

Linking health databases for research

Richard Birtwhistle MD MSc CCFP FCFP Marshall Godwin MD MSc CCFP FCFP Jannet Ann Leggett JD CIPP/C CIPT CIAPP-M Ken Martin MSc CPHIMS

The Canadian Primary Care Sentinel Surveillance Network (CPCSSN) database holds the primary health care records of almost 1 million Canadians. Every health care system interaction for these patients results in data collection and a data set. Although anonymized, these data can be re-identified and the database linked to other provincial or national databases. *Data linkage* temporarily brings 2 or more data sets together to answer a research question or to assess the effectiveness of care delivery, allowing better understanding of health outcomes for chronic diseases and identification of new risk factors for disease. However, it also raises privacy and technical issues.

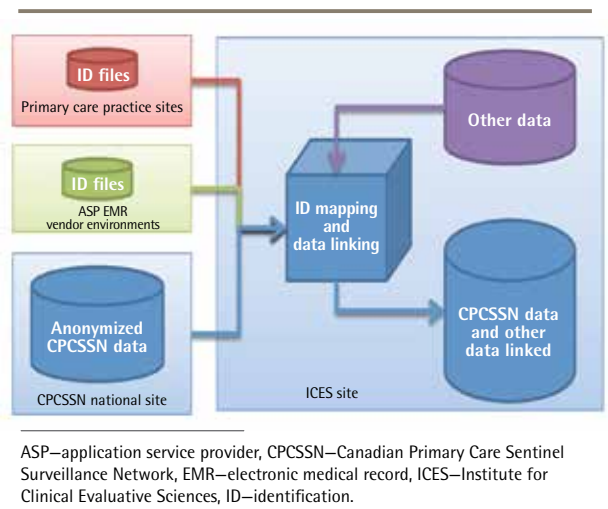
Canada has long maintained databases such as the death registry, hospitalization discharge summaries, disease registries, provincial medical billing databases, census data, and national surveys. Missing from the mix has been what CPCSSN can provide: data from primary care on provider-patient encounters, a list of ongoing health conditions, laboratory and diagnostic imaging investigations, medications, immunizations, in-office procedures, referrals, and more.

By linking CPCSSN's database to mortality and hospitalization databases, we can look at, for example, exposures in primary care (medications, achievement of guideline targets, physician visits) and link them to outcomes, or we can look at primary care provided for ambulatory care-sensitive conditions and how it affects hospitalization outcomes. As genetic databases develop, CPCSSN can link individuals' genetic information with their health conditions over time to better understand the interplay of these factors and the effects on morbidity and mortality. Data linkage will also be useful in postmarketing surveillance of pharmaceuticals. Collecting evidence to associate a new drug with a previously unidentified benefit or adverse effect often takes years; monitoring and linking the CPCSSN database could reduce that time.

Provincial and territorial privacy laws vary but all require patient consent for use of health information beyond the purpose of its original collection. Most permit the use of identifiable health data for research, including data linkage, provided that a research ethics board has approved it; the physician (*data custodian*) has been informed; and patients have been given notice, explanation of their data's use, and a chance to opt out. CPCSSN's fully anonymized data are not readily linkable; CPCSSN has devised a process by which patient identification is only done by a *prescribed entity* (in Ontario, an institution designated secure by the Information

and Privacy Commissioner) with the right to hold identifiable data, such as the Canadian Institute for Health Information or the Institute for Clinical Evaluative Sciences. Each patient has a unique CPCSSN number that can be linked to patient identifiers in the practice or at the electronic medical record vendor application service provider site. These identification key files are then sent separately from the anonymized CPCSSN database to the prescribed entity, using a secure virtual private network Web portal. Patient identifier keys and CPCSSN data are only combined by the prescribed entity and then a common identifier is used to link to other databases (Figure 1).

Figure 1. CPCSSN's anonymized data are linked with other data at secure designated research sites for enhanced research capability



CPCSSN's database has been linked to administrative data in several provincial projects, expanding its utility in the study of health outcomes, social determinants of health, and disease risk factors. CPCSSN looks forward to more such projects, as it safeguards data security and patient and provider privacy. 🍁

Acknowledgment

Funding for this publication was provided by the Public Health Agency of Canada. The views expressed do not necessarily represent the views of the Public Health Agency of Canada.

Dr Birtwhistle is Professor in the Department of Family Medicine and the Department of Community Health and Epidemiology at Queen's University in Kingston, Ont.

Dr Godwin is Professor of Family Medicine and Director of the Primary Healthcare Research Unit in the Faculty of Medicine at Memorial University of Newfoundland in St John's. **Ms Leggett** is Research Privacy and Ethics Officer (CPCSSN Project) at the CFPC in Mississauga, Ont. **Mr Martin** is Information and Technology Manager for CPCSSN.

Competing interests

None declared

La traduction en français de cet article se trouve à www.cfp.ca dans la table des matières du numéro d'avril 2015 à la page e223.

Sentinel Eye is coordinated by CPCSSN, in partnership with the CFPC, to highlight surveillance and research initiatives related to chronic illness prevalence and management in Canada. Please send questions or comments to Dr Richard Birtwhistle, Chair, CPCSSN, at richard.birtwhistle@dfm.queensu.ca.