

## Primary care electronic medical records: a new data source for research in Canada

Richard Birtwhistle MD MSc, Tyler Williamson PhD

Canada was lagging behind in the use of electronic medical records (EMRs) in primary care practice but is now catching up.<sup>1</sup> Although many factors influence the adoption and use of EMRs, there is evidence that they may have a positive influence on patient care at the practice level.<sup>2,3</sup> For clinicians, EMRs provide a new opportunity to assess care given at the practice population level and to improve clinical care. For example, this information helps to understand how care for chronic disease is delivered at the practice population level and reflects on the adequacy of that care. This can result in practice change and improved processes for patients. To do this, practitioners need summary information about their practice population and individual patients (i.e., about care of patients with chronic diseases, how well the disease is controlled and what medications are being used). Providing summary data about the practice population was not considered when EMRs were developed, and they generally lack this capability beyond the most basic queries. At present, practitioners have difficulty getting data back from EMRs. However, several groups in Canada are now extracting reliable clinical data from EMRs for this purpose.<sup>4,5</sup>

As the use of EMR expands, there is an opportunity to use the patient data from EMRs for other purposes. The data from EMRs provide a “laboratory” for primary care researchers and a view of primary care that has never been available in Canada, because patient information was buried in paper charts and only used by the most ardent researcher. Furthermore, by combining EMR data from many practices across the country, it is possible to assess primary care at a regional, provincial and national level.

EMRs are a rich source of data for different types of primary care research, such as longitudinal comparisons over time, cohort studies, assessments of the meaningful use of EMRs and large N clinical trials in primary care. Linking the EMR data with other datasets, such as genomic information, will allow genetic and clinical information to be used in analytics.

EMR data could be used for disease surveillance, in particular for chronic diseases such as diabetes, hypertension, cardiovascular disease and chronic lung disease.<sup>4,6</sup> Opportunities for data analytics and development of predictive models will further the understanding of risk factors for chronic disease. Pharmacosurveillance is another area in which EMR data are being evaluated.<sup>11,12</sup> As near-real-time data extraction comes closer to being a reality, opportunities for acute disease surveillance (e.g., for influenza) become possible.

Finally, EMR data can be used for evaluation and monitoring of health systems. Linking EMR data to administrative data, such as hospital discharges and emergency department visits, will increase the value of both for this purpose. EMR data can be used at the regional and provincial levels for assessing care, identifying gaps in care, and predicting patients who are high-cost users of the system, and developing community interventions that would improve their care and reduce the cost burden.

EMR data are at a similar stage to where administrative data were a few years ago — considered too dirty to provide useful information — but as the quality of administrative data improved and their limitations were understood, they are now seen as a valuable resource for research and evaluation of health systems. However, administrative data, such as admission to hospital, emergency department visits and billings by physicians, have a number of shortcomings. Data for

**Competing interests:** Richard Birtwhistle and Tyler Williamson have received grants from the Public Health Agency of Canada.

This article has been peer reviewed.

**Correspondence to:** Richard Birtwhistle, richard.birtwhistle@dfm.queensu.ca

CMAJ 2015. DOI:10.1503/cmaj.140473

### KEY POINTS

- Data in electronic medical records (EMRs) for primary care are being extracted into secure databases for use in research, disease surveillance and practice quality improvement.
- Regular extraction of EMR data can provide timely and longitudinal information to assess trends in chronic disease and to study risk factors for development of disease.
- EMRs are a source of data, such as blood pressure, weight, laboratory investigations and prescribed medications, that are not currently available in most administrative databases.
- Linking to other databases (e.g., hospital admissions, emergency department and, in time, genetic repositories) has potential for data analytics.

admissions to hospital are not useful for estimating population prevalence of conditions that are most likely to be treated in the community (e.g., asthma and depression). Diagnoses on physician billing are often not accurate and limited to one diagnosis, making multimorbidity difficult to study. Billing patterns can also change depending on the model of care.<sup>8</sup> Other approaches, such as population surveys, are more useful for assessing population prevalence and provide information about participant sociodemographics, physical measurements (e.g., height, weight and blood pressure), but they are expensive to undertake and are not timely.

Electronic medical records provide a new, complimentary source of data that extends what is available from administrative sources or population surveys. The data are both timely and longitudinal, clinically verified, and include health conditions, biometrics (e.g., height, weight and blood pressure), prescribed medications, laboratory tests and risk factors.

However, the usefulness of the data from EMRs is dependent on the structure, extractability and quality of the data. There are three main limitations to the extraction of usable information from EMRs. The first limitation is the input variability of the data, such as differences in where patient information is entered by practitioners (e.g., smoking history may be entered as social history, risk factor or as part of the encounter note). The second limitation is data that are difficult to code or analyze, such as scanned documents or encounter notes.<sup>9</sup> The third limitation is that some of the data needed for population surveillance or research is missing or poorly represented (e.g., important risk factors and modifiers of chronic disease, such as ethnicity, education and income). Solutions to these problems include providing a more structured data entry that has less free text and choosing an EMR capable of communicating with other electronic systems (e.g., in hospitals, laboratories and imaging) to allow direct storage of as much clinical information as possible in the EMR. Currently, much effort must be put into cleaning the data to make them useable. Assessment of data quality is an important ongoing process. Several frameworks have been developed to do this.<sup>10,11</sup> A number of groups are developing case definitions for diseases that can be used to reliably identify patients with those diseases in the EMR. This is an important step forward in ensuring data quality.

A number of EMR databases are being used effectively for research in Europe and elsewhere.

In Canada, EMR databases are emerging and becoming ready for use. The Canadian Primary Care Sentinel Surveillance Network is the first multidisease electronic record surveillance system in Canada.<sup>12</sup> It involves 588 primary care practices and 840 000 patients and continues to grow. The Electronic Medical Record Administrative Data Linked Database (EMRALD) is an EMR database linked to administrative data at the Institute for Clinical Evaluative Sciences in Ontario.<sup>5</sup>

Although there is still work to do to make EMRs a valid and reliable source of data in Canada, they can now provide useful data for primary care research and surveillance, make a valuable contribution to the understanding of chronic disease management in primary care, and help practitioners provide population management and improve the health of Canadians.

## References

1. Rich P. EMR use rates soaring in Canada. *Future Practice*. 2013;5.
2. *The emerging benefits of electronic medical record use in community-based care*. Toronto: Canada Health Infoway; 2013.
3. Holroyd-Leduc JM, Lorenzetti D, Straus SE, et al. The impact of the electronic medical record on structure, process, and outcomes within primary care: a systematic review of the evidence. *J Am Med Inform Assoc* 2011;18:732-7.
4. Birtwhistle RV. Canadian Primary Care Sentinel Surveillance Network: a developing resource for family medicine and public health. *Can Fam Physician* 2011;57:1219-20.
5. Tu K, Mitiku TF, Ivers NM, et al. Evaluation of Electronic Medical Record Administrative Data Linked Database (EMRALD). *Am J Manag Care* 2014;20:e15-21.
6. Klompas M, Murphy M, Lankiewicz J, et al. Harnessing electronic health records for public health surveillance. *Online J Public Health Inform* 2011;3.
7. Stang PE, Ryan PB, Racoosin JA, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med* 2010;153:600-6.
8. Alshammari AM, Hux JE. The impact of non-fee-for-service reimbursement on chronic disease surveillance using administrative data. *Can J Public Health* 2009;100:472-4.
9. Mitiku TF, Tu K. Using data from electronic medical records: theory versus practice. *Healthc Q* 2008;11:23-5.
10. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc* 2013;20:144-51.
11. *The CIHI data quality framework*. Ottawa: Canadian Institutes for Health Information; 2009.
12. Birtwhistle R, Keshavjee K, Lambert-Lanning A, et al. Building a pan-Canadian primary care sentinel surveillance network: initial development and moving forward. *J Am Board Fam Med* 2009;22:412-22.

**Affiliations:** Departments of Family Medicine and Public Health (Birtwhistle), Queen's University, Kingston; Ont.; Department of Community Health Sciences (Williamson), University of Calgary, Calgary, Alta.

**Contributors:** Both of the authors contributed to the conception of the article, interpreted the literature, drafted the article, revised it critically for important intellectual content, approved the final version submitted for publication and agreed to act as guarantors of the work.

**Funding:** The authors received funding from the Public Health Agency of Canada. The views expressed herein do not necessarily represent the views of the Public Health Agency of Canada.