

Editorial

Identifying Atrial Fibrillation From 30,000 Feet: Strengths and Weaknesses of Health Administrative Data

Stephen B. Wilton, MD, MSc, Hongwei Liu, MD, and Danielle A. Southern, MSc

*Libin Cardiovascular Institute of Alberta and O'Brien Institute of Public Health, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada****See article by Hawkins et al., pages 1412–1415 of this issue.***

Health administrative, or "claims" data are collected and stored by government agencies or other payers, primarily for purposes related to their mandates: for instance, to maintain accurate population registers, plan service delivery, and pay physicians. As a result of its universal health care system and unique provincial identifiers allowing linkage of multiple data sets, Canada enjoys a relatively rich health administrative data "ecosystem" compared with many other countries.¹ This system has permitted secondary use of health administrative data for research and quality improvement. Researchers, health authorities, and nonprofit agencies are all naturally interested in using these data, including for cohort definition, comorbidity adjustment, exposure classification, and outcome ascertainment. Compared with primary data collection, using such secondary data sources has the advantages of population-level coverage and typically a much lower (zero in some cases) cost for data collection, increasing the feasibility of many projects. Canadian researchers have used these data to make innumerable important methodological contributions and empirical observations in the fields of epidemiology, health services research, health economics, and randomized clinical trials.²

However, there remain important challenges to the secondary use of Canadian health administrative data for research. These include the large variation in the scope of data collection across provinces, legal barriers to combining inter-provincial data sets, varying degrees of validation for accuracy of the underlying diagnostic codes, and limited granularity of detail compared with clinical records or primary research data. As with any other data source, researchers wishing to use health administrative data sets should become familiar with the modes of data collection and need to understand and report the inherent limitations of the data set.

Of the available health administrative data sets, Canadian researchers have the most experience and confidence with hospital discharge information. Hospital discharge abstracts are coded using consistent standards nationwide and have been reported to the Canadian Institute of Health Information (CIHI) since 1997. No other patient-level data sets are currently available for all Canadian jurisdictions. Accordingly, the majority of publications and CIHI-defined Quality Indicators relate to hospital-based care that can be measured and compared nationally. For example, the 2017 Cardiac Care Quality Indicators Report, coproduced by the Canadian Cardiovascular Society and CIHI, focused on short-term in-hospital mortality and readmission metrics for common cardiovascular procedures.³

However, very few health conditions begin and end with a hospital stay. Like all common cardiovascular conditions, atrial fibrillation (AF) has a preclinical phase, followed by a postdiagnosis course that may include a mix of outpatient, emergency department, and hospital-based care. Relying on hospital-based health administrative data to represent the entire population with AF is suspect, as those using hospital-based services likely have systematically different baseline characteristics and prognoses than those whose care is managed in the outpatient domain. It was partly for this reason that the Canadian Cardiovascular Society Quality Project's AF Working group concluded, in their initial 2016 publication, that none of the 3 top-ranked quality indicators could be adequately measured at a national level with existing data sets.⁴ Thus, using health administrative data to identify those with AF and assess the quality of their care and their outcomes remains a complex challenge.

The brief paper by Hawkins and colleagues in the Health Policy and Promotion section of this issue of the *Canadian Journal of Cardiology* is a welcome empirical demonstration of some of these issues.⁵ In their report, the authors used health administrative data from British Columbia to assess the impact of varying definitions of AF on the population size, the proportion of incident vs prevalent cases, and—importantly—the estimated rate of oral anticoagulation prescription in those with elevated risk for stroke. They used data from hospital discharges (the DAD), ED visits (the National Ambulatory Care Reporting System [NACRS]), and

Received for publication June 3, 2019. Accepted June 18, 2019.

Corresponding author: Dr Stephen B. Wilton, Libin Cardiovascular Institute of Alberta, GE64 3280 Hospital Drive NW, Calgary, Alberta T2N 4Z6, Canada. Tel.: +1-403-210-7102; fax +1-403-210-9180.

E-mail: sbwilton@ucalgary.ca

See page 1290 for disclosure information.

outpatient physician billings (the Medical Services Plan [MSP]) to identify those with AF and their comorbidities and provincial prescription data (Pharmanet) to identify those filling prescriptions for oral anticoagulants. They compared 2 definitions for AF: a "sensitive" definition, requiring a single DAD, NACRS, or MSP entry, vs a more specific definition of 1 or more DAD or NACRS entries or 2 MSP entries 30 to 365 days apart. These definitions were not validated against clinical records but were compared against each other. They also examined the effects of varying the look-back period on identifying incident cases of AF as well as various definitions for excluding potentially secondary AF.

The authors report that the more specific AF definition captured 19% fewer cases, with a predominant loss of patients identified only in ambulatory care. The more specific definition was also associated with an absolute 5% lower rate of oral anticoagulation (53.4% vs 58.7%) among those deemed to be at risk of stroke. When considering all of the permutations of cohort definition for new-onset AF, the rate of oral anticoagulation within 1 year of a first diagnosis ranged from poor (52%) to respectable (72%).

Another interesting aspect of the analysis relates to attempts to exclude transient or secondary AF in these cohorts. Here, they found that, except for diagnoses occurring in the context of cardiac surgery, most cases of so-called transient AF (for example, during hospitalization for myocardial infarction, sepsis, or pulmonary embolism) had preadmission diagnoses of AF that were not captured in ambulatory claims. This finding validates the evolving understanding that most cases of "secondary" AF represent mainly the unmasking of previously undiagnosed AF by a combination of physiologic stress and increased monitoring.⁶ The authors' recommendation to include such cases in AF cohorts derived from health administrative data seems wise.

There are several limitations to this brief report that need to be kept in mind. First, no attempt was made to validate the claims-based diagnoses or treatments against a clinical record. However, a similar recent analysis from Ontario compared various AF definitions based on health administrative data with data extracted from a primary care electronic health record found that the "sensitive" definition used in this study had a sensitivity of 92.2% and specificity of 90.4%, which may be acceptable.⁷ Second, the data come from 1 Canadian province. As noted here, jurisdictional differences in data sets used, and in coding rules within data sets, make interprovincial comparison or generalization problematic.

In the not-too-distant future, efforts by groups, such as the Canadian Institutes of Health Research-funded National Data Platform, to create a national clearinghouse for secondary data

sources—coupled with wider adoption of electronic health records and increased availability of population-based electrocardiographic data—may mitigate some of the current problems inherent in relying on health administrative data for identifying AF and other chronic conditions. However, enriching the available data will also add complexity in cohort definition, so the lessons provided by this paper will remain relevant. As always in research, the devil is in the details.

Despite its limitations, we recommend the article by Hawkins et al. to anyone using or considering the use of health administrative data in research or quality reporting. The results will not come as a surprise to those experienced in the area but rather serve as a quantitative demonstration of the implications of study design choices. This has particular importance when comparing adherence with quality indicators across studies in which the case definition algorithms differ. In their concluding paragraph, the authors remark that, "There is a marked disconnect between the volume of publications in these domains and those examining the underpinning data." We agree and would add that this study provides a useful template for those wishing to test these concepts for other health conditions.

Disclosures

The authors have no conflicts of interest to disclose.

References

1. Quan H, Smith M, Bartlett-Esquilant G, et al. Mining administrative health databases to advance medical science: geographical considerations and untapped potential in Canada. *Can J Cardiol* 2012;28:152-4.
2. Kaczorowski J, Chambers LW, Dolovich L, et al. Improving cardiovascular health at population level: 39 community cluster randomised trial of Cardiovascular Health Awareness Program (CHAP). *BMJ* 2011;342:d442.
3. Canadian Institute of Health Information: Cardiac Care Quality Indicators Report. Ottawa, Ontario, Canada. 2017.
4. Cox JL, Dai S, Gong Y, et al. The development and feasibility assessment of Canadian quality indicators for atrial fibrillation. *Can J Cardiol* 2016;32:1566-9.
5. Hawkins NM, Daniele PR, Humphries KH, et al. Empirical insights when defining the population burden of atrial fibrillation and oral anticoagulation utilization using health administrative data. *Can J Cardiol* 2019;35:1412-5.
6. Verma A. Does "secondary" atrial fibrillation really exist? *J Am Coll Cardiol Clin Electrophysiol* 2018;4:394-6.
7. Tu K, Nieuwlaar R, Cheng SY, et al. Identifying patients with atrial fibrillation in administrative data. *Can J Cardiol* 2016;32:1561-5.