An Introduction to Health Care Administrative Data

Suzanne M Cadarette and Lindsay Wong

INTRODUCTION

Health care administrative data are generated at every encounter with the health care system, whether through a visit to a physician’s office, a diagnostic procedure, an admission to hospital, or receipt of a prescription at a community pharmacy. The terms “health care utilization data”, “administrative health care billing records”, “administrative claims data”, or simply “claims data” are synonymous with “health care administrative data”. These data are collected for administrative or billing purposes, yet may be leveraged to study health care delivery, benefits, harms, and costs. Pharmacists play a key role in the health care system and may be uniquely attuned to identify important pharmacy practice and pharmacotherapy questions that can be answered with health care administrative data. However, before embarking on a new research study, a fundamental understanding of the strengths and limitations of these data for research is imperative. In this primer, we introduce the common types of health care administrative data and how they may be used to understand professional community pharmacy services, drug utilization, and drug safety and effectiveness.

COMMON SOURCES OF HEALTH CARE ADMINISTRATIVE DATA

Canadians are fortunate to have access to publicly funded universal medicare. Each province has its own system to provide remuneration for health care delivery, yet some databases, such as the database for inpatient hospitalizations, are standardized across most provinces and territories in the country. Five of the most common health care administrative databases linked for research purposes in Ontario (summarized in Table 1) cover data for enrolment, drugs, physician services, and hospital services. Identification (ID) variables are anonymized and used for linkage between different databases. For example, in Ontario, the patient ID is an encrypted and scrambled version of the patient’s Ontario Health Insurance Plan number.

Health care administrative data can be supplemented through linkage with other data sources, such as census data to estimate neighbourhood income, clinical registries or electronic medical records to enrich claims data with more clinical data, citizenship data to study immigrants to Canada or Aboriginal peoples, or survey data to add patients’ perspectives and lifestyle variables. It is beyond the scope of this primer to comprehensively review all of the data that may be linked with health care administrative data. Indeed, with funding support and a clear rationale for how the data may help to answer health services or clinical research questions, opportunities for data linkage are endless. Instead, we focus this primer on traditional databases typically used for research.

EXAMPLES OF HEALTH CARE ADMINISTRATIVE DATA RESEARCH

Health care administrative data can be analyzed to understand professional community pharmacy services, drug utilization, and drug effects. A few examples are provided to help pharmacists appreciate some of the different types of pharmacy and pharmacotherapy questions that can be answered with traditional health care administrative data, as well as some of the current challenges in relying on these data for research.

Community Pharmacy Practice Research

Use of health care administrative data to describe professional pharmacy services only recently became possible in Canada with the introduction of remuneration models for professional pharmacy services. In 2013, the Ontario Pharmacy rEsearch collaboratioN (OPEN, www.open-pharmacy-research.ca/) was funded by the Government of Ontario to examine the value of recent and emerging medication management services provided by the province’s pharmacists. The scope of work includes several projects that leverage health care administrative data to describe and examine the delivery of professional pharmacy services. In a recent descriptive analysis, our research group found that 7767 smokers and 1253 pharmacies had participated in the Ontario Pharmacy Smoking Cessation Program within the first 2 years after the program was launched. As depicted in Figure 1, we...
were able to characterize the use of pharmacy smoking cessation services over time and identify peaks and lulls in service delivery. We encourage readers to read the full paper, and focus here on how linkages between different health care administrative databases permitted unique insights. For instance, we documented regional differences in service delivery (where regions were identified from patients’ postal codes in health care enrolment data) and differences in the prevalence of chronic conditions between men and women receiving pharmacy smoking cessation services (from diagnoses coded in physician services and hospital data). We also found that many patients had received another professional pharmacy service within the year before program enrolment (pharmacy data), that over 90% had received a prior smoking cessation service from a physician (physician services data), and that more than two-thirds had received a prescription smoking cessation medication during follow-up (pharmacy data).

Our findings also highlight several missed opportunities for both pharmacies and patients: only an estimated one-third of pharmacies in the province were providing smoking cessation services, only 56% of program enrollees had received a follow-up smoking cessation service after enrollment, and “quit status” (successful, unsuccessful, or unknown) had been reported for only 12% of participants.10 The Pharmacy Smoking Cessation Program has the potential to significantly reduce the burden of tobacco-related morbidity. Our results, which leveraged health care administrative data, point to the need for strategies to help pharmacists take advantage of the program and to improve follow-up services and reporting of whether or not patients have quit smoking.

Some notable limitations at the time of our study related to pharmacy data, specifically, postal code of the pharmacy, pharmacy ID, pharmacy characteristics, and prescriber ID. Although a full postal code is available for each patient, only the first 3 characters of the postal code are available for each pharmacy. This imprecision meant that we had to manually code the location of each pharmacy in estimating the number of pharmacies participating in each region. In addition, although pharmacists have been authorized to prescribe smoking cessation medications in Ontario since 2012, we were not able to identify when pharmacists (as opposed to physicians) were prescribing these medications. In theory, we should have been able to make this distinction through record linkage. However, at the time of our analysis, the prescriber ID field linked primarily to the College of Physicians and Surgeons of Ontario, with no linkage to a pharmacy regulatory body. Finally, pharmacy-specific information such as type (banner, chain, independent) was not

### Table 1. Description of Five Common Administrative Databases in Ontario

<table>
<thead>
<tr>
<th>Database</th>
<th>Description</th>
<th>Key Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered Persons Database (RPDB)</td>
<td>Enrolment data: Includes data on all persons enrolled in the Ontario provincial health care program and a unique encrypted patient identifier that is used for record linkage across all databases</td>
<td>Patient ID, Date of birth, Date of death, Postal code</td>
</tr>
<tr>
<td>Ontario Drug Benefit (ODB) Claims</td>
<td>Pharmacy data: Captures medications dispensed and services billed through the ODB Program: • Drugs listed on the ODB formulary for residents: • ≥ 65 years of age • younger residents on social assistance* • Drugs not listed on formulary but possibly funded through the Exceptional Access Program or Special Access Programme if certain criteria are met • Community pharmacy services (influenza vaccinations, MedsCheck, pharmaceutical opinions, smoking cessation)</td>
<td>Patient ID, Pharmacy ID, Prescriber ID, Date of service, Days supplied, Drug ID number, Dosage form and strength, Quantity dispensed</td>
</tr>
<tr>
<td>Ontario Health Insurance Plan (OHIP) Claims</td>
<td>Physician services: Captured using OHIP codes • Diagnoses (e.g., reason for office visits and diagnoses made) • Procedures (e.g., laboratory and diagnostic tests, vaccinations)</td>
<td>Patient ID, Physician ID, Date of service, Diagnoses, Procedures</td>
</tr>
<tr>
<td>Discharge Abstract Database (DAD)</td>
<td>Hospital inpatient data: Detailed data relating to hospital admissions • Diagnoses (differentiating between the most responsible diagnosis and other diagnoses) • Procedures and interventions (e.g., surgery codes)</td>
<td>Patient ID, Physician ID, Hospital (facility) ID, Date of admission, Date of discharge, Diagnoses, Length of stay, Procedures</td>
</tr>
<tr>
<td>National Ambulatory Care Reporting System (NACRS)</td>
<td>Hospital outpatient data: Detailed data for day surgeries and emergency department services</td>
<td>Patient ID, Physician ID, Hospital (facility) ID, Date of service, Diagnoses, Procedures</td>
</tr>
</tbody>
</table>

*Social assistance programs for residents less than 65 years of age: Ontario Disability Support Program, Ontario Works (employment assistance), Trillium Drug Program (for people whose drug costs are high relative to their income), residence in homes for special care or long-term care, provision of professional home care services.

ID = identification.

≥ 65 years of age
available, and little work has been done to track pharmacies when they receive a new pharmacy ID through change of ownership. Nonetheless, research with health care administrative data can help to gain an understanding of the use and impact of pharmacy services, as well as identify opportunities for improvement. Opportunities for community pharmacy practice research with health care administrative data will improve and broaden as methods are developed and refined, and as data linkages with other pharmacy-specific data sources are created.

**Drug Utilization Research**

Pharmacy claims databases are commonly used to examine drug utilization; to estimate drug exposure or adherence to therapy; and, when linked to medical claims, to study drug effects.

**Prescribing and Dispensing Patterns**

Drug dispensing data can be used to describe drug use and estimate prescribing. Although dispensing data reflect drugs dispensed and not prescriptions written, these data are often used to infer prescribing patterns. As an example, when our team used drug dispensing data to compare incident drug prescribing between provinces, we found significant differences between drug plans. The definition of incident (new), prevalent (current), or past (history) drug use is typically based on a minimum length of time (e.g., a 1-year period) before the first date of drug dispensing that defines study entry. Drug dispensing data can also be used to examine the effect of new drug coverage policies on physicians’ prescribing practices and characterize prescribers (e.g., by age, sex, specialty, and practice region) as well as patients (e.g., by age, sex, drug and health history, and other factors). For example, we recently identified differences in the characteristics of new users of zoledronic acid following a change in drug formulary listing status (e.g., fewer with prior osteoporosis pharmacotherapy). A strength of our study was in linking traditional health care administrative data with physician data to describe physician characteristics; however, physician ID was missing or not linkable for 10% of the prescriptions. We also used interrupted time series analysis to document the significant increase in the numbers of new patients receiving and new physicians prescribing zoledronic acid following the change in drug formulary listing status. Interrupted time series analysis examines aggregate data collected over equally spaced intervals before and after an intervention to estimate and compare drug utilization trends had the intervention not occurred. One of the benefits of time series analysis is in the visual representation of results. For example, Gomes and others include figures in their paper that help illustrate the effect of new narcotics legislation on the prevalence of opioid, benzodiazepine, and stimulant prescriptions that would have triggered warnings of “double-doctoring” and polypharmacy.

Despite the advantages of pharmacy claims data in Canada, drugs dispensed in the hospital setting (emergency department, inpatient wards, same-day surgery clinics) are not captured. Researchers must thus consider gaps in drug information when patients are admitted to hospital. Several approaches can be used...
to adjust for non-measurable drug exposure during a hospital admission, but the validity of these methods is not well described and further research is warranted. Instead, central hospital databases or chart review are currently relied upon to examine drug utilization in the hospital setting. Similarly, although Ontario pharmacy data include drugs dispensed to patients residing in long-term care facilities, pharmacy data in most other provinces do not capture drugs dispensed in this setting. The comprehensiveness of data for drugs dispensed in the community also differs between provinces. For example, the PharmaNet database in British Columbia captures all prescription medications dispensed in the community setting, yet drug data in Ontario typically include only drugs dispensed through the public drug plan. Therefore drugs paid for out-of-pocket or through private insurance are not captured. Efforts are underway to improve data capture, e.g., through the Narcotics Monitoring System in Ontario. As more provinces move toward capturing all drugs for all residents, it will become easier to comprehensively assess and compare drug use and effects.

### Adherence with Therapy

The existence of a pharmacy drug claim does not guarantee consumption of the medication, yet repeated dispensing over regular intervals is often used as a good proxy for adherence to therapy. Adherence is typically quantified with measures of compliance and persistence. Compliance is measured using the medication possession ratio (number of days for which the drug is supplied in an observation period divided by the number of days in the observation period) or, when capped at 100%, the proportion of days covered. Persistence with therapy is defined as the duration of continuous therapy after drug initiation.

Prior research has shown less than 0.5% missing data for days supply or quantity dispensed, and several studies have found high agreement between self-reported use and pill counts or pharmacy claims. However, we recently found errors in the "days supply" field for extended-dose osteoporosis medications. In our study, we compared days supply values for different osteoporosis drug formulations with values expected from the dosing interval. For example, dispensing a single monthly dose of risendronate (one 150-mg pill) should have a days supply value around 30, and an annual infusion of zoledronic acid should have a days supply value of 365. Overall, the days supply values met our "expected value" criteria (as defined a priori) for 90% of osteoporosis drugs dispensed in the community, yet only 59% of those dispensed in long-term care. These differences resulted in the underestimation of adherence to oral bisphosphonate therapy, particularly in long-term care. Inaccurate reporting could result from practical strategies to avoid claim rejections, such as underestimating the days supply slightly to allow a patient to pick up his or her next prescription at a convenient time, before the current medication supply is complete; alternatively, it could result from uncertainty about how to enter the days supply for eye drops or nasal spray, or how to enter a vial of zoledronic acid (annual infusion) when 100 days is the maximum that can be submitted through the public drug plan. When estimating drug persistence, researchers typically include a grace period between dispensing of prescriptions. Nonetheless, the potential for error remains, and data cleaning, in addition to a grace period, is encouraged. We were comforted to find little difference in estimates of adherence after cleaning the data from the community setting, and also believe that our findings of concern are largely related to extended-dose drugs. As other medications with extended or fixed dosing become more common (e.g., oral contraceptives, antipsychotics, and antiretroviral medications), researchers will need to remain mindful of possible errors in data entry when they are estimating drug exposure. Our results do not discount the tremendous opportunities that pharmacy claims data provide for understanding drug adherence and effects, but rather point to the need to examine and possibly “clean” data before estimating drug exposure.

### Drug Safety and Effectiveness

In an earlier Research Primer in this series, covering cohort and case–control studies, Gamble introduced the fundamentals of these key observational study designs, as well as their strengths and limitations for estimating drug safety and effectiveness. The proper application of observational study methodology is imperative. Here, we focus on the importance of local knowledge of the data and drug policies to guide the analysis of health care administrative data in drug safety and effectiveness research.

Drug safety and effectiveness research based on health care administrative data must be interpreted in the context of changes in drug availability over time. Entry of new drugs to the market and changes to reimbursement criteria or coverage can significantly affect drug utilization and patient characteristics. As described in the primer on cohort and case–control studies, “confounding by indication” is natural in any study that examines the effects of pharmacotherapy, since drugs are always prescribed (indicated) to mitigate risk, i.e., to prevent or treat a condition and thus reduce the risk of harm. Confounding is not a problem in the context of complete information, since researchers can use statistical methods to adjust for differences between exposure groups under comparison. However, health care administrative data are often limited in clinical detail, and thus residual confounding (bias left over after adjustment for all of the information measured) is more common. For example, information on height, weight, smoking status, and level of physical activity are not available, yet may be important to consider. Confounding by indication can be particularly problematic when comparing the benefits of 2 or more agents for the same indication. In a recent analysis of claims data from British Columbia and Ontario, we identified policy-induced selection bias in Ontario due to the limited coverage of second-generation bisphosphonates for patients at higher risk of fracture on the basis of age, fracture history, and bone mineral density. Although we could adjust for variables available in health care administrative data,
such as age and fracture history, we were limited in our ability to adjust for bone mineral density. Our prior validation work linking clinical records (dual-energy X-ray absorptiometry [DXA] records, and thus measurement of bone mineral density) to health care administrative data showed excellent validity for identifying patients who had undergone DXA, yet the ability of a medical claim diagnosis of osteoporosis to indicate DXA-defined osteoporosis was poor.24 Residual confounding would thus remain. Intimate, detailed knowledge of local access to particular drugs, as well as data accuracy, is critical to inform studies of drug safety and effectiveness. Indeed, the Canadian Network for Observational Drug Effect Studies brings together local experts from multiple provinces to inform and interpret data analyses.28

Similarly, it is important to consider changes in diagnostics, medical practice, and billing codes (e.g., diagnostic and procedure codes in the 9th and 10th revisions of the International Statistical Classification of Diseases and Related Health Problems) and their validity for outcomes research. Health care administrative data are not specific enough to examine all outcomes. Hard clinical outcomes, such as hip fracture, that typically require hospital admissions and surgery are well captured,29,30 yet outcomes that are clinically less specific or that may be easily misclassified can be problematic. For example, complications that occur during a hospital stay may be incorrectly classified as comorbidities.12,31 Validation studies that compare the ability of claims data to capture true events are thus important.

FEASIBILITY

Studies that leverage health care administrative data are efficient, because the data are collected routinely and thus (in theory) are more readily available than primary data collected through chart review or patient surveys. However, access to these data varies across Canada, and it is typically necessary to work with a research scientist.32 In addition, once data exist, delays in access for research purposes are common. In the context of preapproved rapid-response projects, data access may be as quick as a few days.28 At the other extreme, it may take several years to receive data, making analysis for some projects infeasible. The good news is that data access has been improving over time. In fact, health care administrative data in Ontario are becoming more widely accessible as a result of strategic provincial and federal funding.33

ROLE OF PHARMACISTS IN CREATION OF DATA

As health care providers, pharmacists contribute to the generation of health care administrative data, whether working at a hospital, family health team clinic, long-term care facility, or community health centre; or providing home care or community pharmacy services. Pharmacists are gatekeepers to drug distribution and, as such, are involved in collecting patient information, as well as ordering and dispensing medications. With their expanded scope of practice, pharmacists may provide billable services and submit claims for remuneration. Although data are collected primarily for billing purposes, there may be significant implications for research purposes if the data collected are not accurate. Results from health care administrative data studies may be used to inform clinical practice and health policy decisions. Therefore, pharmacists are encouraged to be mindful of their role in collecting patient, medication, and pharmacy service data, and their significant contributions to research in their daily practice.

CONCLUSIONS

Health care administrative databases are rich sources of information that may be leveraged for research purposes. However, it is important to understand the limitations associated with their use for research. Despite these limitations, the many strengths permit evidence-informed clinical and policy decision-making. As health care providers, pharmacists have the opportunity to contribute to research in their daily practice by accurately collecting patient, medication, and pharmacy service data. Pharmacists interested in becoming involved in research with health care administrative databases are encouraged to seek guidance from a scientist with expertise utilizing these data. Many opportunities exist for hospital and community pharmacists to contribute to research and to help answer important health services and drug therapy–related questions.

References


Suzanne M Cadarette, PhD, is Associate Professor with the Leslie Dan Faculty of Pharmacy, University of Toronto, and Senior Adjunct Scientist with the Institute for Clinical Evaluative Sciences, Toronto, Ontario.

Lindsay Wong, BScPhm, PharmD, was, at the time of writing, a student in the Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario. She is currently a pharmacy intern at St Michael's Hospital, Toronto, Ontario.

Competing interests: None declared.

Address correspondence to:
Dr Suzanne M Cadarette
Leslie Dan Faculty of Pharmacy
University of Toronto
144 College Street
Toronto ON M5S 3M2
e-mail: s.cadarette@utoronto.ca

This article is the 11th in the CJHP Research Primer Series, an initiative of the CJHP Editorial Board and the CSHP Committee. The planned 2-year series is designed to appeal to relatively inexperienced researchers, with the goal of building research capacity among practising pharmacists. The articles, presenting simple but rigorous guidance to encourage and support novice researchers, are being solicited from authors with appropriate expertise.

Previous articles in this series:


