

# Estimating the completeness of physician billing claims for diabetes case ascertainment using population-based prescription drug data

L. M. Lix, PhD (1); J. P. Kuwornu, MSc (1); K. Kroeker, BSc (1); G. Kephart, PhD (2); K. C. Sikdar, PhD (3); M. Smith, MSc (4); H. Quan, PhD (3, 5)

This article has been peer reviewed.

 [Tweet this article](#)

## Abstract

**Introduction:** Changes in physician reimbursement policies may hinder the collection of billing claims in administrative data; this can result in biased estimates of disease prevalence and incidence. However, the magnitude of data loss is largely unknown. The purpose of this study was to estimate completeness of capture of disease cases for Manitoba physicians paid by fee-for-service (FFS) and non-fee-for-service (NFFS) methods.

**Methods:** Manitoba's administrative data were used to identify a cohort ( $\geq 20$  years) with a new diabetes medication between 1 April, 2007, and 31 March, 2009. Cohort members were classified by payment method of the prescribing physician (i.e. FFS vs. NFFS). The cohort was then classified as missing or not missing a diabetes diagnosis using physician claims and hospital records. Then,  $\chi^2$  statistics were used to test for differences in the characteristics of the two groups.

**Results:** The cohort consisted of 12 394 individuals; 86.4% had a prescription for a diabetes medication from an FFS physician. A total of 1172 physicians (81.8% FFS) prescribed these medications for the cohort. Cohort members with a prescription from an FFS physician were older and more likely to reside in the urban Winnipeg health region than those with a prescription from a NFFS physician. A greater percentage of NFFS physicians' cases were missing a diabetes diagnosis (18.7% vs. 14.9% for FFS physicians).

**Conclusion:** The results suggest minimal loss of physician claims associated with remuneration policies in Manitoba. This method of assessing data completeness could be applied to other chronic diseases and jurisdictions to estimate completeness.

**Keywords:** *chronic disease, medical records, surveillance, data quality*

## Introduction

Electronic administrative health data are used extensively in Canada and internationally to conduct population-based chronic disease research and surveillance.<sup>1,2</sup> These data are popular for several reasons: they are timely; they contain information about

numerous individuals; and they are relatively inexpensive to access and use. Electronic physician claims data, which capture billing records or claims for outpatient physician contacts with patients, are a rich resource for research and surveillance because the vast majority of people with a chronic disease will maintain regular con-

## Highlights

- Physician claims data are useful for research into and surveillance of the health of Canadians.
- More doctors are being paid in different ways, and many are moving from fee-for-service (FFS) payment to non-fee-for-service (NFFS) payment. Physician claims data may miss information on patient contacts from physicians receiving NFFS payments.
- We used population-based diabetes prescription drug records to estimate the completeness of physician claims in Manitoba.
- 14.9% of people with a diabetes prescription from an FFS doctor did not have a diabetes diagnosis for 2 years before and 2 years after getting the prescription.
- 18.7% of people with a diabetes prescription from a NFFS physician did not have a diabetes diagnosis for 2 years before and 2 years after getting the prescription.
- The loss of data due to missed claims is probably small. However, this loss of data means disease incidence is underestimated.

tact with a physician for disease management and treatment.

However, the quality of administrative health data available for research and

## Author references:

1. Department of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada
2. Department of Community Health and Epidemiology, Dalhousie University, Halifax, Nova Scotia, Canada
3. Libin Cardiovascular Institute of Alberta, University of Calgary, Calgary, Alberta, Canada
4. Manitoba Centre for Health Policy, University of Manitoba, Winnipeg, Manitoba, Canada
5. Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada

**Correspondence:** Lisa M. Lix, Department of Community Health Sciences, S113-750 Bannatyne Avenue, University of Manitoba, Winnipeg, MB R3E 0W3; Tel: 204-789-3573; Fax: 204-789-3905; Email: [lisa.lix@med.umanitoba.ca](mailto:lisa.lix@med.umanitoba.ca)

surveillance has been questioned, resulting in a number of studies on this topic. Most of these studies have examined the validity of diagnostic information<sup>3-6</sup> even though multiple aspects of quality must be assessed to ensure the data are fit for research and surveillance purposes. For example, complete capture of the population is essential to ensure unbiased research and surveillance results. Gaps in completeness can result in estimation problems similar to those encountered with non-response bias in survey data.<sup>7</sup>

The information in physician claims data is used to remunerate physicians on a fee-for-service (FFS) basis for the services they provide to patients. Non-FFS (NFFS) physicians, who are paid by salary or contract, often use a shadow-billing claims submission process in which they submit parallel claims. However, this process does not always consistently capture information on service encounters. A 2009 study reported underestimation of diabetes incidence due to incomplete capture of shadow-billed claims in Ontario.<sup>8</sup> However, it is not known how widespread this problem is. As national and international health care systems increasingly adopt new models for primary care service delivery and move physicians onto alternate payment plans, the comprehensive capture of physician encounter data is threatened.<sup>9</sup>

The purpose of our study was to estimate completeness of capture of chronic disease cases associated with missed NFFS physician shadow-billed claims in administrative health data. The specific focus of this study is diabetes because diagnosis codes in administrative data have demonstrated good sensitivity and specificity for diabetes case ascertainment and diabetes surveillance using electronic administrative data is of interest worldwide.<sup>5</sup>

## Method

### Data sources

We used data from the province of Manitoba, which has a population of approximately 1.2 million according to the 2011 Statistics Canada Census. The Manitoba Centre for Health Policy houses multiple

electronic administrative databases that can be anonymously linked via a unique personal health identification number. Manitoba, like other Canadian provinces, has a universal health care system, which means that virtually the entire provincial population is registered to receive health care benefits.

Physicians paid via the FFS method submit all billing claims to the provincial ministry of health. Physicians paid by NFFS methods are required to submit parallel shadow-billed claims to the ministry of health, but the extent of compliance with this requirement is unknown given that the remuneration of NFFS physicians does not depend on the submitted claims and the shadow-billed claims are not normally audited. A single diagnosis is recorded on each claim using *International Classification of Diseases, 9<sup>th</sup> revision, Clinical Modification* (ICD-9-CM) codes. Hospital abstracts, which are completed when a patient is discharged from an acute care facility, contain diagnoses recorded using the *International Classification of Diseases, 10<sup>th</sup> revision, Canadian version* (ICD-10-CA) codes. Each record captures up to 25 diagnoses. Prescription drug records are from the Drug Programs Information Network (DPIN), a centralized, electronic, point-of-sale database connecting all retail pharmacies in Manitoba. For each dispensation, information is collected on the date, drug attributes such as dosage and drug identification number, and the prescriber identification number. The population registry contains dates of health insurance coverage as well as demographic information such as date of birth, sex and location of residence. The provider registry contains quarterly snapshots of physician characteristics, including date of birth, specialty, practice location, provider identification number and method of payment. The most recent update to the registry that we used in our study was the first quarter (i.e. January to March) of 2009.

Manitoba's administrative health data have been extensively used in chronic disease research and surveillance studies.<sup>10-12</sup>

Ethics approval was received from the Manitoba Health Research Ethics Board and permission to access the study data

was provided by the Manitoba Health Information Privacy Committee.

### Study cohort

We identified a cohort of incident adult diabetes cases from prescription drug records, which have demonstrated excellent sensitivity for diabetes case ascertainment.<sup>13</sup> The cohort inclusion criteria were (1) at least one prescription with an Anatomical Therapeutic Chemical (ATC) code of A10 (i.e. drugs for diabetes) in the two-year observation period from fiscal year 2007/08 to 2008/09 (a fiscal year extends from 1 April to 31 March); (2) continuous health insurance coverage during the two-year period before and two-year period after the index prescription date, that is the date that a diabetes prescription medication was first identified in the DPIN database during the observation period; and (3) at least 20 years of age at the index prescription date. Individuals were excluded from the study if they had a prescription with an ATC code of A10 within the 730 days prior to their index prescription date.

The ATC code A10 captures blood glucose-lowering drugs such as metformin and insulins and their analogues but not supplies such as glucose test strips. A 2009 study used a one-year period to exclude non-incident cases,<sup>8</sup> but to increase the likelihood that the cohort members were newly prescribed with a diabetes medication we extended this to two years.

The prescriber identification number associated with the index diabetes medication prescription was anonymously linked to the corresponding number in the provider registry to determine physician payment method (i.e. FFS, NFFS). Individuals were excluded if the payment method of the provider who made the index prescription was not recorded in the registry. The cohort members were stratified into two mutually exclusive groups: (1) individuals with an index prescription from an FFS physician, and (2) individuals with an index prescription from a NFFS physician.

### Study variables

Diabetes diagnoses in administrative data were identified using the Canadian Chronic

Disease Surveillance System (CCDSS) diabetes case definition, which requires one hospitalization or two physician billing claims with an ICD-9 code of 250.00 or ICD-10-CA code of E10-E14 within a two-year period.<sup>4,14</sup> The diabetes diagnosis date was the date on which a diagnosis was recorded in hospital records or the date of the first diagnosis in physician billing claims. Cohort members who met the case definition requirements anywhere in the period that extended from two years before the index prescription date to two years after the index prescription date were classified as captured cases. All other cohort members were classified as missed cases.

The cohort was characterized in terms of its sociodemographic characteristics, including age group, sex, income quintile, and health region of residence. Age group was categorized as less than 65 years and 65 years and above. Income quintile was an area-level measure based on average household income from the Statistics Canada Census. Every individual's postal code extracted from the population registry was assigned to a dissemination area (DA), the smallest geographic unit for which census data are reported. The Manitoba population was then divided into five roughly equal groups according to the DA average household income;<sup>15</sup> this was done separately for urban and rural populations. Health region of residence was classified as Winnipeg and non-Winnipeg. The latter encompasses four rural health regions: Interlake-Eastern, Northern, Prairie Mountain, and Southern Health-Santé Sud. First Nations reserves are primarily located outside of the Winnipeg health region. All sociodemographic characteristics were measured at the index prescription date.

Because patients with more advanced diabetes (i.e. greater comorbidity) were more likely to be detected by the CCDSS case definition,<sup>8</sup> we investigated the prevalence of selected cardiovascular comorbid conditions in the cohort using diagnoses in hospital discharge abstracts. These conditions included hypertension (ICD-10-CA I10-I13, I15), congestive heart failure (ICD-10-CA I50), acute coronary syndromes (ICD-10-CA I21, I22, I23, I24.9, I20.0, I20.1), cerebrovascular disease (i.e. stroke: ICD-10-CA I60, I61, I63, I64) and atrial fibrillation (ICD-10-CA I48). However, given that the frequencies were quite small for

some conditions, we classified them as "hypertension only" and "all cardiovascular comorbid conditions."

The physicians who prescribed diabetes medications for study cohort members were characterized by sex, age group (< 35 years, 35–60 years, or ≥ 61 years), health region of practice (Winnipeg vs non-Winnipeg), and specialty (specialist vs general practitioner). All physician characteristics were measured at the index prescription date.

### Statistical analyses

We described the study cohort and their prescribing physicians using frequencies and percentages. A  $\chi^2$  statistic was used to test for differences in characteristics between the FFS and NFFS groups. We estimated the percentage of cohort members with a missing diabetes diagnosis; this was done for the entire cohort, as well as stratified by age group.

We estimated the crude diabetes incidence rate for the adult (≥ 20 years) Manitoba population in the period from 2007/8 to 2008/09 using the CCDSS case definition. The total adult registered population from the Manitoba population registry comprised the denominator. We calculated the observed rate, as well as rates adjusted for missed cases.

All analyses were conducted using SAS version 9.3 (SAS Institute Inc., Cary, NC, USA).<sup>16</sup>

## Results

A total of 73 719 individuals received a diabetes prescription in 2007/08 or 2008/09. After we applied our exclusion criteria (i.e. no continuous provincial health insurance coverage [8.0%]; diabetes medication use in the 730 days prior to the index prescription date [74.3 %]; < 20 years of age at the index date [0.4%]; index prescription written by a physician whose payment method could not be ascertained [0.5%]), 12 394 individuals were retained in the cohort.

The majority (86.4%) of study cohort members received their index prescription from an FFS physician (Table 1). In general,

individuals who received their index prescription from FFS physicians were more likely than those who received their index prescription from NFFS physicians to be older, Winnipeg residents, and from the lowest urban income quintile (all  $p < .001$ ).

A total of 1172 physicians prescribed diabetes medications to at least one cohort member during the observation period (Table 2). On average, each FFS physician prescribed to 11.2 cohort members and each NFFS physician prescribed to 7.9 cohort members. Compared to the NFFS physicians, FFS physicians were more likely to be older, specialists, and to practice in Winnipeg (all  $p < .001$ ). There was no difference between FFS and NFFS physicians on sex ( $p = .086$ ).

A higher percentage of cohort members with a missed diagnosis received their index prescription from a NFFS physician than those who received their prescription from an FFS physician (18.7% vs. 14.9%;  $p < .001$ ; Table 3). When the analysis was stratified by age, a higher percentage of cohort members aged less than 65 years with a missed diagnosis were observed for the NFFS cohort than the FFS cohort (20.4% vs. 16.5%;  $p < .001$ ). The percentages for those individuals aged 65 years or more were not significantly different ( $p = .174$ ).

Cohort members with a missed diagnosis who had their index prescription from a NFFS physician were more likely to be male, non-Winnipeg residents, and in the lowest rural income quintiles than those who received their prescription from an FFS physician ( $p < .001$ ; Table 4).

Of the cohort members who had their index prescription from an FFS physician, those whose diagnosis was captured in administrative data were more likely to be hospitalized for cardiovascular comorbid diseases than those whose diagnosis was missed ( $p < .001$ ; Table 5). Similarly, those of the NFFS cohort whose diagnosis was captured were more likely to be hospitalized for cardiovascular comorbid diseases than those whose diagnosis was missed ( $p < .001$ ).

Based on the CCDSS case definition for diabetes, 12 877 newly diagnosed adult (≥ 20 years) cases of diabetes were identified

**TABLE 1**  
**Characteristics of the study cohort, by method of remuneration for the prescribing physician**

Characteristics	Cohort		
	FFS physicians, n (%)	NFFS physicians, n (%)	All physicians, n (%)
Total <sup>a</sup>	10 714 (86.4)	1680 (13.6)	12 394 (100.0)
Age group, years**			
20–64	7733 (72.2)	1305 (77.7)	9038 (72.9)
≥ 65	2981 (27.8)	375 (22.3)	3356 (27.1)
Sex*			
Male	5468 (51.0)	907 (54.0)	6375 (51.4)
Female	5246 (49.0)	773 (46.0)	6019 (48.6)
Health region**			
Winnipeg	6403 (59.8)	339 (20.2)	6742 (54.4)
Non-Winnipeg	4311 (40.2)	1341 (79.8)	5652 (45.6)
Income quintile**			
Urban Q1/Q2 (lowest)	3079 (28.7)	191 (11.4)	3270 (26.4)
Urban Q3	1206 (11.3)	58 (3.5)	1264 (10.2)
Urban Q4/Q5 (highest)	2136 (19.9)	76 (4.5)	2212 (17.8)
Rural Q1/Q2 (lowest)	1636 (15.3)	754 (44.9)	2390 (19.3)
Rural Q3	785 (7.3)	192 (11.4)	977 (7.9)
Rural Q4/Q5 (highest)	1085 (10.1)	347 (20.7)	1432 (11.6)
Missing	787 (7.4)	62 (3.6)	849 (6.8)

**Abbreviations:** FFS, fee-for-service; NFFS, non-fee-for-service.

<sup>a</sup> Percentages for this row are based on the row total; all other tabled percentages are based on column totals for each variable.

\**p* < .05.

\*\**p* < .001.

**TABLE 2**  
**Characteristics of physicians who prescribed diabetes medications to the study cohort, by method of remuneration**

Characteristics	Cohort		
	FFS physicians, n (%)	NFFS physicians, n (%)	All physicians, n (%)
Total <sup>a</sup>	959 (81.8)	213 (18.2)	1172 (100.0)
Age group, years**			
< 35	124 (12.9)	38 (17.8)	162 (13.8)
35–60	689 (71.9)	167 (78.4)	856 (73.0)
≥ 61	146 (15.2)	8 (3.8)	154 (13.1)
Sex			
Male	664 (69.2)	134 (62.9)	798 (68.1)
Female	295 (30.8)	79 (37.1)	374 (31.9)
Specialty**			
Specialist	211 (22.0)	16 (7.5)	227 (19.4)
General practitioner	727 (75.8)	187 (87.8)	914 (78.0)
Missing	21 (2.2)	10 (4.7)	31 (2.6)
Health region**			
Winnipeg	645 (67.3)	76 (35.7)	721 (61.5)
Non-Winnipeg	314 (32.7)	137 (64.3)	451 (38.5)

**Abbreviations:** FFS, fee-for-service; NFFS, non-fee-for-service.

**Note:** A variable without asterisks indicates no statistically significant difference between the FFS and NFFS groups.

<sup>a</sup> Percentages for this row are based on the row total; all other tabled percentages are based on column totals for each variable.

\*\**p* < .001.

in Manitoba during fiscal years 2007/08 to 2008/09. This resulted in a crude observed incidence rate of 1.6 % (Figure 1). However, when this observed rate was adjusted for missed cases from NFFS physicians, the rate increased to 1.7%, and when it was further adjusted for missed cases from both FFS and NFFS physicians, it increased to 1.9%.

## Discussion

Overall, compared to FFS physicians, fewer NFFS physicians prescribed diabetes medications in Manitoba during the study time period, and the two groups of physicians differed in age, specialty and health region of practice. A greater percentage of NFFS physicians practised in non-Winnipeg health regions than in the Winnipeg health region, probably because provincial programs providing physicians with alternative payment options to encourage them to work in rural and remote parts of the province.

When we used prescription drug records as the reference data source, we found incomplete diabetes diagnosis information for cohort members seen by both FFS and NFFS physicians, although NFFS physicians missed a greater percentage of diagnoses. Potential sources of missing diagnostic information include misclassification bias due to less-than-perfect sensitivity of diagnosis information and missing data bias due to few NFFS physicians shadow billing. Assuming that the sensitivity of diabetes diagnoses is the same for both FFS and NFFS physicians, the results suggest that 3.8% of physician billing claims were missing in Manitoba due to a lack of shadow billing. There were differences in the percentage of missed cases by age group for both FFS and NFFS physicians, suggesting that the sensitivity of ascertainment is different for younger and older cohort members.

An earlier study conducted using Ontario's physician claims data,<sup>8</sup> reported that 23.7% of diabetes diagnoses were missing from NFFS physician billing claims; we found 18.7% missing diagnoses in the NFFS physician billing claims in Manitoba. The Ontario study focussed exclusively on examining diagnoses missing by NFFS physicians,

**TABLE 3**  
The study cohort with a captured and missed diabetes diagnosis by age group and method of remuneration of the prescribing physician

Age group	Cohort	
	FFS physician, n (%)	NFFS physician, n (%)
20–64 years**		
Captured diagnosis	6458 (83.5)	1039 (79.6)
Missed diagnosis	1275 (16.5)	266 (20.4)
≥ 65 years		
Captured diagnosis	2661 (89.3)	326 (86.9)
Missed diagnosis	320 (10.7)	49 (13.1)
All ages**		
Captured diagnosis	9119 (85.1)	1365 (81.3)
Missed diagnosis	1595 (14.9)	315 (18.7)

**Abbreviations:** FFS, fee-for-service; NFFS, non-fee-for-service.

**Notes:** Cases were classified as captured or missed based on the presence/absence of a diabetes diagnosis in physician billing claims and hospital discharge abstracts. Age group without asterisks indicates no statistically significant difference between the FFS and NFFS groups.

\*\* $p < .001$ .

**TABLE 4**  
Characteristics of cohort members with a missed diabetes diagnosis in administrative data by method of remuneration of the prescribing physician

Characteristics	Cohort		
	FFS physicians, n (%)	NFFS physicians, n (%)	All physicians, n (%)
Total <sup>a</sup>	1595 (83.5)	315 (16.5)	1910 (100.0)
Age group, years			
20–64	1275 (79.9)	266 (84.4)	1541 (81.0)
≥ 65	320 (20.1)	49 (15.6)	369 (19.0)
Sex**			
Male	444 (27.8)	146 (46.3)	590 (30.9)
Female	1151 (72.2)	169 (53.7)	1320 (69.1)
Health region**			
Winnipeg	847 (53.1)	56 (17.8)	903 (47.3)
Non-Winnipeg	748 (46.9)	259 (82.2)	1007 (52.7)
Income quintile**			
Urban Q1/Q2 (lowest)	411 (25.8)	34 (10.8)	445 (23.3)
Urban Q3	174 (10.9)	10 (3.2)	184 (9.6)
Urban Q4/Q5 (highest)	269 (16.9)	7 (2.2)	276 (14.5)
Rural Q1/Q2 (lowest)	330 (20.7)	165 (52.4)	495 (25.9)
Rural Q3	112 (7.0)	37 (11.7)	149 (7.8)
Rural Q4/Q5 (highest)	176 (11.0)	53 (16.8)	229 (12.0)
Missing	123 (7.7)	9 (2.9)	132 (6.9)

**Abbreviations:** FFS, fee-for-service; NFFS, non-fee-for-service.

**Note:** A variable without asterisks indicates no statistically significant difference between FFS and NFFS groups.

<sup>a</sup> Percentages for this row are based on the row total; all other tabled percentages are based on column totals for each variable.

\* $p < .05$ .

\*\* $p < .001$ .

while our study compared FFS and NFFS physicians. Inclusion of FFS physicians allows for investigation about the effects of shadow billing on the proportion of missed diabetes diagnosis information in physician billing claims. The Ontario study also used varying follow-up periods to capture diabetes diagnosis before and after the diabetes prescription date, up to a maximum of nine years. In addition, unlike the Ontario study, which focussed only on the older population (i.e. ≥ 65 years), we included all individuals 20 years and older and investigated missing data in both younger and older age groups. We found a higher percentage of missed diagnoses in the younger age group than the older age group, an expected finding because older people are more likely to have a higher disease burden and therefore be hospitalized or have regular contact with a primary care physician.

Given that this study has demonstrated data loss associated with a lack of shadow billing by NFFS physicians, it is important to consider strategies for adjusting prevalence and incidence estimates for possible underestimation. One strategy is to use other population-based data such as electronic medical records, which are increasingly being adopted in population-based chronic disease research and surveillance studies,<sup>17</sup> and prescription drugs data sources to supplement physician claims records for disease surveillance. For example, based on the CCDSS case definition, we estimated a 1.6% crude diabetes incidence rate in the Manitoba population aged 20 years and over during the study period. However, when cases identified in the prescription drugs data were used to adjust for underestimation, the incidence rate increased to 1.9%.

Alternatively, simulation studies and predictive modelling could be used to produce adjusted estimates of disease prevalence, as has been done in previous research.<sup>18,19</sup> Specifically, we found that both patient and physician characteristics differed significantly between the FFS and NFFS groups. Also, we found that the presence of comorbid cardiovascular conditions was associated with whether a patient's diabetes diagnosis was captured by the CCDSS case definition. These factors could

**TABLE 5**  
**Selected comorbid conditions in individuals with a captured and missed diabetes diagnosis**

Comorbid condition	Cohort	
	Captured diagnosis	Missed diagnosis
	FFS physicians, n (%)	
Hypertension**	662 (7.3)	75 (4.7)
Cardiovascular disease <sup>a,**</sup>	839 (9.2)	94 (5.9)
	NFFS physicians, n (%)	
Hypertension**	100 (7.3)	6 (1.9)
Cardiovascular disease <sup>a,**</sup>	122 (8.9)	9 (2.9)

**Abbreviations:** FFS, fee-for-service; NFFS, non-fee-for-service.

<sup>a</sup> Cardiovascular disease included congestive heart failure, hypertension, acute coronary syndromes, cerebrovascular disease (stroke) and atrial fibrillation.

\*\* $p < .001$ .

be included in predictive models to estimate disease prevalence. In a recent study,<sup>20</sup> we included physician characteristics in predictive models to adjust for underestimation of diabetes prevalence due to a loss of physician claims.

### Strengths and limitations

Our study has some limitations. First, physicians were classified as either FFS or NFFS, but some physicians may receive both types of remunerations at the same time or change from one method to another. Given that we only used two fiscal years of diabetes prescription information, the possibility of physicians switching payment method during the study period is likely to be minimal, but the Manitoba provider registry does not allow us to make a clear

distinction between physicians who may be receiving both types of remuneration.

Second, the results may be sensitive to the definitions used to ascertain missed and non-missed cases. We examined the two-year period pre- or post-index prescription date; this period was chosen to align with the CCDSS diabetes case definition from administrative health data. While it is possible that individuals without diabetes might receive a prescription for a diabetes drug, this is unlikely; previous research has shown excellent specificity and sensitivity of prescription drug records for diabetes case ascertainment.<sup>13</sup> Lastly, diabetes patients treated with lifestyle modification only would not be captured in prescription drug data. However, this represents a general problem of underestimation of any chronic

condition treated with only lifestyle modification when using administrative data to estimate prevalence or incidence.

## Conclusions

In this study, we adopted a population-based approach to assess the completeness of physician claims data for chronic disease studies. We relied on prescription drug data to evaluate completeness; this source is known to be sensitive for ascertaining diabetes cases.<sup>13</sup> Our study showed that the loss of data due to missed shadow-billed claims is small. At the same time, this loss of data contributes to an underestimation of disease incidence. The method we used can be readily applied over time and to data from other provinces or territories. The results obtained for diabetes could be compared to the results obtained for other chronic conditions to assess consistency of conclusions about data completeness.

## Acknowledgements

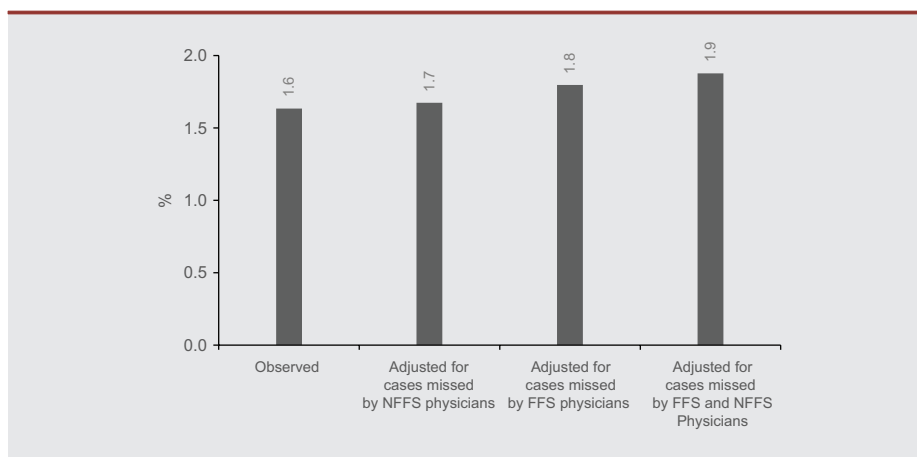
The authors acknowledge the Manitoba Centre for Health Policy for use of data contained in the Population Health Research Data Repository (HIPC project #2012/2013-04). The results and conclusions are those of the authors and no official endorsement by the Manitoba Centre for Health Policy, Manitoba Health, or other data providers is intended or should be inferred.

This research was funded by the Canadian Institutes of Health Research (Funding Reference #123357). Lisa M. Lix is supported by a Research Chair from Research Manitoba. The funders had no involvement in the conduct of the research or in manuscript preparation.

## References

1. Virnig BA, Mcbean M. Administrative data for public health surveillance and planning. *Annu Rev Public Health*. 2001;22:213-30.
2. Dombkowski KJ, Wasilevich EA, Lyon-Callo S, Nguyen TQ, Medvesky MG, Lee MA. Asthma surveillance using Medicaid administrative data: a call for a national framework. *J Public Health Manag Pract*. 2009; 15(6):485-93. DOI: 10.1097/PHH.0b013e3181a8c334.

**FIGURE 1**  
**Crude diabetes incidence (%) for the Manitoba adult (20+ years) population, 2007/08–2008/09**



**Abbreviations:** FFS, fee-for-service; NFFS, non-fee-for-service.

3. Quan H, Khan N, Hemmelgarn BR, et al. Validation of a case definition to define hypertension using administrative data. *Hypertension*. 2009;54(6):1423-8. DOI: 10.1161/HYPERTENSIONAHA.109.139279.
4. Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes Care*. 2002;25(3):512-6.
5. Saydah SH, Geiss LS, Tierney E, Benjamin SM, Engelgau M, Brancati F. Review of the performance of methods to identify diabetes cases among vital statistics, administrative, and survey data. *Ann Epidemiol*. 2004;14(7):507-16.
6. Tu K, Campbell NR, Chen ZL, Cauch-Dudek KJ, McAlister FA. Accuracy of administrative databases in identifying patients with hypertension. *Open Med*. 2007;1(1):s18-26.
7. Gundgaard J, Ekholm O, Hansen EH, Rasmussen NK. The effect of non-response on estimates of health care utilisation: linking health surveys and registers. *Eur J Public Health*. 2008;18(2):189-94.
8. Alshammari AM, Hux JE. The impact of non-fee-for-service reimbursement on chronic disease surveillance using administrative data. *Can J Public Heal*. 2009;100(6):472-4.
9. Wranik DW, Durier-Copp M. Physician remuneration methods for family physicians in Canada: expected outcomes and lessons learned. *Heal Care Anal*. 2010;18(1):35-59. DOI: 10.1007/s10728-008-0105-9.
10. Bernstein CN, Blanchard JF, Rawsthorne P, Wajda A, Revision N, Modification C. Epidemiology of Crohn's Disease and ulcerative colitis in a central Canadian province: a population-based study. *Am J Epidemiol*. 1999;149(10):916-24.
11. Robinson JR, Young TK, Roos LL, Gelskey DE, Robinson JR. Estimating the burden of disease data and self-reports administrative comparing. *Med Care*. 1997;35(9):932-47.
12. Lix LM, Yogendran MS, Leslie WD, et al. Using multiple data features improved the validity of osteoporosis case ascertainment from administrative databases. *J Clin Epidemiol*. 2008;61(12):1250-60.
13. Koleba T, Pohar SL, Johnson JA. Prescription drug data and the National Diabetes Surveillance System case definition. *Can J Diabetes*. 2007;31(1):47-53. DOI: 10.1016/S1499-2671(07)11010-8.
14. Clottey C, Mo F, LeBrun B, Mickelson P, Niles J, Robbins G. The development of the National Diabetes Surveillance System (NDSS) in Canada. *Chronic Dis Can*. 2001;22(2):67-9.
15. Roos NP, Mustard CA. Variation in health and health care use by socioeconomic status in Winnipeg, Canada: does the system work well? Yes and no. *Milbank Q*. 1997;75:89-111.
16. SAS Institute Inc. SAS/STAT 9.3 User's Guide. Cary (NC): SAS Institute Inc.; 2011.
17. Desai JR, Wu P, Nichols GA, Lieu TA, O'Connor PJ. Diabetes and asthma case identification, validation, and representativeness when using electronic health data to construct registries for comparative effectiveness and epidemiologic research. *Med Care*. 2012;50 Suppl:S30-S35. DOI: 10.1097/MLR.0b013e318259c011.
18. Das B, Clegg LX, Feuer EJ, Pickle LW. A new method to evaluate the completeness of case ascertainment by a cancer registry. *Cancer Causes Control*. 2008;19(5):515-25. DOI: 10.1007/s10552-008-9114-0.
19. Silcocks PB, Robinson D. Simulation modelling to validate the flow method for estimating completeness of case ascertainment by cancer registries. *J Public Health*. 2007;29(4):455-62.
20. Lix LM, Yao X, Kephart G, et al. A prediction model to estimate completeness of electronic physician claims databases. *BMJ Open*. 2015;5:e006858. DOI: 10.1136/bmjopen-2014-006858.