PAIN®

OPEN

Assessing the validity of health administrative data compared to population health survey data for the measurement of low back pain

Jessica J. Wong[a,b,*], Pierre Côté[a,b,c,d], Andrea C. Tricco[a,d,e], Tristan Watson[f], Laura C. Rosella[a,f]

Abstract

Low back pain (LBP) is a high-burden condition that lacks routine surveillance data. Health administrative data may be used for surveillance, but their validity for measuring LBP in the general population has not been established. We aimed to (1) determine the validity of health administrative data to measure LBP compared to self-reported LBP in a population-based sample of Ontario adults; and (2) describe the differences in characteristics of LBP cases based on data sources. Adult respondents (≥18 years) of the Canadian Community Health Survey (CCHS) from 2003 to 2012 were included (N = 150,695). Canadian Community Health Survey data were individually linked to health administrative data, including Ontario Health Insurance Plan and hospitalization data. The reference standard was defined as self-reported back problem diagnosed by a health professional in the CCHS. Measurement of LBP from billing records was defined as ≥1 physician billing or procedural code for LBP during the year preceding CCHS interview date. We measured concurrent validity by comparing prevalence, agreement (kappa), and accuracy (sensitivity, specificity, and positive and negative predictive values [PV]) of administrative data to measure LBP. Prevalence of LBP was higher using self-reported (21.2%) than administrative data (10.2%), and agreement was low (kappa = 0.21). Administrative data had sensitivity 23.9% (95% CI 23.1-24.6), specificity 93.4% (95% CI 93.2-93.7), positive PV 50.4% (95% CI 49.1-51.7), and negative PV 82.0% (95% CI 81.7-82.3). Characteristics of LBP cases based on data sources differed in sex, health/behaviour characteristics, and allied health care utilization. Using health administrative data significantly underestimates the prevalence of LBP. This can lead to misclassification bias that is likely nondifferential in epidemiological studies.

Keywords: Low back pain, Health administrative data, Population health survey, Self-reported data, Validity

1. Introduction

Low back pain (LBP) is the leading cause of years lived with disability globally.[17] Low back pain is burdensome to patients, families, and health systems, and is associated with high healthcare utilization and costs.[11,12,17] The lifetime prevalence of LBP is approximately 80%.[6,35] Although most episodes

http://dx.doi.org/10.1097/j.pain.0000000000002003

resolve, 10% to 20% of adults with LBP experience chronic symptoms, functional limitations, or difficulties returning to work.[4,5]

Accurate LBP measurement at the population level is necessary to inform disease surveillance, public health and intervention strategies, healthcare planning, and research. Health administrative data are increasingly being used for disease surveillance. Since health administrative data are routinely collected for other purposes, it serves as a relatively inexpensive data source that provides population-level information over time. Health administrative data can be individually linked to a range of data sources, including vital statistics and population health survey data. Studies found varied agreement between health administrative data and self-reported data for ascertaining chronic diseases, including myocardial infarction, asthma, diabetes, chronic lung disease, stroke, hypertension, congestive heart failure, and depression.[10,23,24,29,34] Agreement was higher for well-defined chronic diseases requiring long-term management, such as diabetes, and lower for those less clearly defined, such as congestive heart failure.[23,34]

Few studies have assessed the validity of health administrative data for measuring LBP in a general population cohort. Lacasse et al.[20] assessed the validity of diagnostic codes for measuring chronic LBP in Régie de l'assurance maladie du Québec administrative database compared to tertiary care patient population in the Quebec Pain Registry. In this study, administrative data (≥1 healthcare encounter for LBP) had 0.65 sensitivity, 0.83 specificity, 0.73 positive predictive value, and

0.78 negative predictive value.[20] Patient populations from tertiary care clinics have a much higher likelihood of LBP, impacting predictive values.[13] Lisi et al.[22] compared primary care data with a chart review in a Veterans database for differentiating acute from nonacute LBP cases, which may also have different background risk for LBP. This study identified an administrative model of 18 variables for differentiating acute vs nonacute LBP case among veterans (C-statistic = 0.819).[22] This study aimed to discriminate LBP duration (acute or nonacute) among veterans diagnosed with LBP. To the best of our knowledge, our study is the first to assess the validity of using health administrative data compared to population health survey data for measuring LBP for an entire province covered under a single health system to inform generalizability of results. This information allows critical insight into the validity of health administrative data for measuring LBP in the general adult population. We need to explore characteristics of LBP cases to elucidate potential selection bias or measurement error when defining LBP populations, particularly when used in epidemiologic studies.

The objective was to determine the validity of health administrative data to identify the presence of LBP using self-reported LBP as reference standard in a population-based sample of adults in Ontario. We aimed to describe the differences in characteristics of LBP cases based on data sources.

2. Methods

We conducted a population-based study of Ontario adult respondents of the Canadian Community Health Survey (CCHS). This project received ethics approval from the Health Sciences Research Ethics Board at the University of Toronto (Ref # 37424).

2.1. Study sample

The sample included all Ontario adults aged 18 years and older who responded to at least 1 of the 5 CCHS cycles (cycle 2.1 [2003-2004], cycle 3.1 [2005-2006], cycle 4.1 [2007-2008], 2009/2010, and 2011/2012). We excluded respondents who could not be linked with health administrative databases or had a death date before the CCHS interview date. We only used data from the first CCHS interview for respondents of multiple survey cycles.

2.2. Data sources

Canadian Community Health Survey data were individually linked to individual-level healthcare utilization data from health administrative databases using unique healthcare identifiers (encrypted Ontario Health Insurance Plan [OHIP] numbers). The CCHS is a cross-sectional survey administered by Statistics Canada that collects data on the distribution of health determinants, outcomes, and healthcare use across Canada.[32] The CCHS uses a multistage sampling survey design to target Canadians aged 12 years and older living in private dwellings, and excludes persons living in institutions (eg, those living in long-term care or complex continuing care facilities), full-time members of the Canadian Forces, and persons living on-reserve and other First Nations settlements.[32] We restricted the sample to respondents aged 18 years and older to focus on adults with LBP. Starting in 2001, the CCHS collected data from a sample of respondents every 2 years until 2007, from which CCHS data were collected annually.[32] Canadian Community Health Survey data are representative of 98% of the Canadian population aged

12 years and older living in private dwellings at national and provincial levels, with response rates greater than 75%.[32] Detailed survey methodology is described elsewhere.[30]

We used health administrative data from OHIP, Canadian Institute for Health Information Discharge Abstract Database and Same-Day Surgeries, and National Ambulatory Care Reporting System to capture physician billing and hospitalization data related to LBP. Ontario Health Insurance Plan covers all Ontario residents, including all CCHS respondents, as a single-payer health insurance system. These data cover all healthcare providers who can claim OHIP (eg, physicians and laboratories) and include service codes, dates of service, and associated diagnosis.[18] Canadian Institute for Health Information Discharge Abstract Database and Same-Day Surgeries collect demographic, administrative, and clinical data on hospital discharges and same-day surgeries, which are received from acute care facilities, health/regional authority, or ministry of health depending on the province. National Ambulatory Care Reporting System captures data on all hospital-based and community-based ambulatory care, collected from specific facilities, regional health authorities, and ministries of health.

Ontario is the largest province by population (~14.2 million in 2017) in Canada, and the most ethnically diverse province with more than 200 ethnicities represented.[31] In Ontario, many healthcare services are publicly funded, including visits to family physicians and specialists, and most basic and emergency healthcare services (eg, surgery and hospital stays).[25] These services are paid through OHIP, which is the government-run provincial health insurance plan.

2.3. Measurement of low back pain using self-reported data from Canadian Community Health Survey (reference standard)

As the reference standard, self-reported LBP was obtained from the CCHS question: "Do you have back problems, excluding fibromyalgia and arthritis?" This CCHS question refers to "conditions diagnosed by a health professional and are expected to last or have already lasted 6 months or more." Individuals who responded yes to this question were classified as having LBP. This definition of self-reported LBP has been used in previous studies.[1,2,7,21,26]

2.4. Measurement of low back pain using health administrative data

Using health administrative data, LBP cases were defined as having at least 1 billing or procedural code related to LBP in the year before the CCHS interview date. Low back pain codes for physician billing, hospital visits, or procedural codes for spinal imaging (radiographs, computed tomography, and magnetic resonance imaging) were used. *International Classification of Diseases* (*ICD*)-10 for LBP-related physician billing and hospital visits included M47, M48, M51, M53, M54, M99, and S33, with similar *ICD-9* codes for LBP (Appendix I, available at http://links. lww.com/PAIN/B121). In addition, we used the ICES equivalency table for *ICD-10* to *ICD-9* codes to inform the mapping of codes for LBP. The primary author (J.J.W.; clinician and epidemiologist) looked up each *ICD-10* code in the equivalency table to retrieve the suggested *ICD-9* code and assessed the description of the *ICD-9* code to ensure that it was coding LBP. Some *ICD-10* codes mapped to the same *ICD-9* code, resulting in fewer *ICD-9* codes listed for LBP. These LBP billing and procedural codes were informed by previous studies.[3,14]

2.5. Sample characteristics

Informed by previous literature,[15] the following characteristics from CCHS were used to describe the sample: sex (male/female); age (years); household income quintile (lowest to highest quintile); education (less than secondary, secondary graduate, more than secondary); body mass index (overweight/obese, normal weight, other); physical activity status (active/moderately active, inactive); alcohol consumption (heavy/moderate drinker, light/never drinker); smoking status (former/current smoker, never smoker); self-reported chronic conditions (at least one of the following: asthma, arthritis, migraines, diabetes, hypertension, heart disease, stroke, cancer, intestinal ulcer, mood disorder, or anxiety); self-rated general health (excellent/very good/good; fair/poor); consulted a physiotherapist in the past 12 months (yes/no); and consulted a chiropractor in the past 12 months (yes/no). Specifically, prevalence of LBP increases with age and is higher in females than males.[15] Adults with physical and mental comorbidities, obesity, or who are smokers are at greater risk of reporting LBP.[15]

2.6. Analysis

We calculated the period prevalence and 95% confidence interval (CI) of LBP among adults based on health administrative and self-reported data. We calculated the period prevalence from 2003 to 2012, and for each CCHS cycle spanning 2 years. For the prevalence of LBP using self-reported data from 2003 to 2012, we combined all 5 CCHS cycles using the pooled approach to identify the total number of prevalent cases.[33] The prevalence was weighted using sampling survey weights from Statistics Canada to determine the represented population in Ontario during the study period (ie, total weighted number of LBP cases divided by total weighted population).

We described LBP cases based on health administrative and self-reported data with respect to sociodemographic characteristics (age, sex, income quintile, and education), health and behaviour characteristics (body mass index, physical activity status, alcohol consumption, smoking status, chronic conditions, and self-rated general health), and allied healthcare utilization (physiotherapy and chiropractic care).

To determine agreement between the 2 data sources, we calculated Cohen's kappa coefficient and 95% CI, and the proportion of positive and negative agreement. Cohen's kappa coefficient measures the proportion of agreement that is beyond what is expected by chance.[8]

To determine validity, we calculated sensitivity, specificity, and 95% CI. We also calculated positive predictive value and negative predictive value and 95% CI to determine the predictive value of health administrative data for measuring LBP.

To assess the potential source of misclassification bias, we conducted sensitivity analyses and recalculated measures of agreement and validity: (1) using a 2-year lookback window for LBP measurement in health administrative data; and (2) including a less specific LBP diagnostic code (ie, DXCODE 847 for sprains, strains of neck, lower back, and coccyx) within a 1-year lookback window. All analyses were performed using SAS V.9.4.[28] All point estimates were based on survey sampling weights, and variance estimates were based on bootstrap weights computed using balanced repeated replication. Statistics Canada provided the survey sampling weights and bootstrap weights to account for the complex survey design of CCHS.[33]

3. Results

The sample included 150,695 adult CCHS respondents (**Table 1**; Appendix II, available at http://links.lww.com/PAIN/B121). The sample has a relatively even distribution by sex, and 16% were adults aged 65 years and older, 49% were considered overweight or obese based on body mass index, 49% were physically active or moderately active, 54% never smoked, and 51% had one or more chronic conditions.

The majority of LBP cases in health administrative data came from physician visits in OHIP data (95%) compared to hospitalizations data (13%) with some overlap between databases. In the OHIP data, 49% of LBP codes were physician billings, whereas 51% were procedural codes, and 24% of individuals had more than 1 LBP code within the 1-year lookback window.

The prevalence of LBP among adults was lower when measured with health administrative data compared to self-reported data (**Table 2**). Specifically, the period prevalence from 2003 to 2012 was 21.2% (95% CI 20.9-21.5) using self-reported data compared to 10.2% (95% CI 10.0-10.5) using health administrative data. The prevalence of LBP was consistent across CCHS cycles, ranging from 19.8% (95% CI 19.0-20.5) in 2011/2012 to 22.6% (95% CI 21.9-23.3) in 2007/2008 for self-reported data, and 9.7% (95% CI 9.1-10.4) in 2011/2012 to 10.7% (95% CI 10.2-11.3) in 2005/2006 in health administrative data.

Agreement between health administrative and self-reported data was kappa = 0.21 (95% CI 0.21-0.21) (**Table 3**).

Sensitivity of using health administrative data with a 1-year lookback window was 24% (95% CI 23.1-24.6), specificity was 93% (95% CI 93.2-93.7), positive predictive value was 50% (95% CI 49.1-51.7), and negative predictive value was 82% (95% CI 81.7-82.3) (**Table 3**). Positive and negative agreements were 32% and 87%, respectively.

When using a 2-year lookback window for defining LBP, the measures of validity for health administrative data were higher for sensitivity (37% vs 24%) and negative predictive value (84% vs 82%) compared to the 1-year lookback window (**Table 3**). However, using a 2-year lookback window for LBP had lower specificity (88% vs 93%) and positive predictive value (45% vs 50%) compared to the 1-year lookback window. When adding a less specific LBP code (ie, DXCODE 847) to define LBP, the measures of validity for health administrative data were higher for sensitivity (26% vs 24%), unchanged for negative predictive value (24%), but lower for specificity (92% vs 93%) and positive predictive value (48% vs 50%) (**Table 3**).

We compared the characteristics of LBP cases based on the data source according to: (1) LBP cases identified using health administrative data that did not self-report LBP (ie, identified in health administrative data only); and (2) LBP cases identified using self-reported data that did not seek medical care (ie, identified in self-reported data only) (**Table 4**). Compared to LBP cases identified in self-reported data only, a higher proportion of LBP cases identified in health administrative data only was female (56% vs 53%), had excellent-to-good self-rated general health (84% vs 79%), or consulted a physiotherapist in the past 12 months (18% vs 12%). A lower proportion of LBP cases identified in health administrative data only were heavy/moderate alcohol drinkers (24% vs 30%), former/current smokers (43% vs 53%), had chronic conditions (60% vs 71%), or consulted a chiropractor in the past 12 months (13% vs 24%).

Table 1

Demographic and clinical characteristics of Ontarian adult respondents of 5 cycles of the Canadian Community Health Survey (2003-2012) linked to health administrative data, N = 150,695.

Characteristic	N	Weighted distribution (%)*†
Sex		
Male	67,935	48.8
Female	82,760	51.2
Age group (y)		
18-34	35,800	29.3
35-49	35,504	30.2
50-64	39,594	24.2
65-74	21,554	9.4
≥75	18,243	6.9
Income quintile		
1 (lowest 20%)	4931	18.4
2	5498	19.2
3	6251	19.6
4	5822	21.1
5 (highest 20%)	6543	21.7
Education		
Less than secondary	16,502	6.1
Secondary graduate	19,594	10.5
More than secondary	108,318	77.5
Body mass index (kg/m²)		
Overweight/obese ≥ 25	79,331	49.1
Normal weight 18.5-24.9	61,812	43.9
Other	9394	7.0
Physical activity status		
Active/moderately active	75,803	49.4
Inactive	71,994	48.4
Alcohol consumption		
Heavy/moderate drinker	44,210	28.8
Never/light drinker	103,992	69.9
Smoking status		
Former or current smoker	72,851	42.8
Never smoker	71,635	53.5
Chronic conditions‡		
Yes	89,277	51.0
No	60,781	48.7
Self-rated general health		
Excellent/very good/good	128,066	87.6
Fair/poor	22,500	12.3
Consulted physiotherapist (in the past 12 mo)§		
Yes	10,790	8.7
No	110,778	91.2
Consulted chiropractor (in the past 12 mo)§		
Yes	15,817	11.5
No	105,772	88.4

* Weighted using Canadian Community Health Survey sampling weights to provide population estimates.
† Percentages may not add up to 100% due to missing values.
‡ Report at least one of the following chronic conditions: asthma, arthritis, migraines, diabetes, hypertension, heart disease, stroke, cancer, intestinal ulcer, mood disorder, or anxiety.
§ Based on 4 Canadian Community Health Survey cycles (2003-2010) due to data availability.

4. Discussion

The prevalence of LBP was lower in health administrative data (10.2%) compared to self-reported data (21.2%), suggesting that using health administrative data underestimates the prevalence of LBP among adults in the general population. The agreement between the 2 data sources was low (kappa = 0.21), which influences the sensitivity and specificity. Health administrative data had 24% sensitivity and 93% specificity, which indicates that agreement was lower for identifying adults as having LBP compared to adults without LBP between data sources. A positive predictive value of 50% suggests that using health administrative data to measure LBP can lead to misclassification bias that is likely non-differential. The positive predictive value is impacted by the prevalence of LBP, which was 21% of adults based on self-reported data. Using health administrative data with a 1-year lookback window for measuring LBP performed better than the 2-year lookback window because it had higher specificity (93% vs 88%) and positive predictive value (50% vs 45%).

Table 2

Prevalence of low back pain among adults in Ontario (2003-2012) based on self-reported data and health administrative data.*†

Time period based on CCHS cycles	N	Prevalence using self-reported data %, 95% CI	Prevalence using health administrative data %, 95% CI
2003-2012	150,537	21.2, 95% CI 20.9-21.5	10.2, 95% CI 10.0-10.5
2003-2004	30,147	21.6, 95% CI 20.9-22.4	9.9, 95% CI 9.3-10.4
2005-2006	30,068	20.8, 95% CI 20.2-21.5	10.7, 95% CI 10.2-11.3
2007-2008	31,267	22.6, 95% CI 21.9-23.3	10.4, 95% CI 9.9-10.9
2009-2010	30,061	21.4, 95% CI 20.6-22.1	10.4, 95% CI 9.8-11.0
2011-2012	28,994	19.8, 95% CI 19.0-20.5	9.7, 95% CI 9.1-10.4

* Missing N = 158 (0.1%) for self-reported data from the Canadian Community Health Survey.
† Weighted using Canadian Community Health Survey sampling weights to provide population estimates.
CCHS, Canadian Community Health Survey; CI, confidence interval.

Characteristics of LBP cases based on the 2 data sources differed in sex, health and behaviour characteristics, and allied healthcare utilization, suggesting that health administrative data identified adults with LBP who were healthier (ie, no chronic conditions or excellent-to-good general health) or leading a healthier lifestyle. Overall, our findings suggest that using health administrative data underestimates the prevalence of LBP in adults and can lead to misclassification bias that is likely nondifferential.

There are potential reasons why fewer LBP cases were captured in the health administrative data. First, some adults with LBP may seek care from allied health professionals, such as chiropractors, rather than visit physicians or the hospital. We found higher utilization of chiropractic care among LBP cases in CCHS compared to health administrative data (24% vs 14%). A scoping review reported that the median 12-month utilization of chiropractic services was 9% (interquartile range 7%-13%), with LBP being the most common reason for seeking chiropractic care.[2] Second, physician billings for adults with LBP may use nonspecific codes for pain instead of LBP codes. Our sensitivity analysis shows that including nonspecific pain codes would lower specificity and positive predictive value when measuring LBP. Third, individuals may overreport their LBP as chronic in surveys. This seems unlikely because the prevalence of 21% based on self-reported data is similar to the global prevalence of 20% for chronic LBP reported in a recent systematic review.[17] Finally, it is possible that more severe LBP cases were captured in health administrative data. This is unlikely when considering chronicity

and impact on general health as measures of LBP severity. The CCHS question refers to back problems of ≥6 months' duration diagnosed by a health professional. In addition, a higher proportion of LBP cases in administrative data had excellent-to-good general health compared to self-reported data (84% vs 79%).

As mentioned previously, Lacasse et al.[20] assessed the validity of health administrative data for measuring chronic LBP in the general population compared to the Quebec Pain Registry. Using at least 1 healthcare encounter in health administrative data for chronic LBP, Lacasse et al.[20] reported a sensitivity of 0.65 (95% CI 0.59-0.71), specificity of 0.83 (95% CI 0.79-0.87), positive predictive value of 0.73 (95% CI 0.66-0.79), and negative predictive value of 0.78 (95% CI 0.73-0.82). These results differ from our study results likely due to differences in the sampling frame and case definition. The study by Lacasse et al. was restricted to a patient population with a much higher likelihood of LBP, given they attended a pain clinic. Lacasse et al.[20] used the Quebec Pain Registry as the reference standard, which is a data registry of patients suffering from chronic noncancer pain within tertiary care clinics offering multidisciplinary pain treatment. These pain clinics are affiliated with the Integrated University Health Networks in Quebec. Patients were identified as having chronic LBP based on diagnoses by the physician of the pain clinic. The sample targeted by Lacasse et al.[20] had a LBP prevalence of 40%, which is higher than the LBP prevalence of our study and would impact the positive predictive value (73% vs 50%).

Table 3

Measures of validity between health administrative data (using 1- and 2-year lookback windows, and less specific low back pain code) compared to self-reported data for measuring low back pain among adults in Ontario from 2003 to 2012, N = 150,537.

	Using health administrative data with 1-y lookback window for LBP	Using health administrative data with 2-y lookback window for LBP	Using health administrative data with less specific diagnostic code* for LBP
Kappa, 95% CI	0.21 (95% CI 0.21-0.21)	0.26 (95% CI 0.26-0.26)	0.22 (95% CI 0.22-0.22)
Sensitivity, 95% CI	23.9% (95% CI 23.1-24.6)	36.5% (95% CI 35.7-37.4)	26.0% (95% CI 25.3-26.8)
Specificity, 95% CI	93.4% (95% CI 93.2-93.7)	88.0% (95% CI 87.7-88.3)	92.3% (95% CI 92.0-92.6)
Positive predictive value, 95% CI	50.4% (95% CI 49.1-51.7)	45.0% (95% CI 44.0-46.0)	47.7% (95% CI 46.5-48.9)
Negative predictive value, 95% CI	82.0% (95% CI 81.7-82.3)	83.7% (95% CI 83.4-84.0)	82.2% (95% CI 81.9-82.5)
Positive agreement	32%	40%	34%
Negative agreement	87%	86%	87%

* Includes Ontario Health Insurance Plan (OHIP) DXCODE 847 (sprains, strains of neck, lower back, and coccyx).
CI, confidence interval; LBP, low back pain.

Table 4

Characteristics of low back pain cases in Ontario (2003-2012) identified in: (1) self-reported data only; (2) health administrative data only; and (3) both health administrative and self-reported data.*†

Characteristic	Self-reported data only (%) N = 27,973	Health administrative data only (%) N = 7864	Both health administrative and self-reported data (%) N = 8833
Sex‡			
Male	47.1	43.9	43.0
Female	52.9	56.1	57.0
Age group (y)‡			
18-34	18.6	19.0	14.7
35-49	30.9	30.3	29.4
50-64	29.9	27.6	31.4
65-74	12.0	12.5	13.2
75+	8.6	10.6	11.2
Income quintile†‡			
1 (lowest 20%)	20.0	21.0	26.5
2	19.2	22.8	23.1
3	19.9	17.8	20.8
4	21.9	20.3	12.8
5 (highest 20%)	19.0	18.0	16.8
Education†			
Less than secondary	8.2	7.7	10.7
Secondary graduate	11.6	11.8	12.5
More than secondary	74.6	75.0	71.6
Body mass index (kg/m^2)‡			
Overweight/obese ≥25	55.3	52.4	57.5
Normal weight 18.5-24.9	38.2	40.1	36.3
Other	6.5	7.6	6.1
Physical activity status			
Active/moderately active	44.6	42.9	40.1
Inactive	53.0	54.2	57.4
Alcohol consumption‡			
Heavy/moderate drinker	29.7	23.8	25.4
Light/never drinker	68.8	75.1	72.6
Smoking status‡			
Former or current smoker	51.6	43.0	53.8
Never smoker	44.8	53.3	42.4
Chronic conditions‡§			
Yes	70.5	60.3	77.5
No	29.2	39.3	22.4
Self-rated general health‡			
Excellent/very good/good	78.6	83.5	65.4
Fair/poor	21.3	16.4	34.5
Consulted physiotherapist (in the past 12 mo)‡‖			
Yes	12.1	18.0	28.4
No	87.8	81.8	71.4
Consulted chiropractor (in the past 12 mo)‡‖			
Yes	23.5	13.1	25.5
No	76.4	86.9	74.4

* Weighted using Canadian Community Health Survey sampling weights to provide population estimates.
† Percentages may not add up to 100% due to missing values.
‡ $P < 0.05$ when comparing self-reported data to health administrative data (χ^2).
§ Report at least one of the following chronic conditions: asthma, arthritis, migraines, diabetes, hypertension, heart disease, stroke, cancer, intestinal ulcer, mood disorder, or anxiety.
‖ Based on 4 Canadian Community Health Survey cycles (2003-2010) due to data availability.

Our study assessed the characteristics of LBP cases identified in each data source to potentially explain differences in measurement of LBP. Characteristics of LBP cases based on the data sources differed in sex, health/behaviour characteristics (alcohol consumption, smoking, and chronic conditions), and allied healthcare use, suggesting that health administrative data captured adults with LBP who were healthier (ie, no chronic conditions or excellent-to-good general health) or leading a healthier lifestyle. Results were consistent with the finding that women are slightly more likely to seek care for LBP than men.[9] Characteristics of LBP cases depending on the data source are important to consider in future epidemiologic studies when defining LBP and examining an exposure–outcome association. This is important because using health administrative to define

LBP may capture healthier individuals (eg, lower proportion of heavy/moderate alcohol drinkers, former/current smokers, those with chronic conditions or excellent-to-good general health), which may confound the exposure–outcome association being studied. Users and researchers should be cautious about the limitations of health administrative data due to underestimating LBP prevalence and potential misclassification bias.

This study offers several advantages that overcome limitations from previous studies. First, each CCHS respondent was linked individually and deterministically to population-based health administrative databases. Deterministic linkage is an all-or-nothing linkage approach where records are matched using an exact match of unique identifying information.[27] Second, CCHS data are representative of 98% of the Canadian population aged 12 years and older living in private dwellings.[32] Third, we used LBP billing and procedural codes informed by previous literature,[3,14] and used both a 1-year lookback window and 2-year lookback window to assess the validity of health administrative data for measuring LBP.

Our study has limitations. First, there is no gold standard for measuring LBP in the general population. Our choice to use self-reported information as a reference standard may result in under-reporting of LBP. The CCHS question on back problems does not distinguish between back problems expected to last 6 months or more by the health professional versus the participant. However, the prevalence of 21% based on self-reported data is similar to the global prevalence of 23% for LBP (20% for chronic LBP) reported in a recent systematic review, suggesting unlikely underreporting or poor recall for LBP in our study.[16] In addition, the CCHS question used for measuring LBP refers to "back problems," which does not differentiate between pain in the midback vs low back region. However, the incidence of midback pain is low in adults, with a 1-month incidence proportion of less than 1% among adults, and is therefore unlikely to affect results.[19] Second, the billing and procedural codes for LBP in the health administrative data exclude services not covered by OHIP, such as allied health services (eg, physiotherapy and chiropractic care). However, we were able to describe LBP cases with respect to consultations to physiotherapists and chiropractors in the past 12 months based on CCHS data availability. Third, the *ICD* codes do not specify LBP duration, and some LBP cases in the administrative data may be acute cases. This would have led to an overestimation of LBP prevalence in administrative data. However, multiple *ICD* codes refer to conditions that are often chronic (eg, spinal stenosis and myelopathy) and we attempted to capture chronic LBP by using both 1- and 2-year lookback windows from the CCHS interview date. Finally, the CCHS sampling frame includes individuals living in private dwellings only, and results may not be generalizable to other populations (eg, persons living in institutions or on reserve and other First Nations settlements).

5. Conclusions

We found that using health administrative data to measure LBP underestimates the prevalence of LBP among adults in the Ontario general population. Moreover, using health administrative data for LBP measurement can lead to misclassification bias that is likely nondifferential. Therefore, users and researchers should be cautious about the limitations of this data source for LBP disease surveillance, healthcare planning, and epidemiologic research. Future epidemiologic studies studying LBP in the Canadian general population should consider using CCHS data for LBP measurement. To inform routine disease surveillance in health systems, future research aimed to develop accurate health administrative data algorithms for measuring LBP may be warranted.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/B121.

References

[1] Alkherayf F, Agbi C. Cigarette smoking and chronic low back pain in the adult population. Clin Invest Med 2009;32:E360–367.

[2] Bielecky A, Chen C, Ibrahim S, Beaton DE, Mustard CA, Smith PM. The impact of co-morbid mental and physical disorders on presenteeism. Scand J Work Environ Health 2015;41:554–64.

[3] Bouck Z, Pendrith C, Chen XK, Frood J, Reason B, Khan T, Costante A, Kirkham K, Born K, Levinson W, Bhatia RS. Measuring the frequency and variation of unnecessary care across Canada. BMC Health Serv Res 2019;19:446.

[4] Carey TS, Evans A, Hadler N, Kalsbeek W, McLaughlin C, Fryer J. Care-seeking among individuals with chronic low back pain. Spine 1995;20:312–17.

[5] Carey TS, Evans AT, Hadler NM, Lieberman G, Kalsbeek WD, Jackman AM, Fryer JG, McNutt RA. Acute severe low back pain. A population-based study of prevalence and care-seeking. Spine 1996;21:339–44.

[6] Cassidy JD, Carroll LJ, Côté P. The Saskatchewan health and back pain survey. The prevalence of low back pain and related disability in Saskatchewan adults. Spine 1998;23:1860–6; discussion 1867.

[7] Cassidy T, Fortin A, Kaczmer S, Shumaker JTL, Szeto J, Madill SJ. Relationship between back pain and urinary incontinence in the Canadian population. Phys Ther 2017;97:449–54.

[8] Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37–46.

[9] Ferreira ML, Machado G, Latimer J, Maher C, Ferreira PH, Smeets RJ. Factors defining care-seeking in low back pain a meta-analysis of population based surveys. Eur J Pain 2010;14:747.e741–747.

[10] Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying individuals with physcian diagnosed COPD in health administrative databases. Copd 2009;6:388–94.

[11] Global Burden of Disease 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;386:743–800.

[12] Global Burden of Disease 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;388:1545–602.

[13] Gordis L. Epidemiology. 5th ed. Philadelphia: Elsevier Saunders, 2014.

[14] Gunz AC, Canizares M, Mackay C, Badley EM. Magnitude of impact and healthcare use for musculoskeletal disorders in the paediatric: a population-based study. BMC Musculoskelet Disord 2012;13:98.

[15] Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, Hoy D, Karppinen J, Pransky G, Sieper J, Smeets RJ, Underwood M. What low back pain is and why we need to pay attention. Lancet 2018;391:2356–67.

[16] Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, Woolf A, Vos T, Buchbinder R. A systematic review of the global prevalence of low back pain. Arthritis Rheum 2012;64:2028–37.

[17] Hoy D, March L, Brooks P, Blyth F, Woolf A, Bain C, Williams G, Smith E, Vos T, Barendregt J, Murray C, Burstein R, Buchbinder R. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014;73:968–74.

[18] ICES. Data dictionary: OHIP Library, 2019. Available at: https://datadictionary.ices.on.ca/Applications/DataDictionary/Library.aspx?Library=OHIP. Accessed April 1, 2019.

[19] Johansson MS, Jensen Stochkendahl M, Hartvigsen J, Boyle E, Cassidy JD. Incidence and prognosis of mid-back pain in the general population: a systematic review. Eur J Pain 2017;21:20–8.

[20] Lacasse A, Ware MA, Dorais M, Lanctot H, Choiniere M. Is the Quebec provincial administrative database a valid source for research on chronic non-cancer pain? Pharmacoepidemiol Drug Saf 2015;24:980–90.

[21] Lim KL, Jacobs P, Klarenbach S. A population-based analysis of healthcare utilization of persons with back disorders: results from the Canadian Community Health Survey 2000-2001. Spine 2006;31:212–18.

[22] Lisi AJ, Burgo-Black AL, Kawecki T, Brandt CA, Goulet JL. Use of department of veterans affairs administrative data to identify veterans with acute low back pain: a pilot study. Spine 2014;39:1151–6.

[23] Lix LM, Yogendran MS, Shaw SY, Burchill C, Metge C, Bond R. Population-based data sources for chronic disease surveillance. Chronic Dis Can 2008;29:31–8.

[24] Muggah E, Graves E, Bennett C, Manuel DG. Ascertainment of chronic diseases using population health data: a comparison of health administrative data and patient self-report. BMC Public Health 2013;13:16.

[25] Ontario Ministry of Health, Long-term Care. Understanding health care in Ontario, 2012. Available at: http://www.health.gov.on.ca/en/ministry/hc_system/. Accessed April 1, 2019.

[26] Patten SB, Williams JV, Wang J. Mental disorders in a population sample with musculoskeletal disorders. BMC Musculoskelet Disord 2006;7:37.

[27] Roos LL, Wajda A. Record linkage strategies. Part I: estimating information and evaluating approaches. Methods Inf Med 1991;30:117–23.

[28] SAS Enterprise 9.4. Cary: SAS Institute Inc.

[29] Shah BR, Manuel DG. Self-reported diabetes is associated with self-management behaviour: a cohort study. BMC Health Serv Res 2008;8:142.

[30] Statistics Canada. Canadian community health survey 2003: user guide for the public use microdata file. Ottawa: Statistics Canada, 2005. (Catalogue no. 82M0013GPE).

[31] Statistics Canada. Canada at a glance 2018: population, 2018. Available at: https://www150.statcan.gc.ca/n1/pub/12-581-x/2018000/pop-eng.htm?HPA=1. Accessed April 1, 2019.

[32] Statistics Canada. Canadian community health survey (CCHS)—annual component. Available at: http://www23.statcan.gc.ca/imdb-bmdi/document/3226_D7_T9_V8-eng.htm. Accessed October 20, 2018.

[33] Thomas S, Wannell B. Combining cycles of the Canadian community health survey. Health Rep 2009;20:53–8.

[34] Tu K, Campbell NR, Chen ZL, Cauch-Dudek KJ, McAlister FA. Accuracy of administrative databases in identifying patients with hypertension. Open Med 2007;1:e18–26.

[35] Walker BF. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. J Spinal Disord 2000;13:205–17.