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

Jessica J. Wong*[1,2], Pierre Côté[1,2,3,4], Andrea C. Tricco[1,4,5], Tristan Watson[6], Laura C. Rosella[1,6]

[1]Dalla Lana School of Public Health, University of Toronto; [2]Centre for Disability Prevention and Rehabilitation, Ontario Tech University; [3]Faculty of Health Sciences, Ontario Tech University; [4]Institute of Health Policy, Management and Evaluation, University of Toronto; [5]Li Ka Shing Knowledge Institute of St. Michael's Hospital; [6]ICES; *jes.wong@mail.utoronto.ca

## BACKGROUND

- Accurate measurement of low back pain (LBP) at the population level is necessary to inform disease surveillance, health planning, and research.
- Few studies have assessed the validity of health administrative data for measurement of LBP in a general population cohort.[1,2]
- 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.

## OBJECTIVES

1. To determine the validity of health administrative data to identify the presence of LBP using self-reported LBP as the reference standard in a population-based sample of adults in Ontario.
2. To describe differences in characteristics of LBP cases based on datasets.

## METHODS

- Respondents (aged ≥18 years) of the Canadian Community Health Survey (CCHS) cycles from 2003 to 2012 were included (N=150,695).
- CCHS data were individually linked to health administrative data, including Ontario Health Insurance Plan and hospitalization data.
- The reference standard was collected in the CCHS and defined as self-reported back problem diagnosed by a health professional.
- LBP measurement from billing records was ≥1 billing/procedural code for LBP in the year preceding CCHS interview date (informed by literature).[3]
- We assessed concurrent validity by comparing prevalence, agreement (Kappa), sensitivity, specificity, positive and negative predictive values.
- Point and variance estimates were based on survey sampling weights and bootstrap weights (using balanced repeated replication), respectively.[4]

**Figure 1.** Representation of 2x2 table based on data sources

| | | Self-reported low back pain (population health survey) | |
|---|---|---|---|
| | | Yes | No |
| Low back pain based on billing codes (administrative data) | Yes | a | b |
| | No | c | d |

## RESULTS

**Table 1.** Prevalence of low back pain among adults in Ontario (N=150,537)[a,b]

| Time period | Self-reported data Prevalence - % (95% CI) | Health administrative data Prevalence - % (95% CI) |
|---|---|---|
| 2003-2012 | 21.2 (20.9-21.5) | 7.4 (7.2-7.6) |

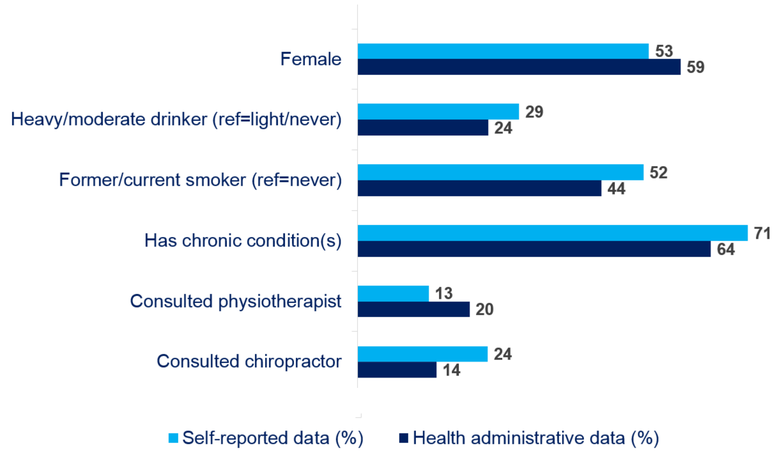
CI - confidence interval; [a]Weighted with CCHS sampling weights
[b]Missing N=158 (0.1%) for Canadian Community Health Survey (CCHS) data

**Table 2**. Measures of validity between health administrative data compared to self-reported data, N=150,537

| | Administrative data with 1-year lookback for LBP | Administrative data with 2-year lookback for LBP |
|---|---|---|
| Kappa, 95% CI | 0.17 (0.17-0.17) | 0.23 (0.23-0.23) |
| Sensitivity, 95% CI | 17.5% (16.9-18.1) | 28.6% (27.9-29.4) |
| Specificity, 95% CI | 95.3% (95.1-95.5) | 91.2% (90.9-91.4) |
| Positive PV, 95% CI | 50.1% (48.5-51.6) | 46.6% (45.5-47.7) |
| Negative PV, 95% CI | 81.1% (80.8-81.4) | 82.6% (82.3-82.9) |
| Positive agreement | 26% | 35% |
| Negative agreement | 88% | 87% |

CI – confidence interval; LBP – low back pain; PV – predictive value

**Figure 2.** Select characteristics (%) of LBP cases identified in 1) self-reported data only or 2) health administrative data only[a], N=150,537




[a]LBP cases did not differ in age, income, education, BMI, physical activity

## KEY FINDINGS

- **Prevalence:** LBP prevalence was lower in administrative data (7.4%) compared to self-reported data (21.2%), suggesting that administrative data underestimates LBP prevalence in adults in the general population.
- **Agreement:** Agreement between the two data sources was low (kappa=0.17), which influences the sensitivity and specificity.
- **Validity measures:** Administrative data had 18% sensitivity and 95% specificity, which indicates that agreement was lower for identifying adults as having LBP compared to adults without LBP between data sources. Positive predictive value of 50% suggests that using administrative data to measure LBP can lead to misclassification bias that is likely non-differential. Positive predictive value is impacted by LBP prevalence, which was 21% of adults based on self-reported data.
- **Characteristics of LBP cases:** Characteristics of LBP cases based on the two data sources differed in sex, health/behaviour characteristics, and allied health care utilization, suggesting that administrative data identified adults with LBP who were healthier.

## SIGNIFICANCE

- Using health administrative data significantly underestimates prevalence of LBP and can lead to misclassification bias that is likely non-differential.
- Users and researchers should be cautious about the limitations of this data source for LBP disease surveillance, health care planning, and epidemiologic research. Future epidemiologic studies studying LBP in the Canadian general population should consider using CCHS data.
- To inform routine disease surveillance in health systems, future research aimed to develop accurate health administrative data algorithms for measuring LBP may be warranted.

## REFERENCES

1. Lacasse A, Ware MA, Dorais M, et al. Is the Quebec provincial administrative database a valid source for research on chronic non-cancer pain? Pharmacoepidemiol Drug Saf 2015;24(9):980-90.
2. Lisi AJ, Burgo-Black AL, Kawecki T, et al. Use of Department of Veterans Affairs administrative data to identify veterans with acute low back pain: a pilot study. Spine 2014;39(14):1151-6.
3. Gunz AC, Canizares M, Mackay C, et al. Magnitude of impact and healthcare use for musculoskeletal disorders in the paediatric: a population-based study. BMC Musculoskelet Disord 2012;13:98.
4. Thomas S, Wannell B. Combining cycles of the Canadian Community Health Survey. Health Rep 2009;20(1):53-8.

UNIVERSITY OF TORONTO
DALLA LANA SCHOOL OF PUBLIC HEALTH

Population Health Analytics Laboratory

IC/ES

Funding for this study was supported by the Canada Research Chair held by LR. LR is funded by a Tier 2 Canada Research Chair in Population Health Analytics. PC is funded by a Tier 2 Canada Research Chair in Disability Prevention and Rehabilitation. AT is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis.