CMAJ OPEN

Research

Using physician billing claims from the Ontario Health Insurance Plan to determine individual influenza vaccination status: an updated validation study

Kevin L. Schwartz MD MSc, Nathaniel Jembere MPH, Michael A. Campitelli MPH, Sarah A. Buchan MSc, Hannah Chung MPH, Jeffrey C. Kwong MD MSc

Abstract

Background: Owing to the absence of a vaccination registry in Ontario, administrative data are currently the best available source to determine population-based individual-level influenza vaccination status. Our objective was to validate physician billing claims for influenza vaccination in the Ontario Health Insurance Plan database against the Canadian Community Health Survey.

Methods: We used self-reported seasonal influenza vaccination status of Ontario residents surveyed between 2007 and 2009 as the reference standard. The survey responses were linked to physician claims database records to validate billing codes for influenza vaccination. We calculated sensitivity, specificity, positive predictive value and negative predictive value with 95% confidence intervals (CIs). We stratified the data by several covariates and comorbidities to determine stratum-specific performance characteristics. We used these estimates to adjust an estimate of influenza vaccine effectiveness for the 2010/11 influenza season.

Results: For the 47 301 patients included in the analysis, the sensitivity for the billing codes was 49.8% (95% CI 49.0%–50.5%), specificity was 95.7% (95% CI 95.5%–96.0%), positive predictive value was 88.4% (95% CI 87.8%–89.0%) and negative predictive value was 74.5% (95% CI 74.0%–74.9%). Performance measures were optimized in patients aged 65 years and older, particularly those with comorbidities.

Interpretation: Although administrative data have limitations for determining influenza vaccination status, owing to the high positive predictive value, they are well suited for self-controlled study designs that are often used to assess vaccine safety. For studies of coverage and effectiveness, restricting the cohort to patients aged 65 years and older will minimize misclassification bias. Performance characteristics from this study can be used to mitigate misclassification bias.

Influenza continues to pose a major public health burden in Canada. It is estimated that 5%–10% of the population has a symptomatic influenza infection each year.[1] Since 2000, the province of Ontario has offered free influenza vaccines to residents aged 6 months and older in a variety of settings, including physician offices, community-based public health clinics, healthcare facilities, workplaces, schools and pharmacies. However, the absence of a comprehensive vaccination registry that captures influenza vaccines delivered in all settings has hindered efforts to evaluate the program in terms of vaccine safety, effectiveness and coverage.

We previously validated physician billing claims for influenza vaccination submitted to the Ontario Health Insurance Plan (OHIP) against self-reported influenza vaccination from the Canadian Community Health Survey cycle 1.1, conducted in 2000/01.[2] We found high specificity (97%) and positive predictive value (91%), moderately high negative predictive value (79%), but lower sensitivity (56%). Sensitivity was higher for adults aged 65 years and older and for patients who reported having chronic medical conditions. Previous studies have found self-reported vaccination status to be valid.[3–10] The low sensitivity of physician billing claims is partially explained by patients receiving influenza vaccines outside of physician offices.[2]

The objective of this study was to update the previous validation with more recent data, and to estimate performance

Competing interests: None declared.

This article has been peer reviewed.

Correspondence to: Jeffrey Kwong, jeff.kwong@utoronto.ca

***CMAJ Open* 2016. DOI:10.9778/cmajo.20160009**

measures of OHIP billing claims for patients with a more comprehensive (and more rigorously determined) list of risk factors for serious influenza infections.

Methods

Study population and setting

This study included Ontario residents who responded to the Canadian Community Health Survey between Jan. 1, 2007, and Sept. 30, 2009, and who agreed to have their survey data linked with provincial health administrative data. These data sets were linked using unique encoded identifiers and analyzed at the Institute for Clinical Evaluative Sciences (ICES). We excluded patients surveyed on Oct. 1, 2009, or later because 2 vaccines were used during the 2009/10 influenza season (the monovalent pandemic A/H_1N_1 vaccine and the trivalent seasonal influenza vaccine), and we were unable to differentiate between them using the OHIP data because the same billing codes were used for both vaccines. Data from more recent cycles of the survey were not yet available in linked format at ICES at the time of manuscript submission. This study was approved by the institutional review board at Sunnybrook Health Sciences Centre, Toronto, Ontario.

Data sources

Canadian Community Health Survey

The Canadian Community Health Survey is a national cross-sectional survey that collects health-related information on patients aged 12 years and older through telephone and in-person interviews. The first 3 iterations in 2000/01 (cycle 1.1), 2003 (cycle 2.1) and 2005 (cycle 3.1) were biennial surveys of about 130 000 respondents. In 2007, Statistics Canada changed the survey design so that data would be collected from about 65 000 respondents each year. The survey excludes people residing on aboriginal settlements, full-time members of the Canadian Armed Forces and people living in institutions (less than 3% of total population). Details of the survey methodology have been described elsewhere.[11] The response rates for the 2007/08 and 2009/10 cycles were 77.6% and 73.2%, respectively. The linkage rate between Canadian Community Health Survey and ICES data was 83%.

Ontario Health Insurance Plan database

The OHIP database contains billing information from about 94% of Ontario's physicians.[12] It excludes physicians not paid through fee-for-service methods. OHIP provides virtually the entire Ontario population with universal insurance coverage for physician services and hospital care, excluding new residents during their initial 3 months in the province.

Outcomes

Influenza immunization status from the Canadian Community Health Survey

Respondents were asked "Have you ever had a flu shot?" Those who stated that they had were then asked "When did you have your last flu shot?" Respondents who specified that they had received the vaccine within the previous 12 months were then asked which month; if they answered the current month, they were asked "Was that this year or last year?"

We classified patients who reported receiving a flu shot within the last 12 months as having been immunized as follows: those whose month of vaccination differed from the month of interview; and those whose month of vaccination matched the month of interview and was in the same year. Because the questionnaire did not determine the exact date of vaccine receipt, patients whose month of vaccination matched the month of interview, but in the previous year, may have received their vaccine more than 365 days earlier. For these specific patients, we classified those whose interview occurred during the first 15 days of the month as having been immunized and those interviewed after the first 15 days of the month as not immunized. We conducted a sensitivity analysis restricting the survey dates from Feb. 1 to Aug. 31 of each year to minimize the risk of immunization year misclassification for patients surveyed during influenza vaccination campaigns in Ontario (usually September to January).

Influenza immunization status from OHIP

To identify influenza vaccination status in the OHIP database, we used the billing codes for vaccination with influenza vaccines, G590 (influenza vaccination plus visit) and G591 (influenza vaccination only). We also included the tracking code Q130 (influenza vaccine tracking code), which is used when a patient has undergone vaccination elsewhere. Physicians belonging to certain remuneration plans receive financial incentives for attaining prespecified targets for influenza vaccination of their patients aged 65 years and older, and all 3 codes are included in the numerator for those calculations. Using the Canadian Community Health Survey interview date as the reference date, we considered the presence of any of the influenza vaccination codes over the previous 365 days to be actively immunized.

Other variables

We determined neighbourhood income quintile using residential postal codes, and defined rural residence as a community size of fewer than 10 000 residents. Having a regular physician was determined from the Canadian Community Health Survey question "Do you have a regular medical doctor?" We evaluated patients for the presence of a number of potential risk factors for serious influenza infection, including chronic cardiovascular diseases (congestive heart failure, history of acute myocardial infarction or acute ischemic stroke, and hypertension), chronic respiratory diseases (asthma and chronic obstructive pulmonary disease), diabetes, chronic kidney disease, cancer, immunosuppression (resulting from infection with HIV or from immunosuppressive therapies), dementia, morbid obesity (body mass index > 40 calculated from the height and weight provided in the survey) and pregnancy (derived from the MOMBABY database). Most of these conditions were defined using previously validated algorithms applied to administrative data sets housed at ICES, including

the OHIP database, the Canadian Institute for Health Information Discharge Abstract Database, the National Ambulatory Care Reporting System, the Same Day Surgery database, the Canadian Organ Replacement Register, the Ontario Renal Reporting System, the Ontario Diabetes Database, the Ontario Cancer Registry, the Ontario Myocardial Infarction Database, the Chronic Obstructive Pulmonary Disease database, the Ontario Drug Benefits database, the Ontario Congestive Heart Failure database, the Ontario Hypertension Database and the Ontario HIV database.[13–29] These databases and the definitions used are described in Appendix 1 (available at www.cmajopen.ca/content/4/3/E463/suppl/DC1).

Statistical analyses

We set self-reported influenza immunization status from the Canadian Community Health Survey as the reference standard. We calculated performance measures (sensitivity, specificity, positive predictive value and negative predictive value) with 95% confidence intervals (CIs) for OHIP physician billing claims for influenza vaccination. We stratified the results by survey cycle, age group, sex, rural versus urban residence, having a regular physician and presence of risk factors for serious influenza infections. We further stratified some of these groups by age (< 65 v. ≥ 65 yr). Statistical analyses were conducted using SAS Enterprise Guide version 6.1 (SAS Institute Inc., Cary, NC).

Application example

To illustrate the applicability of our results, we used the values for sensitivity and specificity to correct the bias arising from misclassification of influenza vaccination status based on OHIP physician billing claims. We applied a SAS macro developed by Fox and associates[30] to results from a previous study by Kwong and associates that assessed vaccine effectiveness against admissions to hospital for laboratory-confirmed influenza among older adults during the 2010/11 influenza season.[31] This macro uses a probabilistic method for conducting a sensitivity analysis using individual-level data. Using the overall sensitivity and specificity from this study's results for influenza immunization status in Ontario patients aged 65 years and older, we calculated vaccine effectiveness corrected for misclassification of the exposure variable (i.e., influenza vaccination).[30]

Results

There were a total of 48 426 survey responses, of which 1122 were excluded for either refusal or an inability to answer the influenza vaccine question and 3 were excluded for invalid birthdates, leaving 47 301 Ontarians included in our analysis (Table 1). Based on survey results, about 40% of patients reported having undergone vaccination against influenza, ranging from less than 25% among patients aged less than 50 years to 68% among those aged 65 years and older. Vaccine coverage was higher among women, older adults, patients with a regular physician and patients with risk factors for serious influenza infection, except for pregnancy.

Table 1: Demographic characteristics of the study population and percentage receiving influenza vaccine within the previous 12 months

Characteristic	Patients, no. (%) *n* = 47 301	Patients who received vaccine, % *n* = 18 684
Canadian Community Health Survey cycle		
2007/08	33 840 (71.5)	38.9
2009/10	13 461 (28.5)	41.1
Sex		
Female	25 904 (54.8)	42.4
Male	21 397 (45.2)	36.1
Age group, yr		
12–17	4071 (8.6)	24.8
18–49	20 247 (42.8)	23.6
50–64	11 815 (25.0)	44.6
≥ 65	11 168 (23.6)	68.4
Residence*		
Urban	37 406 (79.1)	39.3
Rural	9802 (20.7)	40.4
Neighbourhood income quintile†		
1 (lowest)	9342 (19.7)	40.3
2	9402 (19.9)	39.6
3	9467 (20.0)	39.5
4	9568 (20.2)	38.7
5 (highest)	9330 (19.7)	39.7
Has regular doctor		
Yes	43 110 (91.1)	41.3
No	4191 (8.9)	21.7
Risk factors for serious influenza infections‡		
Hypertension	13 826 (29.2)	61.7
Asthma	6225 (13.2)	44.0
Diabetes	4877 (10.3)	63.7
Cancer	2572 (5.4)	64.4
COPD	1596 (3.4)	70.5
Congestive heart failure	1235 (2.6)	74.0
Myocardial infarction	928 (2.0)	69.6
Chronic kidney disease	886 (1.9)	68.9
Morbid obesity§	866 (1.8)	43.7
Stroke	861 (1.8)	66.0
Immunosuppression	679 (1.4)	74.7
Pregnancy¶	512 (1.1)	18.0
Dementia	245 (0.5)	64.5
None of the above	25 924 (54.8)	27.3

Note: COPD = chronic obstructive pulmonary disease.
*Missing data on 95 patients.
†Missing data on 192 patients.
‡May add up to more than 100%; patients may have more than 1 risk factor.
§Body mass index > 40.
¶Date of delivery between Nov. 1 and June 1.

The combined sensitivity for influenza OHIP billing codes was 49.8% (95% CI 49.0%–50.5%), specificity was 95.7% (95% CI 95.5%–96.0%), positive predictive value was 88.4% (95% CI 87.8%–89.0%) and negative predictive value was 74.5% (95% CI 74.0%–74.9%) (Table 2). The sensitivity ranged from 20.3% in adolescents (aged 12–17 yr) to 68.9% in patients aged 65 years and older, whereas specificity was high for those less than 65 years of age (≥ 96.0%) and declined to 82.7% for those aged 65 years and older. Similarly, positive predictive value increased with age, whereas negative predictive value decreased.

Having access to a regular physician substantially improved the sensitivity of OHIP influenza vaccine billing codes, but with some decrease in specificity. The validity of the OHIP influenza immunization codes was fairly consistent across a variety of risk factors, as long as the cohort was restricted to patients aged 65 years and older. For chronic conditions, the sensitivity ranged from 68.9% to 74.3% and dropped to 60.5% in patients without any comorbidities. The specificity ranged from 73.8% to 90.0%. The sensitivity decreased for all conditions to 40.4%–57.4% in patients less than 65 years of age, but remained significantly higher than for younger patients without any comorbidities (29.1%). The specificity was high across all comorbid conditions in the younger cohort. The positive predictive value was high for all groups except those aged 12–17 years.

In the sensitivity analysis restricted to those who were surveyed between February and August, the overall positive predictive value increased from 88.4% to 93.2% and the specificity increased marginally from 95.7% to 97.5% (Table 3). The improvements in both specificity and positive predictive value were seen in all subgroups. Results for patients aged less than 65 years who have risk factors for serious influenza infection are not presented owing to the presence of numerous small cells (i.e., cell size < 6 patients).

We incorporated our results into a misclassification bias adjustment sensitivity analysis to show the utility of these results when using administrative data for determining individual-level influenza immunization status for vaccine effectiveness studies. We input a sensitivity of 68.6% and a specificity of 89.9% (from Table 3). Figure 1 shows a significant underestimation of influenza vaccine effectiveness for the 2010/11 season before adjusting for the misclassification of immune status. Vaccine effectiveness increased from 42% (95% CI 29%–53%) to 68% (95% CI 61%–78%) after adjustment.

Interpretation

We found that OHIP billing claims had only moderate performance characteristics to correctly identify influenza vaccination status in Ontario, compared with self-report. For children and adults less than 65 years of age, sensitivity was under 50%, but specificity was greater than 90%. Among those aged 65 years and older, sensitivity was higher, but with lower specificity. The sensitivity was generally higher for patients with comorbid conditions and those with a regular physician. These subpopulations had the most accurate OHIP influenza vaccination billing claims. The performance characteristics were better when restricted to Canadian Community Health Survey respondents who were surveyed between February and August, suggesting the presence of some misclassification by influenza season when including respondents surveyed during months when influenza vaccines are generally given.

There are a number of potential explanations for the low sensitivity of OHIP billing claims. A substantial minority of individuals are vaccinated outside of physician offices, and we would not expect their vaccinations to be captured in health administrative data, despite the existence of an influenza vaccine tracking code. These people include those who received the vaccine at workplaces, schools or public health clinics. In 2012, pharmacists began providing vaccination, and these are captured in the Ontario Drug Benefits database, which may improve the performance of Ontario administrative data in subsequent years. In addition, remuneration per vaccination is low (ranging from $0.68 to $9.60, depending on the family practice funding model), possibly resulting in missed billings.

The lower specificity in the older population is more difficult to explain. It is possible that proportionally more older adults forget while responding to the survey that they had received the influenza vaccine that year. However, previous studies have found self-reported influenza vaccination status to be reasonably valid.[3–9] Alternatively, billing errors or medical fraud could explain a proportion of the false-positive results.

Limitations

The Canadian Community Health Survey excludes children younger than 12 years and older adults living in long-term care facilities. These are important high-risk groups to include in studies, and it is unfortunate that we are unable to quantify the validity of influenza vaccination in these groups. However, our study does characterize the validity of influenza vaccination, as captured by administrative data, in virtually all other high-risk groups.

We used self-report as the reference standard in this analysis, and although verification of responses was not possible, 8 previous validation studies comparing self-report to medical records suggest that sensitivity of self-report is high (86%–100%); both specificity and positive predictive value are more variable, but are generally lower (38%–98% and 62%–96%, respectively).[3–10] However, the specificity and positive predictive value of self-report may be artificially reduced when using medical records as the reference standard if patients can receive influenza vaccine through alternative providers (e.g., workplaces, pharmacies).

Conclusion

This study updates the performance characteristics from our previous study,[2] with a much larger sample size, more recent iterations of the Canadian Community Health Survey and a far more extensive list of risk factors for serious influenza infection defined using validated methods. We quantified the sensitivity, specificity, positive predictive value and negative predictive value across a variety of variables, including

Table 2: Performance measures of Ontario Health Insurance Plan physician billing claims compared with self-reported influenza vaccination using Canadian Community Health Survey data

Characteristic	Immunity, no.				Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
	TP	FP	FN	TN				
Total	9303	1218	9395	27 385	49.8 (49.0–50.5)	95.7 (95.5–96.0)	88.4 (87.8–89.0)	74.5 (74.0–74.9)
Canadian Community Health Survey cycle								
2007/08	6506	1023	6666	19 645	49.4 (48.5–50.2)	95.1 (94.8–95.3)	86.4 (85.6–87.2)	74.7 (74.1–75.2)
2009/10	2797	195	2729	7740	50.6 (49.3–51.9)	97.5 (97.2–97.9)	93.5 (92.6–94.4)	73.9 (73.1–74.8)
Age group, yr								
12–17	205	71	803	2992	20.3 (17.9–22.8)	97.7 (97.1–98.2)	74.3 (69.1–79.4)	78.8 (77.5–80.1)
18–49	1510	275	3272	15 190	31.6 (30.3–32.9)	98.2 (98.0–98.4)	84.6 (82.9–86.3)	82.3 (81.7–82.8)
50–64	2319	263	2947	6285	44.0 (42.7–45.4)	96.0 (95.5–96.5)	89.8 (88.6–91.0)	68.1 (67.1–69.0)
≥ 65	5269	609	2373	2918	68.9 (67.9–70.0)	82.7 (81.5–84.0)	89.6 (88.9–90.4)	55.2 (53.8–56.5)
Sex								
Female	5536	708	5448	14 211	50.4 (49.5–51.3)	95.3 (94.9–95.6)	88.7 (87.9–89.4)	72.3 (71.7–72.9)
Male	3767	510	3947	13 174	48.8 (47.7–49.9)	96.3 (96.0–96.6)	88.1 (87.1–89.0)	76.9 (76.3–77.6)
Residence								
Urban	7474	1016	7229	21 687	50.8 (50.0–51.6)	95.5 (95.3–95.8)	88.0 (87.3–88.7)	75.0 (74.5–75.5)
Rural	1812	201	2152	5638	45.7 (44.2–47.3)	96.6 (96.1–97.0)	90.0 (88.7–91.3)	72.4 (71.4–73.4)
Has a regular doctor (patient age, yr)								
Yes (< 65)	3924	585	6430	21 505	37.9 (37.0–38.8)	97.4 (97.1–97.6)	87.0 (86.0–88.0)	77.0 (76.5–77.5)
Yes (≥ 65)	5205	600	2229	2632	70.0 (69.0–71.1)	81.4 (80.1–82.8)	89.7 (88.9–90.4)	54.1 (52.7–55.5)
No (< 65)	110	24	592	2962	15.7 (13.0–18.4)	99.2 (98.9–99.5)	82.1 (75.6–88.6)	83.3 (82.1–84.6)
No (≥ 65)	64	9	144	286	30.8 (24.5–37.0)	96.9 (95.0–98.9)	87.7 (80.1–95.2)	66.5 (62.1–71.0)
Risk factors for serious influenza infections (age ≥ 65 yr)								
Hypertension	4012	448	1589	1671	71.6 (70.4–72.8)	78.9 (77.1–80.6)	90.0 (89.1–90.8)	51.3 (49.5–53.0)
Asthma	651	63	270	244	70.7 (67.7–73.6)	79.5 (75.0–84.0)	91.2 (89.1–93.3)	47.5 (43.2–51.8)
Diabetes	1 40	159	544	554	71.1 (69.1–73.2)	77.7 (74.6–80.8)	89.4 (87.8–91.0)	50.5 (47.5–53.4)
Cancer	858	81	342	359	71.5 (68.9–74.1)	81.6 (78.0–85.2)	91.4 (89.6–93.2)	51.2 (47.5–54.9)
COPD	586	55	265	234	68.9 (65.7–72.0)	81.0 (76.4–85.5)	91.4 (89.3–93.6)	46.9 (42.5–51.3)
Congestive heart failure	571	51	202	178	73.9 (70.8–77.0)	77.7 (72.3–83.1)	91.8 (89.6–94.0)	46.8 (41.8–51.9)
Myocardial infarction	329	34	130	120	71.7 (67.6–75.8)	77.9 (71.4–84.5)	90.6 (87.6–93.6)	48.0 (41.8–54.2)
Chronic kidney disease	323	36	134	116	70.7 (66.5–74.9)	76.3 (69.6–83.1)	90.0 (86.9–93.1)	46.4 (40.2–52.6)
Morbid obesity*	67	< 6	< 30	45	72.0 (62.9–81.2)	90.0 (81.7–98.3)	93.1 (87.2–98.9)	63.4 (52.2–74.6)
Stroke	334	39	124	137	72.9 (68.9–77.0)	77.8 (71.7–84.0)	89.5 (86.4–92.6)	52.5 (46.4–58.5)
Immunosuppression	337	25	133	123	71.7 (67.6–75.8)	83.1 (77.1–89.1)	93.1 (90.5–95.7)	48.0 (41.9–54.2)
Dementia	107	21	37	59	74.3 (67.2–81.4)	73.8 (64.1–83.4)	83.6 (77.2–90.0)	61.5 (51.7–71.2)
No risk factors	733	95	478	856	60.5 (57.8–63.3)	90.0 (88.1–91.9)	88.5 (86.4–90.7)	64.2 (61.6–66.7)
Risk factors for serious influenza infections (age < 65 yr)								
Hypertension	1460	176	1471	2999	49.8 (48.0–51.6)	94.5 (93.7–95.3)	89.2 (87.7–90.7)	67.1 (65.7–68.5)
Asthma	734	87	1084	3092	40.4 (38.1–42.6)	97.3 (96.7–97.8)	89.4 (87.3–91.5)	74.0 (72.7–75.4)
Diabetes	649	76	572	983	53.2 (50.4–56.0)	92.8 (91.3–94.4)	89.5 (87.3–91.7)	63.2 (60.8–65.6)
Cancer	215	19	243	459	46.9 (42.4–51.5)	96.0 (94.3–97.8)	91.9 (88.4–95.4)	65.4 (61.9–68.9)
COPD	152	21	122	161	55.5 (49.6–61.4)	88.5 (83.8–93.1)	87.9 (83.0–92.7)	56.9 (51.1–62.7)
Congestive heart failure	81	11	60	81	57.4 (49.3–65.6)	88.0 (81.4–94.7)	88.0 (81.4–94.7)	57.4 (49.3–65.6)
Myocardial infarction	94	10	93	118	50.3 (43.1–57.4)	92.2 (87.5–96.8)	90.4 (84.7–96.1)	55.9 (49.2–62.6)
Chronic kidney disease	81	< 6	< 75	120	52.9 (45.0–60.9)	96.8 (93.7–99.9)	95.3 (90.9–99.8)	62.5 (55.7–69.3)
Morbid obesity*	144	27	141	411	50.5 (44.7–56.3)	93.8 (91.6–96.1)	84.2 (78.7–89.7)	74.5 (70.8–78.1)
Stroke	63	< 6	< 50	113	57.3 (48.0–66.5)	96.6 (93.3–99.9)	94.0 (88.4–99.7)	70.6 (63.6–77.7)
Pregnancy†	67	13	93	565	41.9 (34.2–49.5)	97.8 (96.5–99.0)	83.8 (75.7–91.8)	85.9 (83.2–88.5)
No risk factors	1725	321	4209	17 447	29.1 (27.9–30.2)	98.2 (98.0–98.4)	84.3 (82.7–85.9)	80.6 (80.0–81.1)

Note: CI = confidence interval, COPD = chronic obstructive pulmonary disease, FN = false negative, FP = false positive, NPV = negative predictive value, PPV = positive predictive value, TN = true negative, TP = true positive.
*Body mass index > 40.
†Date of delivery between Nov. 1 and June 1.

multiple high-risk influenza groups. These results can be used to correct for underdetermination of vaccine coverage levels at the aggregate level, and to account for misclassification bias of vaccination status at the individual level (e.g., in studies of influenza vaccine effectiveness). We have shown the importance of quantifying misclassification bias, with substantial underestimation of influenza vaccine effectiveness when using OHIP physician billing claims data to determine vaccination status. Nondifferential misclassification is generally expected to bias results toward the null hypothesis and thus underestimate effect sizes. However, this may not always be true; therefore it is important to quantify the degree of systematic error in observational studies.[30] In addition, the high positive predictive value and spec-

Table 3: Performance measures of Ontario Health Insurance Plan physician billing claims compared with self-reported influenza vaccination using Canadian Community Health Survey data, restricted to patients surveyed between Feb. 1 and Aug. 31

Characteristic	Immunity, no.				Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
	TP	FP	FN	TN				
Total	5639	408	5744	16 017	49.5 (48.6–50.5)	97.5 (97.3–97.8)	93.2 (92.6–93.9)	73.6 (73.0–74.2)
Canadian Community Health Survey cycle								
2007/08	3745	276	3831	10 414	49.4 (48.3–50.6)	97.4 (97.1–97.7)	93.1 (92.4–93.9)	73.1 (72.4–73.8)
2009/10	1894	132	1913	5603	49.8 (48.2–51.3)	97.7 (97.3–98.1)	93.5 (92.4–94.6)	74.5 (73.6–75.5)
Age group, yr								
12–17	126	34	507	1745	19.9 (16.8–23.0)	98.1 (97.5–98.7)	78.8 (72.4–85.1)	77.5 (75.8–79.2)
18–49	960	112	2025	9060	32.2 (30.5–33.8)	98.8 (98.6–99.0)	89.6 (87.7–91.4)	81.7 (81.0–82.5)
50–64	1401	83	1770	3610	44.2 (42.5–45.9)	97.8 (97.3–98.2)	94.4 (93.2–95.6)	67.1 (65.8–68.4)
≥ 65	3152	179	1442	1602	68.6 (67.3–70.0)	89.9 (88.6–91.3)	94.6 (93.9–95.4)	52.6 (50.9–54.4)
Sex								
Female	3345	221	3320	8226	50.2 (49.0–51.4)	97.4 (97.0–97.7)	93.8 (93.0–94.6)	71.2 (70.4–72.1)
Male	2294	187	2424	7791	48.6 (47.2–50.0)	97.7 (97.3–98.0)	92.5 (91.4–93.5)	76.3 (75.4–77.1)
Residence								
Urban	4485	348	4360	12 628	50.7 (49.7–51.7)	97.3 (97.0–97.6)	92.8 (92.1–93.5)	74.3 (73.7–75.0)
Rural	1142	59	1376	3 357	45.4 (43.4–47.3)	98.3 (97.8–98.7)	95.1 (93.9–96.3)	70.9 (69.6–72.2)
Has a regular doctor (patient age, yr)								
Yes (< 65)	2424	218	3941	12 705	38.1 (36.9–39.3)	98.3 (98.1–98.5)	91.7 (90.7–92.8)	76.3 (75.7–77.0)
Yes (≥ 65)	3119	176	1349	1441	69.8 (68.5–71.2)	89.1 (87.6–90.6)	94.7 (93.9–95.4)	51.6 (49.8–53.5)
No (< 65)	63	11	361	1710	14.9 (11.5–18.2)	99.4 (99.0–99.7)	85.1 (77.0–93.2)	82.6 (80.9–84.2)
No (≥ 65)	33	< 6	< 95	161	26.2 (18.5–33.9)	98.2 (96.1–100.0)	91.7 (82.6–100.0)	63.4 (57.5–69.3)
Risk factors for serious influenza infections (age ≥ 65 yr)								
Hypertension	2403	121	977	896	71.0 (69.6–72.6)	88.1 (86.1–90.1)	95.2 (94.4–96.0)	47.8 (45.6–50.1)
Asthma	368	18	159	134	69.8 (65.9–73.7)	88.2 (83.0–93.3)	95.3 (93.2–97.4)	45.7 (40.0–51.4)
Diabetes	808	52	332	293	70.9 (68.2–73.5)	84.9 (81.2–88.7)	94.0 (92.4–95.5)	46.9 (43.0–50.8)
Cancer	494	19	196	186	71.6 (68.2–75.0)	90.7 (86.8–94.7)	96.3 (94.7–97.9)	48.7 (43.7–53.7)
COPD	348	11	164	128	68.0 (63.9–72.0)	92.1 (87.6–96.6)	96.9 (95.2–98.7)	43.8 (38.1–49.5)
Congestive heart failure	345	10	117	94	74.7 (70.7–78.6)	90.4 (84.7–96.1)	97.2 (95.5–98.9)	44.5 (37.8–51.3)
Myocardial infarction	204	12	78	66	72.3 (67.1–77.6)	84.6 (76.6–92.6)	94.4 (91.4–97.5)	45.8 (37.7–54.0)
Chronic kidney disease	187	10	81	69	69.8 (64.3–75.3)	87.3 (80.0–94.7)	94.9 (91.9–98.0)	46.0 (38.0–54.0)
Morbid obesity*	40	0	13	26	75.5 (63.9–87.1)	100.0 (100.0–100.0)	100.0 (100.0–100.0)	66.7 (51.9–81.5)
Stroke	191	14	79	70	70.7 (65.3–76.2)	83.3 (75.4–91.3)	93.2 (89.7–96.6)	47.0 (39.0–55.0)
Immunosuppression	194	7	84	63	69.8 (64.4–75.2)	90.0 (83.0–97.0)	96.5 (94.0–99.0)	42.9 (34.9–50.9)
Dementia	59	11	19	31	75.6 (66.1–85.2)	73.8 (60.5–87.1)	84.3 (75.8–92.8)	62.0 (48.5–75.5)
No risk factors	415	35	276	473	60.1 (56.4–63.7)	93.1 (90.9–95.3)	92.2 (89.7–94.7)	63.2 (59.7–66.6)

Note: CI = confidence interval, COPD = chronic obstructive pulmonary disease, FN = false negative, FP = false positive, NPV = negative predictive value, PPV = positive predictive value, TN = true negative, TP = true positive.
*Body mass index > 40.

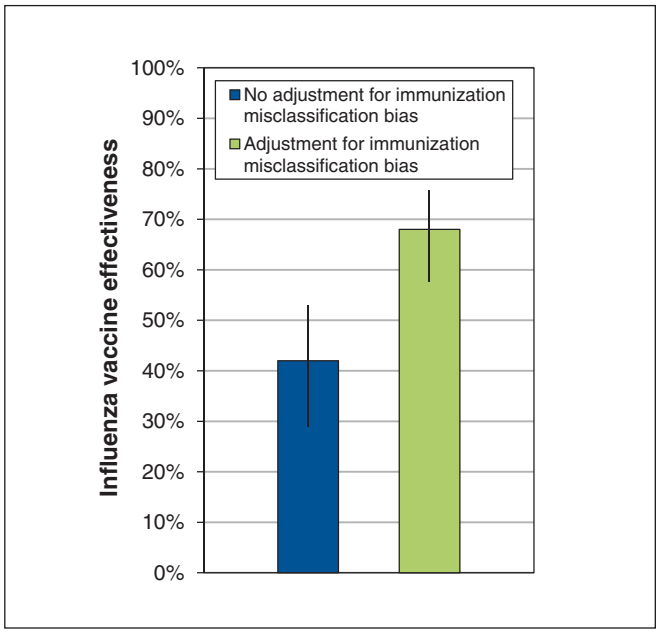

Figure 1: Influenza vaccine effectiveness estimates for the 2010/11 season, with 95% confidence intervals, from Kwong and associates[31] before and after adjustment for misclassification bias using the validation results from this study.

ificity suggests that the database can accurately identify patients who have truly undergone vaccination, allowing these data to be used to study influenza vaccine safety using self-controlled study designs.[32]

In the absence of a vaccination registry in Ontario, administrative data represent the best available data source to study influenza vaccines at a population level. However, we fully support the creation of a registry in Ontario to permit optimal evaluations of our publicly funded vaccination programs, particularly since vaccinations given at public health and workplace clinics are not captured by physician billing claims data. Despite the limitations of administrative data, the results of this study will enable adjustments for systematic error in future studies.

References

1. Kuster SP, Shah PS, Coleman BL, et al. Incidence of influenza in healthy adults and healthcare workers: a systematic review and meta-analysis. *PLoS One* 2011;6:e26239.
2. Kwong JC, Manuel DG. Using OHIP physician billing claims to ascertain individual influenza vaccination status. *Vaccine* 2007;25:1270-4.
3. Hutchison BG. Measurement of influenza vaccination status of the elderly by mailed questionnaire: response rate, validity and cost. *Can J Public Health* 1989;80:271-5.
4. Mac Donald R, Baken L, Nelson A, et al. Validation of self-report of influenza and pneumococcal vaccination status in elderly outpatients. *Am J Prev Med* 1999;16:173-7.
5. Martin LM, Leff M, Calonge N, et al. Validation of self-reported chronic conditions and health services in a managed care population. *Am J Prev Med* 2000;18:215-8.
6. Zimmerman RK, Raymund M, Janosky JE, et al. Sensitivity and specificity of patient self-report of influenza and pneumococcal polysaccharide vaccinations among elderly outpatients in diverse patient care strata. *Vaccine* 2003;21:1486-91.
7. Skull SA, Andrews RM, Byrnes GB, et al. Validity of self-reported influenza and pneumococcal vaccination status among a cohort of hospitalized elderly inpatients. *Vaccine* 2007;25:4775-83.
8. Rolnick SJ, Parker ED, Nordin JD, et al. Self-report compared to electronic medical record across eight adult vaccines: do results vary by demographic factors? *Vaccine* 2013;31:3928-35.
9. Mangtani P, Shah A, Roberts JA. Validation of influenza and pneumococcal vaccine status in adults based on self-report. *Epidemiol Infect* 2007;135:139-43.
10. Irving SA, Donahue JG, Shay DK, et al. Evaluation of self-reported and registry-based influenza vaccination status in a Wisconsin cohort. *Vaccine* 2009;27:6546-9.
11. Béland Y. Canadian community health survey–methodological overview. *Health Rep* 2002;13:9-14.
12. Chan B. Supply of physicians' services in Ontario. *Hosp Q* 1999-2000;3:17.
13. Fleet JL, Dixon SN, Shariff SZ, et al. Detecting chronic kidney disease in population-based administrative databases using an algorithm of hospital encounter and physician claim codes. *BMC Nephrol* 2013;14:81.
14. Quinn RR, Laupacis A, Austin PC, et al. Using administrative datasets to study outcomes in dialysis patients: a validation study. *Med Care* 2010; 48:745-50.
15. Lam NN, McArthur E, Kim SJ, et al. Validation of kidney transplantation using administrative data. *Can J Kidney Health Dis* 2015;2:20.
16. Hux JE, Ivis F, Flintoft V, et al. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes Care* 2002;25:512-6.
17. Starfield B, Weiner J, Mumford L, et al. Ambulatory care groups: a categorization of diagnoses for research and management. *Health Serv Res* 1991; 26:53-74.
18. Hall S, Schulze K, Groome P, et al. Using cancer registry data for survival studies: the example of the Ontario Cancer Registry. *J Clin Epidemiol* 2006;59:67-76.
19. Austin PC, Daly PA, Tu JV. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. *Am Heart J* 2002;144:290-6.
20. Schultz SE, Rothwell DM, Chen Z, et al. Identifying cases of congestive heart failure from administrative data: a validation study using primary care patient records. *Chronic Dis Inj Can* 2013;33:160-6.
21. Gill SS, Anderson GM, Fischer HD, et al. Syncope and its consequences in patients with dementia receiving cholinesterase inhibitors: a population-based cohort study. *Arch Intern Med* 2009;169:867-73.
22. Juurlink D, Preyra C, Croxford R, et al. *Canadian Institute for Health Information Discharge Abstract Database: a validation study*. Toronto: Institute for Clinical Evaluative Sciences; 2006.
23. Gurevich Y, McFarlane A, Morris K, et al. Estimating the number of coronary artery bypass graft and percutaneous coronary intervention procedures in Canada: a comparison of cardiac registry and Canadian Institute for Health Information data sources. *Can J Cardiol* 2010;26:e249-53.
24. Lee DS, Tu JV, Austin PC, et al. Effect of cardiac and noncardiac conditions on survival after defibrillator implantation. *J Am Coll Cardiol* 2007;49: 2408-15.
25. Gershon AS, Wang C, Guan J, et al. Identifying patients with physician-diagnosed asthma in health administrative databases. *Can Respir J* 2009;16:183-8.
26. Gershon AS, Wang C, Guan J, et al. Identifying individuals with physcian diagnosed COPD in health administrative databases. *COPD* 2009;6:388-94.
27. Jaakkimainen RL, Bronskill SE, Tierney MC, et al. Identification of physician-diagnosed Alzheimer's disease and related dementias in population-based administrative data: a validation study using family physicians' electronic medical records. *J Alzheimer Dis*. In press.
28. Antoniou T, Zagorski B, Loutfy MR, et al. Validation of case-finding algorithms derived from administrative data for identifying adults living with human immunodeficiency virus infection. *PLoS One* 2011;6:e21748.
29. Tu K, Campbell NR, Chen ZL, et al. Accuracy of administrative databases in identifying patients with hypertension. *Open Med* 2007;1: e18-26.
30. Fox MP, Lash TL, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. *Int J Epidemiol* 2005;34:1370-6.
31. Kwong JC, Campitelli MA, Gubbay JB, et al. Vaccine effectiveness against laboratory-confirmed influenza hospitalizations among elderly adults during the 2010-2011 season. *Clin Infect Dis* 2013;57:820-7.
32. Kwong JC, Vasa PP, Campitelli MA, et al. Risk of Guillain-Barré syndrome after seasonal influenza vaccination and influenza health-care encounters: a self-controlled study. *Lancet Infect Dis* 2013;13:769-76.

Affiliations: Institute for Clinical Evaluative Sciences (Schwartz, Jembere, Campitelli, Chung, Kwong); Institute of Health, Policy, Management, and Evaluation (Schwartz), University of Toronto; Dalla Lana School of Public Health (Buchan, Kwong), University of Toronto; Public Health Ontario (Kwong), Toronto Ont.

Contributors: All of the authors contributed substantially to the conception, design, data analysis plan and interpretation of the data. In addition, Nathaniel Jembere performed the data analysis and Kevin Schwartz drafted the manuscript. All of the authors critically reviewed the manuscript for important intellectual content, approved the final version to be published and agreed to act as guarantors of the work.

Acknowledgements: This work was supported by the Canadian Institutes of Health Research (CIHR) (grant no. MOP 130568). Kevin Schwartz is supported by a Fellowship Award from CIHR. Jeffrey Kwong is supported by a New Investigator Award from CIHR and a Clinician Scientist Award from the University of Toronto Department of Family and Community Medicine. The study sponsors did not participate in: the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; or the decision to submit the manuscript for publication.

Disclaimer: Parts of this material are based on data and information compiled and provided by the Canadian Institute for Health Information (CIHI) and Cancer Care Ontario (CCO). However, the analyses, conclusions, opinions and statements expressed herein are those of the authors, and not necessarily those of CIHI or CCO. This study was supported by the Institute for Clinical Evaluative Sciences (ICES) which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results and conclusions reported in this article are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intented or should be inferred.

Supplemental information: For reviewer comments and the original submission of this manuscript, please see www.cmajopen.ca/content/4/3/E463/suppl/DC1