

Validation of infant immunization billing codes in administrative data

Kevin L Schwartz^{1,2}, Karen Tu^{1,2,3,4}, Laura Wing¹, Michael A Campitelli¹, Natasha S Crowcroft^{5,6,7}, Shelley L Deeks^{5,7}, Sarah E Wilson^{5,7}, Kumanan Wilson^{1,8}, Ian Gemmill⁹, and Jeffrey C Kwong^{1,3,4,5,7,*}

¹Institute for Clinical Evaluative Sciences; Toronto, Ontario Canada; ²Institute of Health Policy Management and Evaluation; University of Toronto; Toronto, Ontario, Canada; ³Department of Family & Community Medicine; University of Toronto; Toronto, Ontario, Canada; ⁴University Health Network; Toronto, Ontario, Canada; ⁵Dalla Lana School of Public Health; University of Toronto; Toronto, Ontario, Canada; ⁶Department of Laboratory Medicine & Pathobiology; University of Toronto; Toronto, Ontario, Canada; ⁷Public Health Ontario; Toronto, Ontario, Canada; ⁸Ottawa Hospital Research Institute; University of Ottawa; Ottawa, Ontario, Canada; ⁹Kingston, Frontenac, and Lennox & Addington Public Health; Kingston, Ontario, Canada

Keywords: diphtheria-tetanus-pertussis vaccine, health services, meningococcal vaccines, measles-mumps-rubella vaccine, pneumococcal vaccines, vaccine combined, validation studies

Abbreviations: EMRALD, Electronic Medical Record Administrative data Linked Database; OHIP, Ontario Health Insurance Plan; DTaP-IPV-Hib, diphtheria, tetanus, acellular pertussis, inactivated polio, *Haemophilus influenzae* type B vaccine; PC, pneumococcal conjugate vaccine; MMR, measles, mumps, rubella vaccine; MenC, meningococcal conjugate serogroup C vaccine; PPV, positive predictive value; NPV, negative predictive value; EMR, electronic medical records; ICES, Institute for Clinical Evaluative Sciences; COC, continuity of care; CIHI-DAD, Canadian Institute of Health Information Discharge Abstract Database; CIC, Citizen and Immigration Canada; RPDB, Registered Persons Database

Ontario has a single payer provincial health insurance program. Administrative data may provide a potentially robust source of information for post-marketing vaccine studies. Vaccine-specific immunization billing codes were introduced in 2011. Our objective was to validate Ontario's universal health care administrative datasets to assess infant immunization status. Electronic medical record data from the Electronic Medical Record Administrative data Linked Database (EMRALD) was used as the reference standard to calculate performance characteristics of the Ontario Health Insurance Plan (OHIP) database vaccine-specific and general immunization codes for 4 primary infant immunizations: diphtheria, tetanus, acellular pertussis, inactivated polio, *Haemophilus influenzae* type B (DTaP-IPV-Hib) combination vaccine, pneumococcal conjugate vaccine, measles, mumps, rubella (MMR) vaccine, and meningococcal conjugate serogroup C vaccine. OHIP billing claims had specificity ranging from 81% to 92%, sensitivity 70% to 83%, positive predictive value (PPV) 97% to 99%, and negative predictive value (NPV) 13% to 46% for identifying the various specific vaccines in administrative data. For cohorts vaccinated in the new code introduction phase, using both the vaccine-specific and general codes had higher sensitivity than the vaccine-specific codes alone. In conclusion, immunization billing claims from administrative data in Ontario had high specificity and PPV, moderate sensitivity, and low NPV. This study identifies some of the applications of utilizing administrative data for post-marketing vaccine studies. However, limitations of these data decrease their utility for measuring vaccine coverage and effectiveness. Therefore, the establishment of a comprehensive and linkable immunization registry should be a provincial priority.

Introduction

Vaccines are one of the greatest public health achievements of the last century.¹ Public confidence in vaccines is important for ensuring continued success of this achievement. Therefore, enhanced post-marketing vaccine effectiveness, coverage, and safety surveillance is a critical component of any immunization program.²⁻⁴

In Ontario, Canada's largest province (population 13.5 million with 142,448 births in 2013),⁵ infant and toddler immunizations

are almost exclusively administered through physician offices and funded under the Ontario Health Insurance Plan (OHIP).⁶ Measles, mumps, rubella (MMR) and meningococcal conjugate serogroup C (MenC) vaccines are recommended at 12 months of age, while diphtheria, tetanus, acellular pertussis, inactivated polio, *Haemophilus influenzae* type B (DTaP-IPV-Hib) combination vaccine and pneumococcal conjugate vaccine (PC) are given at 2, 4, 6, and 18 months, and at 2, 4, and 12 months, respectively. In addition, rotavirus vaccine is recommended at 2 and 4 months of age, varicella at

© Kevin L Schwartz, Karen Tu, Laura Wing, Michael A Campitelli, Natasha S Crowcroft, Shelley L Deeks, Sarah E Wilson, Kumanan Wilson, Ian Gemmill, and Jeffrey C Kwong

*Correspondence to: Jeffrey C Kwong; Email: jeff.kwong@utoronto.ca

Submitted: 01/13/2015; Revised: 04/02/2015; Accepted: 04/16/2015

<http://dx.doi.org/10.1080/21645515.2015.1043499>

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

15 months of age, and booster doses of MMR-varicella vaccine and DTaP-IPV are currently recommended between 4 to 6 y of age.

Physicians submit billing claims to OHIP for providing immunizations, which are captured in the OHIP database. Prior to 2011, physicians submitted claims using general immunization codes for any vaccine administered. In September 2011, 9 vaccine-specific codes were introduced. Due to the lack of a comprehensive centralized immunization registry that includes pre-school aged children, administrative data sets are a potential data source to assess vaccine coverage as they contain individual-level data spanning the whole province. However, the validity of using both the newer vaccine-specific codes and the earlier general immunization codes recorded in the administrative data is unknown.

Across Canada, the use of electronic medical record (EMR) systems by family physicians has been increasing.⁷ The objective of this study was to utilize primary care EMR data recorded in family physician offices to validate OHIP physician billing claims for general and vaccine-specific immunization codes.

Results

We initially identified 2204 children for the MMR/MenC 2009 birth cohort, 1823 for the DTaP-IPV-Hib/PC 2010 birth cohort, 3329 children for the MMR/MenC 2011 birth cohort, and 2661 for the DTaP-IPV-Hib/PC 2012 birth cohort. Application of the exclusion criteria left final cohorts of 1604, 1557, 2404, and 2194, respectively (Fig. 1). These cohorts were drawn from the practices of 147, 167, 246, and 257 family physicians, respectively.

Characteristics of the study children, mothers, and physicians were summarized and compared to the entire Ontario 2011 birth cohort in Table 1. Compared to the Ontario population, rural patients are overrepresented in the Electronic Medical Record Administrative data Linked Database (EMRALD); 10% Ontario infants live in a rural setting compared to 21% in EMRALD. The immigrant population is underrepresented in EMRALD and physicians using this EMR system are more likely to have graduated in the last 1–2 decades and less likely to be foreign trained, compared to all Ontario physicians (Table 1).

The performance measures of OHIP immunization codes were relatively similar across all 4 vaccines (Table 2). For the general immunization codes, sensitivity was 79.4%–83.2%, specificity 81.0%–92.2%, PPV 96.8%–99.5%, and NPV 18.9%–46.4% for the various vaccines. For the vaccine-specific codes, the sensitivity was 70.1%–72.4%, specificity 88.5%–91.5%, PPV 96.9–99.4%, and NPV 12.9%–38.8%. Comparing the performance of the vaccine-specific codes (post-2011) with the general vaccine codes (pre-2011), overall sensitivity declined. Specificity decreased for DTaP-IPV-Hib and PC but increased for MMR and MenC. PPV was unchanged, and NPV decreased.

The performance measures of OHIP immunization codes were relatively similar across all 4 vaccines (Table 2). For the general immunization codes, sensitivity was 79.4%–83.2%, specificity 81.0%–92.2%, PPV 96.8%–99.5%, and NPV 18.9%–46.4% for the various vaccines. For the vaccine-specific codes, the sensitivity was 70.1%–72.4%, specificity 88.5%–91.5%, PPV 96.9–99.4%, and NPV 12.9%–38.8%. Comparing the performance of the vaccine-specific codes (post-2011) with the general vaccine codes (pre-2011), overall sensitivity declined. Specificity decreased for DTaP-IPV-Hib and PC but increased for MMR and MenC. PPV was unchanged, and NPV decreased.

In the sensitivity analysis that included both the general and vaccine-specific

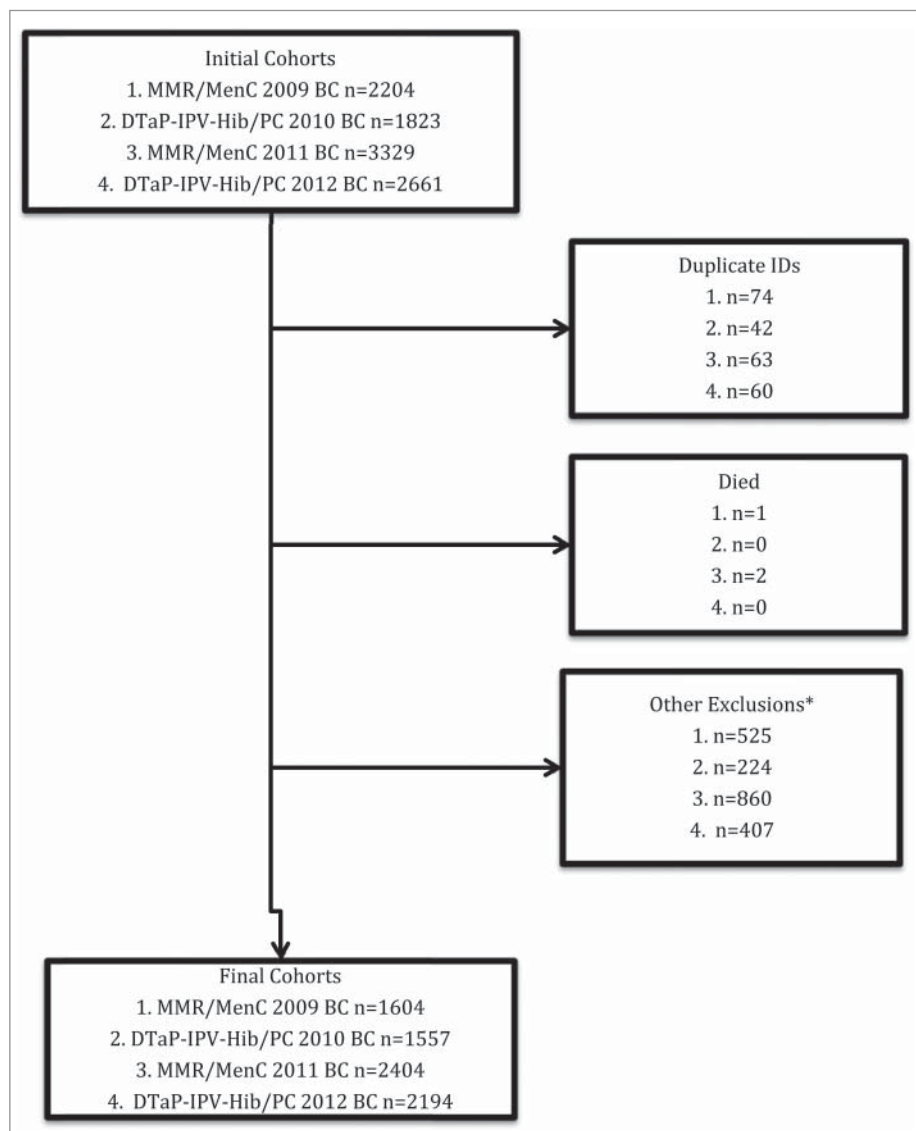


Figure 1. Flow diagram describing the application of the exclusion criteria to the 4 cohorts used in the analysis. MMR = measles, mumps, rubella vaccine; MenC = Meningococcal C vaccine; DTaP-IPV-Hib = Diphtheria, tetanus, acellular pertussis, inactivated polio, Haemophilus influenzae type b vaccine; PC = pneumococcal conjugate vaccine; BC = birth cohort; ID = Identification number *No billed primary care assessments or fewer than 4 visits to their family physician during the first year of life.

Table 1. Child, maternal, and physician characteristics of the combined four cohorts used for validation analysis compared to the 2011 Ontario population

Characteristic	All Ontario 2011 birth cohort N = 132,433	All study birth cohorts N = 7,759	p-value
<i>Child characteristics</i>			
Male	67,922 (48.7%)	4,010 (51.7%)	0.50
Rural residence	13,524 (10.2%)	1,610 (20.8%)	<0.001
Neighbourhood income quintile			<0.001
1 (lowest)	27,544 (20.8%)	1,427 (18.4%)	
2	25,649 (19.4%)	1,463 (18.9%)	
3	26,388 (19.9%)	1,624 (20.9%)	
4	28,162 (21.3%)	1,735 (22.4%)	
5 (highest)	21,866 (16.5%)	1,397 (18.0%)	
Any chronic medical condition	3,670 (2.8%)	137 (1.8%)	<0.001
Low birth weight (<2500 grams)	8,580 (6.5%)	392 (5.1%)	<0.001
No. primary care visits in the first year of life, mean \pm SD	11.3 \pm 6.1	11.3 \pm 4.6	0.91
Low continuity of care*	20,372 (15.4%)	1,358 (17.5%)	<0.001
<i>Maternal characteristics</i>			
Mother is a recent immigrant (<5 years)	14,483 (10.9%)	529 (6.8%)	<0.001
Age at first delivery			<0.001
Younger than 20 years	11,948 (9.0%)	586 (7.6%)	
20–34 years	105,546 (79.7%)	5,812 (74.9%)	
35 years or older	14,939 (11.3%)	1,020 (13.1%)	
Influenza vaccination in year following delivery	16,540 (12.5%)	1,228 (15.8%)	<0.001
Primiparous	66,019 (49.9%)	3,900 (50.3%)	<0.001
<i>Physician characteristics</i>			
Male	72,521 (54.8%)	3,127 (40.3%)	<0.001
Rural practice	9,380 (7.1%)	1,264 (16.3%)	<0.001
Decade of graduation from medical school			<0.001
1970s	31,186 (23.5%)	1,054 (13.6%)	
1980s	40,092 (30.3%)	1,468 (18.9%)	
1990s	34,252 (25.9%)	1,828 (23.6%)	
2000s	19,905 (15.0%)	3,202 (41.3%)	
Foreign-trained	46,917 (35.4%)	611 (7.9%)	<0.001
No. patients aged \leq 6 years on July 1, 2011, mean \pm SD	128.7 \pm 91.7	126.9 \pm 96.5	0.09

*<50% of primary care visits made to EMR physician.

immunization codes post-2011, we observed a marginal improvement in sensitivity for all vaccines, with modest reductions in specificity compared to using the vaccine-specific codes alone (Table 2).

We identified predictors of discordance between immunization billing claims and documentation in the EMR (Table 3). Children with lower continuity of care scores tended to have higher rates of discordant results ($p < 0.05$). Physicians who saw proportionally more young children in their practice had higher concordance ($p < 0.05$). Maternal influenza immunization status also correlated with concordance ($p < 0.05$); however, the magnitude was small. Predictors of discordance for each vaccine and cohort can be seen in Appendix Tables 2A–G.

We evaluated performance measures stratifying by COC and pediatric roster size (Table 4). As an example, the sensitivity of MMR in the 2009 birth cohort with high COC was 85.4% compared to 69.1% with low COC. Specificity was 92.4% and 43.8% in the high and low COC groups, respectively.

Discussion

By utilizing EMR patient charts as the reference standard for immunization status, we identified that OHIP billing claims

have high specificity and PPV, moderate sensitivity, and low NPV in establishing immunization status of children in Ontario. Newer, vaccine-specific codes have worse sensitivity than previous general immunization codes.

The most likely explanation for the moderate sensitivity of OHIP claims is that physicians do not bill because remuneration is small. Ontario's schedule of benefits reimburses CDN \$4.50 for each immunization (the fee is independent of the vaccine given), however depending on the practice remuneration model, a physician may receive as little as \$0.68 per vaccine billing claim. Furthermore, primary care reform has been a major initiative in Ontario that has led to the creation of multidisciplinary team practices.⁸ In this model, more vaccines are administered by nursing staff, so physicians may not always submit billing claims for these services. This factor may partially explain the drop in sensitivity from 2009 and 2010 to 2011 and 2012 that persisted in the sensitivity analysis, using both general and vaccine-specific codes. This sensitivity analysis also revealed minimal change from the primary vaccine-specific codes analysis, confirming that most EMERALD physicians have adopted these new codes. However, multiple vaccines are administered at the 2-month visit. Documentation of the vaccine of interest (e.g., DTaP-IPV-Hib) corresponding to a general OHIP billing code resulted in

Table 2. Evaluating Ontario Health Insurance Plan physician billing claims performance characteristics with 95% confidence intervals compared with electronic medical records as the reference standard

	Cell Counts				Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
	TP	FP	FN	TN				
General immunization codes								
DTaP-IPV-Hib 2010 birth cohort	1175	6	305	71	79.4 (77.2–81.4)	92.2 (83.8–97.1)	99.5 (98.9–99.8)	18.9 (15.1–23.2)
PC 2010 birth cohort	1174	7	302	74	79.5 (77.4–81.6)	91.4 (83.0–96.5)	99.4 (98.8–99.8)	19.7 (15.8–24.1)
MMR 2009 birth cohort	1161	39	238	166	83.0 (80.9–84.9)	81.0 (74.9–86.1)	96.8 (95.6–97.7)	41.1 (36.3–46.1)
MenC 2009 birth cohort	1162	38	234	170	83.2 (81.2–85.2)	81.7 (75.8–86.7)	96.8 (95.7–97.8)	46.4 (41.2–51.5)
Vaccine-specific codes								
DTaP-IPV-Hib 2012 birth cohort	1500	11	595	88	71.6 (69.7–73.5)	88.9 (82.7–95.1)	99.3 (98.8–99.7)	12.9 (10.4–15.4)
PC 2012 birth cohort	1511	9	577	97	72.4 (70.5–74.3)	91.5 (86.2–96.8)	99.4 (99.0–99.8)	14.4 (11.7–17.0)
MMR 2011 birth cohort	1411	45	601	347	70.1 (68.1–72.1)	88.5 (85.4–91.7)	96.9 (96.0–97.8)	36.6 (33.5–39.7)
MenC 2011 birth cohort	1389	44	594	377	70.1 (68.0–72.1)	89.6 (86.6–92.5)	96.9 (96.0–97.8)	38.8 (35.8–41.9)
Sensitivity analysis: general and vaccine-specific codes								
DTaP-IPV-Hib 2012 birth cohort	1569	16	526	83	74.9 (73.0–76.8)	83.8 (76.6–91.1)	99.0 (98.5–99.5)	13.6 (10.9–16.4)
PC 2012 birth cohort	1577	14	511	92	75.5 (73.9–77.4)	86.8 (80.4–93.2)	99.1 (98.7–99.6)	15.3 (12.4–18.1)
MMR 2011 birth cohort	1493	59	519	333	74.2 (72.3–76.1)	85.0 (81.4–88.5)	96.2 (95.3–97.2)	39.1 (35.8–42.4)
MenC 2011 birth cohort	1469	58	514	363	74.1 (72.1–76.0)	86.2 (82.9–89.5)	96.2 (95.2–97.2)	41.4 (38.1–44.7)

TP = True Positive (immunization in OHIP database and EMRALD).

FP = False Positive (immunization in OHIP database but not in EMRALD).

FN = False Negative (immunization in EMRALD but not in OHIP database).

TN = True Negative (no vaccination in OHIP database or EMRALD).

DTaP-IPV-Hib = diphtheria, tetanus, acellular pertussis, inactivated polio, and *Haemophilus influenzae* type b vaccine; PC = pneumococcal conjugate vaccine; MMR = measles, mumps, rubella vaccine; MenC = meningococcal conjugate serogroup C vaccine; OHIP = Ontario Health Insurance Plan; EMR = Electronic Medical Record; PPV = positive predictive value; NPV = negative predictive value; CI = confidence interval.

classification as a true positive. But since the general code can be used for any vaccine, there is a small chance of misclassification and this may have inflated the sensitivity and PPV.

Immunizations of young children are mostly administered through physician offices in Ontario, and it is unlikely that children who are seeing an EMRALD-participating physician on a regular basis (at least 4 times during the first year of life) would receive their vaccines in a walk-in-clinic or other office besides their primary care physician. This assumption is exemplified by the very high specificity; however a small percentage of children may have been vaccinated elsewhere or by salaried physicians.

Important limitations of using an EMR as the reference standard for immunizations should be noted. EMR records are dependent on the clinician documenting administration of the vaccine. There are multiple areas of free text within the EMR, and while we attempted to capture all relevant text entries, missed recordings are possible. The data available through EMRALD are a voluntary sample of Ontarian physicians who all use one type of EMR system and practice under some type of primary care reform model of care, and therefore may not be entirely representative of all physicians in the province. Patients captured by EMRALD are more likely to live in rural areas and to be cared for by younger physicians who have adopted EMR systems.

This study provides important information on the validity of using province-wide administrative data for assessing immunization status, particularly given the absence of a comprehensive all-of-life vaccine registry in Ontario. These findings will help to support a number of future applications. First, immunization status in the OHIP database can be linked to health

services utilization to study potential adverse effects of specific vaccines better.² The excellent PPV makes this an optimal data source for such studies, because they often use self-controlled designs that require inclusion only of individuals who have had both the exposure (i.e., immunization) and the outcome of interest. The high specificity and PPV values for these immunization codes indicate that they lend themselves well to these analyses.

The second application is to assess immunization coverage across different jurisdictions and populations, in particular coverage in early infancy and at the crucial milestone of 2 y of age. Ontario's immunization coverage estimates, derived from its centralized repository, are limited to immunization records received following school entry (generally at age 4 to 6 years). However, the moderate sensitivity seen in this study suggests that administrative data will underestimate vaccine coverage in the population.

Finally, by linking immunization data to microbiologic results, we can evaluate vaccine effectiveness on a population level. This analysis is becoming increasingly important with outbreaks of vaccine-preventable diseases such as measles⁹ and pertussis.¹⁰ This application may however underestimate vaccine effectiveness because of the non-differential misclassification bias in immunization status that we have demonstrated using administrative data.¹¹

EMR data contained more immunizations than administrative data, although some records observed in the administrative data were not recorded in the EMR. Therefore a provincial immunization registry drawn from combining EMR data with

Table 3. Predictors of discordant results for the 2-month dose of diphtheria, tetanus, acellular pertussis, inactivated polio, and *Haemophilus influenzae* type b vaccine using Ontario Health Insurance Plan vaccine-specific billing claims for the 2012 birth cohort

Characteristic	DTaP-IPV-Hib vaccine		p-value
	Discordant	Concordant	
Total	606 (27.6%)	1,588 (72.4%)	
Child characteristics			
Male	322 (53.1%)	792 (49.9%)	0.17
Rural residence	91 (15.0%)	272 (17.1%)	0.30
Income quintile			0.30
1 (lowest)	130 (21.5%)	278 (17.5%)	
2	108 (17.8%)	288 (18.1%)	
3	122 (20.1%)	333 (21.0%)	
4	135 (22.3%)	351 (22.1%)	
5 (highest)	97 (16.0%)	288 (18.1%)	
Any chronic medical condition	14 (2.3%)	24 (1.5%)	0.20
Low birth weight (<2500 grams)	31 (5.1%)	83 (5.2%)	0.46
No. primary care visits in the first year of life, mean ± SD	10.91 ± 4.2	11.45 ± 4.4	0.009
Low continuity of care [^]	156 (25.7%)	258 (16.2%)	<0.001
Maternal characteristics			
Mother is a recent immigrant (<5 years)	47 (7.8%)	121 (7.6%)	0.06
Maternal age at first delivery			0.004
Younger than 20 years	48 (7.9%)	119 (7.5%)	
20–34 years	424 (70.0%)	1,214 (76.4%)	
35 years or older	111 (18.3%)	223 (14.0%)	
Maternal immunization in year following delivery	62 (10.2%)	267 (16.8%)	<0.001
Primiparous	346 (57.1%)	820 (51.6%)	0.001
Physician characteristics			
Male physician	214 (35.3%)	599 (37.7%)	0.49
Physician rural practice	85 (14.0%)	212 (13.4%)	0.68
Decade of medical school graduation			<0.001
1970s	54 (8.9%)	183 (11.5%)	
1980s	138 (22.8%)	286 (18.0%)	
1990s	87 (14.4%)	359 (22.6%)	
2000s	297 (49.0%)	726 (45.7%)	
Foreign-trained	38 (6.3%)	134 (8.4%)	0.09
No. patients aged ≤6 years on July 1, 2011, mean ± SD	96.3 ± 90.5	142.6 ± 122.0	<0.001

[^]<50% of primary care visits made to EMR physician.

Note: Not all variables add up to 100% due to missing data.

administrative data may provide a more complete picture of immunization coverage. Several approaches could be undertaken to improve the quality of EMRALD data. ICES currently provides semi-annual feedback to physicians using EMRALD in the form of quality indicators related to several chronic diseases, including diabetes and ischemic heart disease.¹² Expanding this feedback to include up-to-date immunization status for each vaccine preventable disease may provide additional incentive to physicians using EMRs to improve data quality and/or coverage. Providing continuing education to clinicians on methods to improve the completeness and accuracy of this data source may enhance the utility of these data for research purposes. We strongly encourage further expansion of efforts to link EMR data to administrative data, in order to increase our knowledge about coverage of the population until a comprehensive provincial immunization registry is established.

In summary, OHIP's new vaccine-specific (and older general immunization) billing codes have high specificity and PPV, but only moderate sensitivity and low NPV for predicting immunization status of children in Ontario. OHIP billing claims are a

potentially rich source of information for post-marketing vaccine safety surveillance. We plan to use the results of this analysis to evaluate waning immunity from pertussis immunization as well as measles vaccine effectiveness. These results will also be used in ongoing studies of vaccine safety and to evaluate immunization coverage in infants. However, establishment of a comprehensive vaccine registry that can be linked to administrative data should be a priority, given the limitations of this dataset.

Materials and Methods

We conducted a validation study of vaccine billing codes submitted by physicians compared to the reference standard of documentation in a primary care EMR.

Study population and setting

We included pre-specified birth cohorts of Ontario children under the care of family physicians who share their practices' EMR data with EMRALD, a centralized repository of EMR data

Table 4. Performance measures for select birth cohorts stratified by continuity of care and physician pediatric roster size

	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Strata				
<i>MMR (2009 birth cohort)</i>				
Continuity of care				
Low	69.1 (62.8–75.5)	43.8 (29.7–57.8)	83.9 (78.4–89.5)	25.0 (15.7–34.3)
High	85.4 (83.4–87.4)	92.4 (88.2–96.5)	98.8 (98.2–99.5)	45.3 (39.9–50.8)
Roster size				
<120 patients younger than 6 years	77.4 (74.2–80.6)	75.2 (67.0–83.5)	95.0 (93.1–96.9)	35.4 (29.2–41.7)
≥120 patients younger than 6 years	87.7 (85.3–90.0)	87.0 (80.4–93.6)	98.1 (97.1–99.1)	48.1 (40.8–55.3)
<i>MMR (2011 birth cohort)</i>				
Continuity of care [^]				
Low	56.5 (51.1–61.9)	67.7 (58.2–77.2)	85.9 (81.2–90.5)	31.0 (24.7–37.4)
High	72.7 (70.6–74.9)	95.0 (92.5–97.5)	98.8 (98.2–99.4)	38.1 (34.6–41.6)
Roster size				
<120 patients younger than 6 years	62.0 (59.1–65.0)	85.4 (80.3–90.5)	96.0 (94.6–97.5)	28.4 (24.6–32.1)
≥120 patients younger than 6 years	79.0 (76.4–81.6)	91.3 (87.5–95.1)	97.7 (96.6–98.7)	48.3 (43.4–53.3)
<i>DTaP-IPV-Hib (2010 birth cohort)</i>				
Continuity of care				
Low	63.0 (57.1–68.9)	70.0 (50.0–90.1)	96.4 (93.6–99.2)	12.8 (6.6–19.1)
High	82.8 (80.7–84.9)	100.0 (100.0–100.0)	100.0 (100.0–100.0)	21.4 (16.4–26.3)
Roster size				
<120 patients younger than 6 years	74.1 (71.0–77.3)	94.9 (88.0–100.0)	99.6 (99.2–100.0)	16.0 (11.2–20.7)
≥120 patients younger than 6 years	84.9 (82.2–87.5)	89.5 (79.7–99.2)	99.4 (98.7–99.9)	23.6 (16.7–30.6)
<i>DTaP-IPV-Hib (2012 birth cohort)</i>				
Continuity of care				
Low	61.8 (57.0–66.7)	67.7 (51.9–83.4)	95.5 (93.0–98.1)	13.7 (8.5–18.9)
High	73.8 (71.7–75.8)	100.0 (100.0–100.0)	100.0 (100.0–100.0)	12.6 (9.8–15.5)
Roster size				
<120 patients younger than 6 years	65.0 (62.3–67.7)	84.6 (75.8–93.4)	98.8 (98.0–99.5)	11.4 (8.5–14.2)
≥120 patients younger than 6 years	80.9 (78.3–83.5)	97.1 (91.4–100.0)	99.9 (99.6–100.0)	16.6 (11.4–21.8)

[^]<50% of primary care visits made to EMR physician.

PPV = positive predictive value, NPV = negative predictive value, MMR = measles mumps rubella vaccine, DTaP-IPV-Hib = Diphtheria tetanus acellular pertussis inactivated polio *Haemophilus influenzae* type b vaccine.

used for research and evaluation.¹³ EMERALD has been previously used for administrative data validation studies to identify patients with a variety of disease conditions.^{14–16}

We created 4 cohorts of children. For the MMR and MenC vaccines, we enrolled children born between January 1 2009 to December 31 2009 and January 1 2011 to December 31 2011 to evaluate the general and vaccine-specific immunization codes, respectively. Similarly, for the first DTaP-IPV-Hib and PC vaccine doses, we enrolled children born between January 1 2010 to December 31 2010 and January 1 2012 to December 31 2012 to evaluate the general and vaccine-specific immunization codes, respectively. We selected these study periods to validate the general codes before the introduction of the vaccine-specific codes in September 2011, and the vaccine-specific codes after that date, allowing for a period of several months for physicians to adjust to using the newer codes.

We excluded children who had duplicate identification numbers (preventing 1:1 linkage between EMERALD and the OHIP database) and those who died during the observation period. We also excluded children with no billed primary care assessments, and those with fewer than 4 visits to their family physician during the first year of life. These children were likely to have resided in Ontario intermittently or received their primary care from a salaried physician who did not submit billing claims to OHIP.

Data source for immunization reference standard

EMERALD currently comprises EMR data from over 350 Ontario family physicians who use Practice Solutions Suite[®] software, the most widely used EMR in Ontario.¹⁷ Individual-level data from EMERALD are collected semi-annually and linked to other administrative databases at the Institute for Clinical Evaluative Sciences (ICES). All clinically relevant information is collected in EMERALD, including clinical encounters, the cumulative patient profile, family history, allergies, immunizations, diagnostic tests, prescriptions, discharge summaries, and consultation notes. Participating physicians contribute to EMERALD on a voluntary basis. This data set was used as the reference standard for immunization status.

We conducted text searches for each vaccine of interest in the prescriptions and immunizations fields in EMERALD using a series of keywords (e.g., “MMR,” “measles, mumps, and rubella vaccine;” full list in **Supplemental Table 1**). We also searched EMERALD for records of immunization billing claims submitted through the EMR software.

We used different observation windows depending on the vaccine of interest. Since both MMR and MenC vaccines are scheduled at 12 months of age, the observation window spanned 335 d to 455 d after birth (approximately 11–15 months of age) to capture children vaccinated before or after the recommended

age. We limited the validation of the DTaP-IPV-Hib and PC vaccines to the 2-month dose (i.e., the first dose in the series). The observation window was 53 d (2 months less 1 week) to 112 d (4 months less 1 week) after birth. We used the same observation windows to evaluate the general immunization codes.

Data source for immunization codes

The OHIP database contains physician billing claims paid for by OHIP, which covers virtually all of Ontario's approximately 13.5 million residents, except for recent immigrants and migrants (i.e., residing in Ontario for <3 months) as well as a very small percentage of patients who see non-billing physicians.¹⁸ We used these data to identify infant immunization codes. We searched for physician billing claims for MMR vaccine (G845), MenC vaccine (G844), DTaP-IPV-Hib vaccine (G841), PC vaccine (G846), and general immunizations with (G538) and without (G539) physician consultation.

Data sources for covariates

The data were linked to administrative datasets housed at ICES to identify factors related to concordance between EMR and OHIP records.

The Registered Persons Database (RPDB) was used to identify the child's sex, vital status, and socio-demographic data. The child's postal code was linked to Canadian census data to attribute mean household income quintile of their neighborhood and rural residence (community size under 10 000).

The MOMBABY database is an administrative data set maintained at ICES comprised of admission records of delivering mothers and their newborn babies, which are linked through a unique matching number on each hospitalization record. We used this dataset and the Canadian Institute of Health Information's Discharge Abstract Database (CIHI-DAD) to identify the mothers of the children in our cohort and to determine whether each mother had a previous delivery (primiparous vs. not), to measure the maternal age at first parity, and to identify infants of low birth weight. The CIHI-DAD was also used to identify chronic medical conditions among children, as described by Feudtner et al.,¹⁹ within the first year of life.

The OHIP database was used to identify the number of primary care visits in the first year of life and to calculate a continuity of care (COC) score. COC is defined as the number of visits to an individual's primary care physician divided by the total number of physician visits during the first year of life. We defined a COC of less than 50% as low.^{20,21} We also used the OHIP database to identify maternal influenza immunization status during the year following delivery using the influenza-specific billing codes G590 (influenza immunization with a physician consultation) and G591 (influenza immunization without a physician consultation).

The Citizen and Immigration Canada (CIC) database contains information on individuals who have landed in Ontario since 1985, and we used this data set to ascertain maternal immigration status. Mothers were considered recent immigrants if they arrived to Canada within the previous 5 y

The ICES Physician Database contains information on physician demographics and specialization. We obtained the following physician covariates: sex, rural practice, decade of graduation, and place of medical training (foreign vs. domestic).

The Client Agency Program Enrolment (CAPE) and OHIP databases were used to determine the volume of patients aged 6 y or younger on a physician's roster.

Analysis

EMERALD immunization status was set as the reference standard for immunization, and we calculated performance measures of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for OHIP immunization codes for each vaccine, during each time period (i.e., before and after introduction of the vaccine-specific codes). In a sensitivity analysis, we analyzed the 2011 birth cohort for the MMR and MenC vaccines and the 2012 birth cohort for DTaP-IPV-Hib and PC vaccines to include both general and vaccine-specific immunization codes. We also examined discordant results between EMERALD and OHIP to ascertain the nature of the discordance. We stratified selected variables to demonstrate performance measures within each stratum.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Funding

This work was supported by the Ontario Ministry of Health and Long-Term Care (MOHLTC). This study was also supported by the Institute for Clinical Evaluative Sciences (ICES) and Public Health Ontario (PHO), which are funded by annual grants from the MOHLTC. KS is supported by a Fellowship Award from the Canadian Institutes of Health Research (CIHR). KT is supported by a Clinician Scientist Award from the University of Toronto Department of Family and Community Medicine (DFCM). JCK is supported by a New Investigator Award from CIHR and a Clinician Scientist Award from the University of Toronto DFCM. 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

The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by ICES, PHO, or MOHLTC is intended or should be inferred.

Supplemental Material

Supplemental data for this article can be accessed on the publisher's website.

References

1. The Canadian Public Health Association: 12 great achievements. Accessed January 12, 2015. Available from: <http://www.cpha.ca/en/programs/history/achievements.aspx>.
2. Wilson K, Hawken S, Kwong JC, Deeks SL, Crowcroft NS, Manuel D. Vaccine and immunization surveillance in ontario (VISION) - using linked health administrative databases to monitor vaccine safety. *Vaccine* 2012; 30:6115-20; PMID: 22709951; <http://dx.doi.org/10.1016/j.vaccine.2012.06.004>
3. Chen RT, Glasser JW, Rhodes PH, Davis RL, Barlow WE, Thompson RS, Mullooly JP, Black SB, Shinefield HR, Vadheim CM, et al. Vaccine Safety Datalink project: a new tool for improving vaccine safety monitoring in the United States. The Vaccine Safety Datalink team. *Pediatrics* 1997; 99:765-73; PMID:9164767; <http://dx.doi.org/10.1542/peds.99.6.765>
4. Lankinen KS, Pastila S, Kilpi T, Nohynek H, Makela PH, Olin P. Vaccinovigilance in Europe—need for timeliness, standardization and resources. *Bull W H O* 2004; 82:828-35; PMID:15640918
5. Statistics Canada: Accessed November 18, 2014. Available from: <http://www.statcan.gc.ca/tables-tableaux/sum-som/101/cst01/demo02a-eng.htm>.
6. Publicly funded immunization schedule in Ontario (2011): Accessed May 30 2014. Available from: <http://www.health.gov.on.ca/en/public/programs/immunization/docs/schedule.pdf>.
7. College of Family Physicians of Canada CMA, Royal College of Physicians and Surgeons of Canada. National physician survey 2014. Results 2014. [accessed 2014 December 11] Available from: <http://nationalphysiciansurvey.ca/surveys/2014-survey/survey-results-2/>
8. Hutchison B, Glazier R. Ontario's primary care reforms have transformed the local care landscape, but a plan is needed for ongoing improvement. *Health Aff (Millwood)* 2013; 32:695-703; PMID:23569049; <http://dx.doi.org/10.1377/hlthaff.2012.1087>
9. Bird C. Measles outbreaks threaten those averse to vaccine. *Canad Med Associat J* 2013; 185:E393-E4; PMID:23649422; <http://dx.doi.org/10.1503/cmaj.109-4480>
10. Waters V, Jamieson F, Richardson SE, Finkelstein M, Wormsbecker A, Halperin SA. Outbreak of atypical pertussis detected by polymerase chain reaction in immunized preschool-aged children. *Pediatr Infect Dis J* 2009; 28:582-7; PMID:19561423; <http://dx.doi.org/10.1097/INF.0b013e318197fac1>
11. Ozasa K. The effect of misclassification on evaluating the effectiveness of influenza vaccines. *Vaccine* 2008; 26:6462-5; PMID:18573297; <http://dx.doi.org/10.1016/j.vaccine.2008.06.039>
12. Ivers NM, Tu K, Young J, Francis JJ, Barnsley J, Shah BR, Upshur RE, Moineddin R, Grimshaw JM, Zwarenstein M. Feedback GAP: pragmatic, cluster-randomized trial of goal setting and action plans to increase the effectiveness of audit and feedback interventions in primary care. *Implement Sci* 2013; 8:142; PMID:24341511; <http://dx.doi.org/10.1186/1748-5908-8-142>
13. Tu K, Mitiku TF, Ivers NM, Guo H, Lu H, Jaakkimainen L, Kavanagh DG, Lee DS, Tu JV. Evaluation of electronic medical record administrative data linked database (EMRALD). *Am J Manag Care* 2014; 20:e15-21; PMID:24669409.
14. Tu K, Mitiku T, Lee DS, Guo H, Tu JV. Validation of physician billing and hospitalization data to identify patients with ischemic heart disease using data from the Electronic Medical Record Administrative data Linked Database (EMRALD). *Can J Cardiol* 2010; 26:e225-8; PMID:20847968; [http://dx.doi.org/10.1016/S0828-282X\(10\)70412-8](http://dx.doi.org/10.1016/S0828-282X(10)70412-8)
15. Tu K, Wang M, Young J, Green D, Ivers NM, Butt D, Jaakkimainen L, Kapral MK. Validity of administrative data for identifying patients who have had a stroke or transient ischemic attack using EMRALD as a reference standard. *Can J Cardiol* 2013; 29:1388-94; PMID:24075778; <http://dx.doi.org/10.1016/j.cjca.2013.07.676>
16. Tu K, Wang M, Jaakkimainen RL, Butt D, Ivers NM, Young J, Green D, Jette N. Assessing the validity of using administrative data to identify patients with epilepsy. *Epilepsia* 2014; 55:335-43; PMID:24417710; <http://dx.doi.org/10.1111/epi.12506>
17. Ontario MD. EMR Enhanced Use Program. [accessed 2014 December 11]; Available from: https://http://www.ontariomd.ca/portal/server.pt/community/emr_funding/emr_enhanced_use_program/.
18. Chan B. Supply of physicians' services in ontario. *Hosp Q* 1999; 3:17; PMID:11570359
19. Feudtner C, Christakis DA, Connell FA. Pediatric deaths attributable to complex chronic conditions: a population-based study of Washington State, 1980-1997. *Pediatrics* 2000; 106:205-9; PMID:10888693
20. Campitelli MA, Inoue M, Calzavara AJ, Kwong JC, Guttman A. Low rates of influenza immunization in young children under Ontario's universal influenza immunization program. *Pediatrics* 2012; 129:e1421-30; PMID:22585770; <http://dx.doi.org/10.1542/peds.2011-2441>
21. Guttman A, Manuel D, Dick PT, To T, Lam K, Stukel TA. Volume matters: physician practice characteristics and immunization coverage among young children insured through a universal health plan. *Pediatrics* 2006; 117:595-602; PMID:16510636; <http://dx.doi.org/10.1542/peds.2004-2784>