Limitations of administrative data to identify measles cases in Ontario, Canada: a cautionary tale

Caitlin Johnson¹ · Cynthia Chen^{1,2} · Laura Rosella^{1,2,3} · Heather Rilkoff¹ · Alex Marchand-Austin¹ · Jonathan B. Gubbay^{1,4} · Tina Kozlowski¹ · Shelley L. Deeks^{1,3} · Tony Mazzulli¹ · Natasha Crowcroft^{1,3,4}

Received: 13 December 2016 / Accepted: 28 October 2017 / Published online: 26 February 2018
© Crown 2018

Abstract

Objective To explore the utility of laboratory data and determine the validity of healthcare administrative data for describing the frequency of measles in Ontario.

Methods We linked Ontario healthcare utilization administrative data to Public Health Ontario (PHO) laboratory data from 01 January 2006 to 30 November 2012.

Results The sensitivity of the administrative data was 54% and the positive predictive value was 1% when compared with 50 cases identified in laboratory data as a gold standard.

Conclusions As measles is no longer endemic in Ontario, the high number of measles-coded healthcare visits found in Ontario health administration data exceeds the true number of cases. Great caution should be taken in using administrative data to calculate the incidence of measles in areas where it has been eliminated.

Résumé

Objectif Explorer l'utilité des données de laboratoire et déterminer la validité des données administratives sur les soins de santé pour décrire la fréquence de la rougeole en Ontario.

Méthode Nous avons maillé les données administratives sur l'utilisation des soins de santé en Ontario aux données des laboratoires de Santé publique Ontario (SPO) du 1^{er} janvier 2006 au 30 novembre 2012.

Résultats La sensibilité des données administratives était de 54 %, et leur valeur prédictive positive était de 1 % lorsque nous les avons comparées aux 50 cas indiqués comme étant l'étalon de référence dans les données de laboratoire.

Conclusions Comme la rougeole n'est plus endémique en Ontario, le nombre élevé de visites médicales codées comme étant liées à la rougeole dans les données administratives sanitaires de la province dépasse le nombre réel de cas. Il convient donc de faire preuve d'une grande prudence en utilisant les données administratives pour calculer l'incidence de la rougeole dans les régions où elle a été éliminée.

Keywords Administrative data · Healthcare · Burden · Measles · Ontario

Mots-clés Données administratives · Soins de santé · Charge de morbidité · Rougeole · Ontario

✉ Caitlin Johnson
Caitlin.johnson@oahpp.ca

¹ Public Health Ontario, 480 University Avenue, Suite 300, Toronto, ON M5G 1V2, Canada

² Institute for Clinical Evaluative Sciences, Toronto, ON, Canada

³ Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

⁴ Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada

Introduction

Measles has been eliminated in Canada, with the last endemic case reported in 1997 (Public Health Agency of Canada 2011). Measles cases in Canada occur through importation of the virus from regions where measles is still endemic. While there is no sustained endemic transmission of measles in Ontario (Lim et al. 2014), importations and limited local transmission still require resources for healthcare and public health response (Orenstein and Hinman 2012). Understanding the frequency

of measles is important to inform and evaluate effective prevention and control measures, such as immunization programs and public health outbreak response. Measles is legally notifiable by physicians and laboratories in Ontario, Canada (population 13.8 million in 2015) (Statistics Canada 2015). Cases are reported to public health authorities but reports may miss details of healthcare burden including disease severity and hospitalization. Laboratory data usually do not include such information either, and physician reports are often also incomplete at the time of reporting to public health. Comparisons of measles surveillance data with administrative data have been done in other settings (Takla et al. 2014); however, at the time of this study, linkage of the surveillance data collected by public health authorities with health administrative data was not permitted. We therefore linked Ontario healthcare utilization administrative data to Public Health Ontario (PHO) laboratory data from 01 January 2006 to 30 November 2012 with the objectives to enhance the laboratory data and to validate the use of administrative data for describing the frequency of measles in Ontario.

Methods

We applied the Ontario case definition of a confirmed measles case that requires laboratory confirmation of infection with clinically compatible symptoms in the absence of recent measles immunization (within 5–42 days of sample collection) (Ontario Ministry of Health and Long Term Care 2013). Laboratory confirmation is defined as a positive viral culture, positive polymerase chain reaction (PCR), or seroconversion (a significant (fourfold or greater) rise in measles IgG titer between acute and convalescent sera). In addition, a symptomatic person can be designated as a confirmed case if they test positive for IgM antibody and are epidemiologically linked to a laboratory-confirmed case or have travelled recently to an area with known measles activity.

In Ontario, all testing for measles is done at the PHO laboratory. We extracted all PHO laboratory data that had a positive viral culture, a positive PCR, and/or a positive IgM from the PHO laboratory information management system from 01 January 2006 to 30 November 2012. Due to restriction of our laboratory information system, we were unable to extract seroconversion (IgG) data. Therefore, a case was considered laboratory-confirmed if they had at least one positive culture or PCR result. Positive culture or PCR tests are routinely genotyped, which can discern if the detected virus is a vaccine-related strain. Genotyping information was available for PHO tests done after 2010, allowing us to exclude persons who were found to have a positive test due to vaccine strain rather than a wild-type virus. As the PHO data do not contain reliable information about travel history or epidemiological links to other laboratory-confirmed cases, we excluded individuals with only a positive IgM test result. We therefore limited our

laboratory-confirmed cases to persons who had a positive viral culture or PCR test result not known to be from a vaccine genotype.

Several sources of health administrative data are available and held at the Institute for Clinical Evaluative Sciences (ICES), including the Ontario Health Insurance Plan (OHIP) claims database, which captures physician billings, and the Canadian Institute for Health Information's (CIHI) National Ambulatory Care Reporting System (NACRS) and the CIHI Discharge Abstract Database (DAD), which capture hospitalization and emergency department visits. We included physician office visits from the OHIP database with the diagnostic code for measles (055), emergency visits from CIHI-NACRS, and hospitalizations from the CIHI-DAD Alldx with any diagnostic code for measles (International Classification of Diseases Tenth Revision [ICD-10] codes: B05, B05.0 to B05.4, B05.8, B05.81, B05.89, B05.9).

We linked the PHO data to the health administrative data using deterministic methods using Ontario health card number. In the absence of a valid Ontario health card number, we used probabilistic linkage using name, age, date of birth, sex, postal code, and health unit of residence. Following linkage, we made further effort to exclude cases who may have received a positive laboratory result as a consequence of shedding virus due to a recent immunization. To do this, we searched the physician billing claims for MMR vaccine (G845) and general immunizations with (G538) and without (G539) physician consultation. Persons who had one or more of the mentioned vaccination codes within 60 days prior to their laboratory diagnosis were excluded from our analysis.

Cases were considered true positives if they had at least one administrative measles-specific diagnostic code 7 days before or after their laboratory diagnoses. Laboratory diagnosed cases not associated with a clinical measles diagnosis code within this time window were considered to be false negatives. False-positive cases were individuals who had at least one measles-specific clinical diagnostic code during the study period in the absence of a positive culture or PCR laboratory diagnosis. Measurement of the sensitivity of the administrative data provides an estimate of its ability to correctly capture measles cases. Calculating an accurate idea of healthcare burden of measles requires the sensitivity of the administrative data to be high. The proportion of the cases with measles codes in the administrative data that actually have measles speaks to its positive predictive value. As the prevalence of measles is so low in Ontario, we expected the positive predictive value of the administrative data to be low.

The sensitivity and positive predictive value of the administrative data were calculated overall and by data source compared with the PHO data. We were unable to calculate specificity, as we only had information about positive laboratory tests.

We also examined the average length of stay of any measles-related hospitalizations in this group. The healthcare encounters of laboratory-confirmed cases without a measles clinical diagnosis code (the false negatives) were explored so that we might understand if they did have symptoms consistent with a measles infection.

Results

We used PHO data to determine that six confirmed cases had tested positive for the vaccine-related genotype and thus were excluded. This resulted in 62 laboratory-confirmed measles cases identified. The majority of these (58 cases, 94%) were linked to administrative data; however, four cases (6%) were missing information that was required for linkage. Eight persons were excluded as they had a recent immunization-related visit in the 60 days prior to the laboratory diagnosis. This resulted in 50 laboratory-confirmed measles cases linked to administrative data. There were 3636 individuals that had measles-related healthcare visits that were not identified in the laboratory data. Of these 3636 individuals, 1053 (29%) had a physician visit related to immunization within 60 days prior to these measles-related healthcare encounters and were subsequently excluded (Fig. 1).

The sensitivity of the health administrative data with all sources combined was 54% when compared with the laboratory data (Table 1). Therefore, approximately half of the laboratory-confirmed measles cases were captured by the administrative data (true-positive cases), with the remaining half designated as false negatives, as they did have a corresponding laboratory

measles diagnosis from the laboratory that tests all measles cases in the province. The positive predictive value of the administrative data was low at 1.0% when compared with the laboratory-confirmed cases. While OHIP diagnostic billing is the most sensitive of the three data sources (38%), it also had the poorest positive predictive value (0.8%).

We found that nine of the 50 laboratory-confirmed cases were hospitalized with a measles-specific B05 ICD-10 Alldx code. These hospitalizations also contained codes for fever (R509), isolation (Z290), dehydration (A099), unspecified erythematous condition (A86), rash (E860), gastroenteritis (B09), pneumonia (B349), unspecified viral infection with skin and mucous membrane lesions (L539), unspecified viral infection (R21), and unspecified viral encephalitis (J189). The length of stay of these hospitalizations ranged from 1 to 6 days, with a mean of 3 days. The ages of persons hospitalized for measles ranged from 1 to 59 years, with a median of 37 years old.

We explored the healthcare encounters of the laboratory-confirmed cases without a measles code 7 days before and after their measles-positive culture or PCR test. We found that 13 of the 23 false-negative individuals accessed healthcare with respiratory, gastrointestinal, and other symptoms that could have been consistent with an active measles infection, but not coded as such (Table 2).

Discussion

Although measles is eliminated in Ontario, healthcare system and public health resources are still required for imported

Fig. 1 Flow diagram of Ontario laboratory and administrative measles data from January 1, 2006 to November 30, 2012

Table 1 Ontario measles administrative data performance characteristics compared with laboratory diagnosis as the reference standard

Data source	Cell counts			Sensitivity (%)	PPV (%)
	TP	FN	FP		
All administrative data sources	27	23	2583	54	1.0
Inpatient hospitalizations (CIHI-DAD)	9	41	16	18	36.0
Emergency visits (CIHI-NACRS)	10	40	257	20	3.7
Physician billing (OHIP)	19	31	2378	38	0.8

TP, true positive (measles case in laboratory data and in administrative data); *FN*, false negative (measles case in laboratory data but not in administrative data); *FP*, false positive (measles case in administrative data but not in laboratory data); *PPV*, positive predictive value (proportion of persons with a code who actually have measles)

cases and to prevent the local spread of infection. We have shown that only half of laboratory-confirmed measles cases were coded as measles in Ontario health administrative data. Despite not receiving a measles-specific code, several individuals with a measles laboratory confirmation accessed healthcare with symptoms that may have been related to a measles infection, such as fever, pneumonia, and rash. These false-negative cases may reflect simple miscoding or a lack of recognition of measles by clinicians in a setting where elimination has been achieved (Hutchins et al. 2004), and highlight the limitations of administrative data in describing measles incidence in these settings.

As linkage of Ontario's reportable disease database was not permitted at the time of this study, we were limited to using laboratory confirmation by culture and PCR as a reference standard. Not all measles cases are severe enough to require hospital admission, which explains the lower sensitivity of the hospitalization and emergency room data. Use of laboratory data alone meant that we were unable to identify true symptomatic measles cases confirmed through epidemiological links to laboratory-confirmed cases. Ontario reportable disease data captured 90 measles cases, 56 of which were confirmed with a positive PCR test and 34 of which were epidemiologically linked cases without laboratory confirmation

Table 2 Diagnostic codes of laboratory-confirmed PCR-positive measles cases without a measles-related code that could be related to an active measles infection ($N = 13$ persons)

Symptom type	Healthcare type	Description	Diagnostic code
Respiratory	Emergency (CIHI-NACRS)	Influenza with other respiratory manifestations, virus not identified	J111
		Acute upper respiratory infection, unspecified	J069
		Pneumonia, unspecified	J189
		Bronchitis, not specified as acute or chronic	J40
	Physician office (OHIP)	Acute nasopharyngitis, common cold	460
		Pneumonia—all types	486
		Influenza	487
		Other chronic obstructive pulmonary disease	496
		Other disorders of skin and subcutaneous tissue	709
Gastrointestinal, eye, ear	Physician office (OHIP)	Diarrhea, gastroenteritis, viral gastroenteritis	009
		Other disorders of the eye	379
		Serous otitis media, eustachian tube disorders	381
		Other disorders of skin and subcutaneous tissue	709
		Other disorders of skin and subcutaneous tissue	709
Infection—general	Hospitalizations (CIHI-DAD)	Rotaviral enteritis	A080
		<i>E. coli</i> as the cause of diseases classified	B962
		Urinary Tract infection	N390
	Emergency (CIHI-NACRS)	Viral infection, unspecified	B349
		Fever, unspecified	R509
	Physician office (OHIP)	Other viral diseases	079
		Streptococcal sore throat, scarlet fever	034
		Septicemia, blood poisoning	038
		German measles, rubella	056
		Infectious mononucleosis, glandular fever	075

during the study period [Unpublished data, integrated Public Health Information System (iPHIS) extract, via personal communication, December 2nd, 2015]. Even if the 34 symptomatic epidemiologically linked cases were incorrectly classified as false instead of true positives, this would only be 1.3% (34/2583) of the observed false positives and, thus, does not explain the low PPV we observed. We excluded several cases with a vaccine genotype and other cases with a recent OHIP immunization billing code, but genotyping information was limited to 2010–2012. Also, immunization can be given as prophylaxis following an exposure to measles. If this post-exposure prophylaxis failed, and any individual went on to develop a true measles infection, our lack of complete genotyping information may have resulted in our classifying true cases as false-positive cases. Again, this would only explain a small number. In addition, the high number of false positives may be due to uncoded immunizations, or that persons with a fever and rash illness were recorded as having measles as part of a differential diagnosis but not confirmed.

We could not be sure whether a measles code had been assigned in error because of recent immunization rather than an active infection. A recent study has demonstrated that OHIP immunization billing claims have only moderate sensitivity and may underestimate vaccine coverage (Schwartz et al. 2015). It is therefore likely that recently immunized persons without an active measles infection were included among many false positives and may partially explain the low PPV.

Conclusions

Our results are a cautionary tale about using healthcare administrative data to understand the frequency of measles in settings where measles is eliminated. False-positive measles codes may occur when a diagnostic code for measles is recorded as the final diagnosis when it was only part of the differential diagnosis, or when a measles diagnosis code is used instead of a vaccine code. For rare diseases that are associated with common conditions such as rash/fever, a low proportion of recording errors may lead to numbers of false-positive cases that greatly exceed true cases of an eliminated

disease such as measles. This would be particularly likely to occur in primary care-based settings where such common conditions are generally managed and where immunization programs are also delivered, explaining the lowest observed PPV in the OHIP data. Our findings underline the importance of using reportable disease data for data linkage studies of rare infectious diseases because cases are validated by public health units using standard case definitions. Our study was the first in Ontario to link administrative with PHO laboratory data. We have since completed a number of successful linkage studies for other infectious disease with higher incidence, but it is important to flag that this approach is clearly not appropriate for every disease.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

- Hutchins, S. S., Papania, M. J., Amler, R., Maes, E. F., Grabowsky, M., Bromberg, K., et al. (2004). Evaluation of the measles clinical case definition. *J Infect Dis*, 189(Suppl 1), S153–S159.
- Lim, G. H., Deeks, S. L., Fediurek, J., Gubbay, J., Crowcroft, N. S. (2014). Documenting the elimination of measles, rubella and congenital rubella syndrome in Ontario: 2009–12. *Canadian Communicable Disease Report*, 40(8).
- Ontario Ministry of Health and Long Term Care. (2013). Ontario Public Health Standards Infectious Diseases Protocol. Appendix B – Provincial Case Definitions (Measles).
- Orenstein, W. A., & Hinman, A. R. (2012). Measles: the burden of preventable deaths. *Lancet*, 379(9832), 2130–2131.
- Public Health Agency of Canada. (2011). *Elimination of measles, rubella and congenital rubella syndrome in Canada*. Documentation and verification report: Public Health Agency of Canada, Ottawa.
- Schwartz, K. L., Tu, K., Wing, L., Campitelli, M. A., Crowcroft, N. S., Deeks, S. L., et al. (2015). Validation of infant immunization billing codes in administrative data. *Human Vaccines & Immunotherapeutics*, 11(7), 1840–1847.
- Statistics Canada. (2015). Table 051-0001 - Estimates of population, by age group and sex for July 1, Canada, provinces and territories, annual persons, CANSIM (database). Accessed: Dec 17, 2015.
- Takla, A., Wichmann, O., Rieck, T., & Matysiak-Klose, D. (2014). Measles incidence and reporting trends in Germany, 2007–2011. *Bulletin of the World Health Organization*, 92(10), 742–749.