BMJ Open

Using population-wide administrative and laboratory data to estimate type- and subtype-specific influenza vaccine effectiveness: a surveillance protocol

Allison Nicole Scott,[1,2] Sarah A Buchan,[3,4,5] Jeffrey C Kwong,[3,4,5,6] Steven J Drews,[7] Kimberley A Simmonds,[1,8] Lawrence W Svenson [1,8,9,10]

For numbered affiliations see end of article.

Correspondence to
Dr Allison Nicole Scott;
AScott@policywise.com

ABSTRACT

Introduction The appropriateness of using routinely collected laboratory data combined with administrative data for estimating influenza vaccine effectiveness (VE) is still being explored. This paper outlines a protocol to estimate influenza VE using linked laboratory and administrative data which could act as a companion to estimates derived from other methods.

Methods and analysis We will use the test-negative design to estimate VE for each influenza type/subtype and season. Province-wide individual-level records of positive and negative influenza tests at the Provincial Laboratory for Public Health in Alberta will be linked, by unique personal health numbers, to administrative databases and vaccination records held at the Ministry of Health in Alberta to determine covariates and influenza vaccination status, respectively. Covariates of interests include age, sex, immunocompromising chronic conditions and healthcare setting. Cases will be defined based on an individual's first positive influenza test during the season, and potential controls will be defined based on an individual's first negative influenza test during the season. One control for each case will be randomly selected based on the week the specimen was collected. We will estimate VE using multivariable logistic regression.

Ethics and dissemination Ethics approval was obtained from the University of Alberta's Health Research Ethics Board—Health Panel under study ID Pro00075997. Results will be disseminated by public health officials in Alberta.

Strengths and limitations of this study

► A strength of this protocol is that it provides timely estimation of vaccine effectiveness to assist public health in allocating resources and determining the appropriate policies and public messaging during the influenza season.

► Vaccine effectiveness estimates use a test negative design, taking advantage of linked administrative health records for the entire population.

► While many confounders are included in the vaccine effectiveness estimates, not all known confounders can be measured using administrative health data.

INTRODUCTION

Influenza is a respiratory viral disease associated with significant morbidity and mortality globally. Infections range from relatively mild presentations (eg, cough, sore throat) to severe lower respiratory tract infections (eg, pneumonia). Severe cases may be associated with hospitalisation, intensive care admission and death; young children, the elderly and individuals with chronic conditions are at highest risk of severe outcomes.[1] In Canada, rates of laboratory-confirmed influenza infections are, on average, approximately 200 cases per 100 000 population, with approximately 50% of cases occurring in patients aged ≤18 years.[2] The causative agents, influenza A (subtypes H3N2 and H1N1pdm(09)) and influenza B (Yamagata and Victoria lineages), are under strong selective pressure to mutate genetically; significant genetic changes can occur in relatively short periods of time (ie <1 year).[3]

Influenza prevention relies, in part, on annual vaccination campaigns. Selection of viral strains for inclusion in the vaccine occurs approximately 9 months prior to the onset of the influenza season; by the time the vaccines are administered, the predominant circulating strains may have mutated to the point such that the effectiveness of the vaccine has diminished or has become completely ineffective.[4 5]

Influenza vaccine effectiveness (VE) is commonly estimated using the test-negative design, a variation of the case-control design where cases and controls are selected from a pool of individuals who have been tested for influenza.[6–10] Several research groups use sentinel physician networks to recruit patients: influenza testing is performed on patients who meet a case definition for influenza-like illness, and cases and controls are

selected from that pool.[6–8] While this has become an established method, there are some limitations to using sentinel physicians. As the physicians are often volunteers, there can be bias in the geographic distribution, leading to clustering of sampling in certain areas and not others. This can lead to inaccuracies as predominant circulating influenza strains vary geographically.[7 11] Immunisation information is commonly self-reported, potentially leading to recall and social desirability biases[12]; volunteer physicians may be more likely to have strong views on influenza immunisation, potentially making it more difficult for the patient to admit to not being immunised. Finally, as these studies are labour-intensive for clinic staff, physician recruitment is often low, resulting in small sample sizes and wide confidence intervals (CI). Estimates are, therefore, typically available after the peak of the influenza season, decreasing their usefulness for public health messaging and resource and operational planning.[6–8 11]

Using administrative data and routinely collected clinical specimens for estimating VE is currently under debate.[13] VE estimates generated using linked health administrative and laboratory data in the province of Ontario have been shown to be comparable to previously published estimates.[14] There has been one published estimate of Alberta-specific VE using a sentinel surveillance system[11]; however, because of the small sample size the CI was large, ranging from 8% to 72%. Estimating VE in a large jurisdiction with near-real-time data on all influenza laboratory testing and influenza vaccination in the population has the potential to provide more precise and timely VE estimates than has previously been possible. We present a protocol to estimate influenza VE using individually linked laboratory and administrative data.

METHODS AND ANALYSIS
Study setting
Alberta is a province in Canada with a publicly funded universal healthcare system; each of the 4.25 million residents is assigned a unique personal health number (PHN) at birth or on immigration to the province.[15] The PHN is recorded each time a person accesses the healthcare system, allowing for deterministic linkage across multiple administrative data sets held by the Ministry of Health.

In 2009, influenza vaccination became universally available to all Albertans aged ≥6 months, regardless of comorbidities or other risk conditions.[16] Influenza vaccines are available at no cost to the patient at public health clinics, pharmacies, physician offices, long-term care facilities, university health centres, and workplaces. Annual vaccine campaigns begin in October, with approximately 60% of all influenza vaccinations given by the end of the second week of the campaign. While the peak of influenza activity has varied widely since 2010, the median influenza peak in Alberta is in mid-January, approximately 3 months after the vaccination campaigns begin.

Laboratory methods for influenza A and B detection and influenza A subtyping
All influenza testing in Alberta is performed at a single diagnostic lab, the Provincial Laboratory for Public Health (ProvLab) and stored in a single laboratory information system, along with test and patient identifiers. Clinical specimens (eg, nasopharyngeal swabs, nasopharyngeal aspirates, bronchoalveolar lavages) are processed at ProvLab using previously published protocols. Nucleic acid extraction utilises the easyMAG extractor and reagents (bioMerieux).[17] Nucleic acid from clinical specimens is then tested using a series of respiratory detection assays as described below. Prior to May 2017, a real-time influenza A/B reverse-transcriptase PCR was used to diagnose influenza using a protocol previously described.[18 19] After May 2017, ProvLab has been using a Luminex respiratory pathogen panel for the identification of influenza A (including subtype), influenza B and other respiratory viruses (eg, coronavirus and parainfluenza).[15] Results of the laboratory testing were imported into specific laboratory information systems depending on the testing time period.

Study design
We will use the test-negative design to estimate VE for the 2011/12–2019/20 influenza seasons. The results of all respiratory virus tests conducted at ProvLab will be sent to the Ministry of Health for deterministic linkage to health administrative databases, in order to determine eligibility for inclusion in the analysis, influenza vaccination status and the following covariates: age, sex, socioeconomic status, geographic zone of residence, history of immunocompromising comorbidities, healthcare setting (inpatient or outpatient setting) and month at the time of specimen submission. The presence of a diagnostic code for an acute respiratory illness (ARI) at the time of specimen collection will be used in a sensitivity analysis.

Isolates will be considered eligible for inclusion in the analysis if they meet all of the following criteria: a valid PHN is recorded, the isolate is not from a resident of a long-term care facility, the isolate was collected at least 4 weeks after the initiation of the public influenza vaccination programme and the isolate was collected during the influenza season, as determined using the method recommended by the WHO.[20–22]

It is important to ensure that the population has the chance to be exposed to influenza and that there is sufficient time for immunity to the vaccine to be developed. Residence in a long-term care facility will be determined via the Alberta Continuing Care Information System, which contains information on admissions and discharges from long-term care facilities.[23] PHN validity will be assessed using the Alberta Health Care Insurance Plan (AHCIP) Adjusted Population Registry, which contains records of all individuals registered for healthcare insurance.[23 24]

Individuals can have multiple laboratory tests over the course of their illness; therefore, only the first positive

influenza test during the influenza season will be used, and potential control samples will be selected from among those who only tested negative for influenza during that influenza season, using the first negative test. Cases and controls tested <14 days after vaccination will be excluded from the analysis.

Influenza vaccination status will be determined from the influenza vaccination registry. The registry combines data from four databases that record influenza vaccination events (see below).

The following administrative data sets will be used in this study.

► Alberta Health Immunization and Adverse Reaction to Immunization system (Imm/ARI) contains records of all publicly funded vaccines administered through public health, including influenza vaccines administered at mass influenza vaccination clinics, public health clinics and vaccinations administered by public health nurses in long-term care facilities. Data submission is mandatory and guidelines exist to support complete and accurate vaccination records with descriptions of each, including notes.[25 26]

► The Supplemental Enhanced Service Event (SESE) database captures physician claims for billing purposes; International Classification of Diseases, Ninth Revision (ICD-9) diagnosis codes, procedure codes (Canadian Classification of Procedures), codes indicating location of service delivery and a number of other administrative elements used to support the payment for each patient encounter.[24 27 28]

► Alberta Blue Cross (ABC) administers the pharmacist component of the universal vaccination programme. Pharmacists administering influenza vaccines through this programme submit claims to ABC for each vaccine provided; they are required to submit patient information such as PHN, date of service, name and address.

► The Pharmaceutical Information Network (PIN) database records dispensed pharmacological products, regardless of payer, including the rare instances when an influenza vaccine is purchased rather than administered through the public programme (eg, purchased by travellers prior to the launch of the public campaign). PIN captures approximately 95% of all dispense events in the province.[23]

► Provincial Vaccine Registry combines influenza vaccinations given in the province and recorded in four source databases (PIN, ABC, SESE and Imm/ARI).

► AHCIP Population Registry contains demographic variables, age, sex, socioeconomic status and geographic zone of residence. Neighbourhood-level socioeconomic status is derived from census dissemination area income quintiles using postal code.

► Morbidity and Ambulatory Care Abstracting Reporting (MACAR) system contains the International Classification of Diseases, 10th Revision, Canada (ICD-10-CA) diagnostic codes, procedure codes, the date of admission and date of discharge

for every visit to hospitals, emergency rooms and outpatient clinics.

The quality of administrative data sets in Alberta has been extensively reviewed.[29–31]

Individuals will be considered inpatients if they have at least one physician claim for inpatient services on the same day as specimen collection or if specimen collection occurred during an inpatient stay; all others will be considered outpatients. Individuals with an immunocompromising condition will be defined as those who have a diagnosis of HIV, who received an organ transplant or received oral corticosteroids (for ≥30 days), antineoplastic agents or another immunocompromising drug from a community pharmacist in the past 6 months. (online supplementary appendix 1 and 2).[32] HIV diagnosis and ARI will be determined through physician claims and MACAR. Organ transplantation will be determined using MACAR, and immunocompromising drug dispensations will be identified through PIN.

Statistical analysis

VE data will be refreshed and the analysis completed every 2 weeks until the peak of the influenza season and monthly thereafter. We will use multivariable logistic regression to estimate influenza VE as $(1 - \text{adjusted OR}) \times 100\%$ and will compare the results to historical values of VE for the predominate subtype. We will estimate VE separately by influenza season and influenza subtype (ie, A(H3N2), A(H1N1)pdm09, and influenza B).[33] When there is a large enough sample size in a particular season to provide adequate power, VE will be estimated for specific age groups such as children under the age of 5 and seniors over the age of 65. The following covariates will be included in the adjusted model, regardless of statistical significance: age, sex, socioeconomic status, geographic zone of residence, history of immunocompromising comorbidities, healthcare setting (inpatient or outpatient setting) and month of specimen submission within the influenza season. SAS V.9.4 will be used for all statistical analysis (SAS Institute). VE estimates will be compared with published estimates of VE.[6 7 11 13 34 35]

As shedding of influenza virus continues for approximately 4–5 days after symptom onset, bias can result if specimens that are collected too long after symptom onset are used.[36] Most studies use a threshold of 7 days.[37] To test the robustness of the findings, a sensitivity analysis will be performed; controls will be restricted to those specimens positive for a different respiratory virus (ie, coronavirus, human respiratory syncytial virus) (as suggested by Sullivan *et al* 2016).

A potential limitation to this study is that the samples utilised here are clinical isolates taken through the course of normal patient care, and are not from a standard case definition as is utilised in some other studies.[12] To test the robustness of the findings, the analysis will be repeated using only cases and controls that were given a diagnosis code for ARI on the same day as specimen collection, as

per the SESE database or MACAR. Online supplementary appendix 3 lists the ICD-9 and ICD-10 codes used to define ARIs.

PATIENT AND PUBLIC INVOLVEMENT

Patients and the public were not involved in the design of the study, including the development of the research question, outcomes measures, recruitment to or conduct of the study. The results of the study will be disseminated to the public as deemed appropriate by public health officials.

DISCUSSION

This protocol describes the estimation of seasonal influenza VE using specimens collected for routine influenza diagnostics as well as administrative data and vaccination records.

A key strength of this approach is the large sample size. This approach allows calculation of timely, precise influenza VE estimates weeks prior to the influenza season peak, creating an early warning system for public health if, as in the 2014–2015 season, the vaccine is found to have exceedingly low effectiveness. Early notification of VE can assist public health in determining policies, messaging and allocation of resources (antiviral agents, staffing emergency departments) to counter a potentially more severe influenza season.[37][38] The large sample size also allows for stratified analyses of VE based on product, age group or region.

Whereas sentinel physician networks rely primarily on self-reported measures of influenza vaccination,[34] a significant strength of this study is the use of the near-real-time influenza vaccination registry that contains individual-level, linkable data for most influenza vaccinations administered in the province. Use of this registry reduces the likelihood of recall error and information biases such as social desirability bias and reduces non-differential misclassification, which would bias the OR towards the null, thus underestimating VE.[12]

Finally, we are certain to capture the results of all respiratory virus testing in the province, as all respiratory virus testing is centralised at ProvLab and there is limited use of point-of-care testing.

There are some limitations to this methodology compared with the traditional method of VE estimation using sentinel physician networks, because a standardised clinical case definition cannot be applied to determine study eligibility. A sensitivity analysis restricting to healthcare encounters with a diagnosis code for ARI will be used as a proxy for a standard case definition.

While the inclusion of confounders is important for VE estimate adjustment, not all known confounders can be measured using administrative data. Frailty has been demonstrated to be a potential confounder of VE.[39–41] Frailty cannot be included in the multivariable model because no validated indices of frailty generated from standard administrative data exist at this time. However,

this may not affect the results significantly as a previous study indicated that inclusion of frailty in the multivariate model increased VE estimates only slightly.[42]

Laboratory requisitions in Alberta do not contain illness onset date. Ideally this would be used to ensure that the negative laboratory test results were representative of an acute infectious period and that test-negative specimens were not collected after viral shedding had ceased. Sullivan *et al* 2016 have indicated this bias may be accounted for by selecting influenza test-negative controls that were positive for another respiratory virus. Requiring controls to be positive for another virus excludes individuals who are tested long after their acute infectious period. However, a recent systematic review found no differences when using different groups of controls.[43]

Comparison of the VE results using administrative data to previously published studies, specifically sentinel surveillance for the same seasons (2011/12–2019/20), will help to identify further areas of refinement.

This approach could successfully allow for the generation of early influenza VE estimates which could facilitate tailoring of public health messaging and assist in public health operations planning for the peak of the influenza season.

Author affiliations
[1]Ministry of Health, Government of Alberta, Edmonton, Alberta, Canada
[2]Department of Public Health, Concordia University of Edmonton, Edmonton, Alberta, Canada
[3]Populations and Public Health Research Program, ICES, Toronto, Ontario, Canada
[4]Public Health Sciences, Public Health Ontario, Toronto, Ontario, Canada
[5]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
[6]Department of Family and Community Medicine, University of Toronto, Toronto, Ontario, Canada
[7]Canadian Blood Services, Edmonton, Alberta, Canada
[8]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[9]Division of Preventive Medicine, University of Alberta, Edmonton, Alberta, Canada
[10]School of Public Health, University of Alberta, Edmonton, Alberta, Canada

ORCID iD
Lawrence W Svenson http://orcid.org/0000-0002-3391-578X

REFERENCES

1. Mertz D, Kim TH, Johnstone J, *et al*. Populations at risk for severe or complicated influenza illness: systematic review and meta-analysis. *BMJ* 2013;347.
2. Alberta Health Services2017
3. Wong S-S, Webby RJ. Traditional and new influenza vaccines. *Clin Microbiol Rev* 2013;26:476–92.
4. Skowronski DM, Chambers C, Sabaiduc S, *et al*. A perfect storm: impact of genomic variation and serial vaccination on low influenza vaccine effectiveness during the 2014-2015 season. *Clin Infect Dis* 2016;63:21–32.
5. World Health Organization. Influenza update N°309. Geneva World Health Organization; 2018: 1–8.
6. Skowronski DM, De Serres G, Crowcroft NS, *et al*. Association between the 2008-09 seasonal influenza vaccine and pandemic H1N1 illness during spring-summer 2009: four observational studies from Canada. *PLoS Med* 2010;7:e1000258.
7. Chambers C, Skowronski DM, Sabaiduc S, *et al*. Interim estimates of 2015/16 vaccine effectiveness against influenza A(H1N1)pdm09, Canada, February 2016. *Euro Surveill* 2016;21:30168.
8. 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.
9. Orenstein WA, Bernier RH, Dondero TJ, *et al*. Field evaluation of vaccine efficacy. *Bull World Health Organ* 1985;63:1055–68.
10. Public Health Agency of Canada. Effectiveness of vaccine against medical consultation due to laboratory-confirmed influenza: results from a sentinel physician pilot project in British Columbia, 2004-2005. *Can Commun Dis Rep* 2005;31:181–91.
11. Skowronski DM, Chambers C, Sabaiduc S, *et al*. Interim estimates of 2016/17 vaccine effectiveness against influenza A(H3N2), Canada, January 2017. *Eurosurveillance* 2017;22.
12. World Health Organization. Evaluation of influenza vaccine effectiveness: a guide to the design and interpretation of observational studies. Geneva World Health Organization; 2017: 1–47.
13. Belongia EA, Simpson MD, King JP, *et al*. Variable influenza vaccine effectiveness by subtype: a systematic review and meta-analysis of test-negative design studies. *Lancet Infect Dis* 2016;16:942–51.
14. Kwong JC, Buchan SA, Chung H, *et al*. Can routinely collected laboratory and health administrative data be used to assess influenza vaccine effectiveness? assessing the validity of the flu and other respiratory viruses research (forever) cohort. *Vaccine* 2019;37:4392–400.
15. Fathima S, Simmonds K, Invik J, *et al*. Use of laboratory and administrative data to understand the potential impact of human parainfluenza virus 4 on cases of bronchiolitis, croup, and pneumonia in Alberta, Canada. *BMC Infect Dis* 2016;16:402.
16. Alberta Health. History of immunization in Alberta. Edmonton Government of Alberta; 2017.
17. Pabbaraju K, Tokaryk KL, Wong S, *et al*. Comparison of the Luminex xTAG respiratory viral panel with in-house nucleic acid amplification tests for diagnosis of respiratory virus infections. *J Clin Microbiol* 2008;46:3056–62.
18. Pabbaraju K, Wong S, Wong AA, *et al*. Design and validation of real-time reverse transcription-PCR assays for detection of pandemic (H1N1) 2009 virus. *J Clin Microbiol* 2009;47:3454–60.
19. Chaudhry A, Bastien N, Li Y, *et al*. Oseltamivir resistance in an influenza A (H3N2) virus isolated from an immunocompromised patient during the 2014-2015 influenza season in Alberta, Canada. *Influenza Other Respi Viruses* 2016;10:532–5.
20. World Health Organization. Global epidemiological surveillance standards for influenza. Geneva World Health Organization; 2013.
21. World Health Organization. A manual for estimating disease burden associated with seasonal influenza. Geneva World Health Organization; 2015.
22. Alberta Health. Alberta health seasonal influenza in Alberta 2016/2017 summary report. Edmonton Government of Alberta; 2017.
23. Alberta Health. Overview of administrative health datasets. Edmonton Government of Alberta ; 2017.
24. Russell ML, Schopflocher DP, Svenson L, *et al*. Secular trends in the epidemiology of shingles in Alberta. *Epidemiol Infect* 2007;135:908–13.
25. Government of Alberta. *Immunization data submission and response guidelines*, 2017.
26. MacDonald SE, Dover DC, Simmonds KA, *et al*. Risk of febrile seizures after first dose of measles-mumps-rubella-varicella vaccine: a population-based cohort study. *Can Med Assoc J* 2014;186:824–9.
27. Alberta Health. Alberta health claims assessment, 2017. Available: https://open.alberta.ca/publications/alberta-health-diagnostic-codes [Accessed 23 Feb 2018].
28. Lix LM, Walker R, Quan H, *et al*. Features of physician services databases in Canada. *Chronic Dis Inj Can* 2012;32:186–93.
29. Hinds A, Lix LM, Smith M, *et al*. Quality of administrative health databases in Canada: a scoping review. *Can J Public Health* 2016;107:e56–61.
30. Shiff NJ, Jama S, Boden C, *et al*. Validation of administrative health data for the pediatric population: a scoping review. *BMC Health Serv Res* 2014;14:236.
31. Quan H, Smith M, Bartlett-Esquilant G, *et al*. Mining administrative health databases to advance medical science: geographical considerations and untapped potential in Canada. *Can J Cardiol* 2012;28:152–4.
32. Schwartz KL, Jembere N, Campitelli MA, *et al*. Using physician billing claims from the Ontario health insurance plan to determine individual influenza vaccination status: an updated validation study. *CMAJ Open* 2016;4:E463–E470.
33. Sullivan SG, Cowling BJ. "Crude Vaccine Effectiveness" Is a Misleading Term in Test-negative Studies of Influenza Vaccine Effectiveness. *Epidemiology* 2015;26:e60.
34. Skowronski DM, Janjua NZ, De Serres G, *et al*. Low 2012-13 influenza vaccine effectiveness associated with mutation in the egg-adapted H3N2 vaccine strain not antigenic drift in circulating viruses. *PLoS One* 2014;9:e92153.
35. Skowronski D, Chambers C, Sabaiduc S, *et al*. Interim estimates of 2013/14 vaccine effectiveness against influenza A(H1N1)pdm09 from Canada s sentinel surveillance network, January 2014. *Euro Surveill* 2014;19. doi:10.2807/1560-7917.ES2014.19.5.20690. [Epub ahead of print: 06 Feb 2014].
36. Sullivan SG, Tchetgen Tchetgen EJ, Cowling BJ. Theoretical basis of the test-negative study design for assessment of influenza vaccine effectiveness. *Am J Epidemiol* 2016;184:345–53.
37. Orton L, Lloyd-Williams F, Taylor-Robinson D, *et al*. The use of research evidence in public health decision making processes: systematic review. *PLoS One* 2011;6:e21704.
38. Savel TG, Foldy S, Centers for Disease Control and Prevention. The role of public health informatics in enhancing public health surveillance. *MMWR Suppl* 2012;61:20–4.
39. Nelson JC, Jackson ML, Weiss NS, *et al*. New strategies are needed to improve the accuracy of influenza vaccine effectiveness estimates among seniors. *J Clin Epidemiol* 2009;62:687–94.
40. Jackson LA, Jackson ML, Nelson JC, *et al*. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol* 2006;35:337–44.
41. Jackson LA, Nelson JC, Benson P, *et al*. Functional status is a confounder of the association of influenza vaccine and risk of all cause mortality in seniors. *Int J Epidemiol* 2006;35:345–52.
42. Talbot HK, Nian H, Chen Q, *et al*. Evaluating the case-positive, control test-negative study design for influenza vaccine effectiveness for the frailty bias. *Vaccine* 2016;34:1806–9.
43. Feng S, Cowling BJ, Kelly H, *et al*. Estimating influenza vaccine effectiveness with the test-negative design using alternative control groups: a systematic review and meta-analysis. *Am J Epidemiol* 2018;187:389–97.