Vaccine 37 (2019) 4392–4400

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

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

Jeffrey C. Kwong [a,b,c,d,e,*], Sarah A. Buchan [a,b,c], Hannah Chung [a], Michael A. Campitelli [a], Kevin L. Schwartz [a,b,c], Natasha S. Crowcroft [a,b,c,f], Michael L. Jackson [g], Timothy Karnauchow [h,i], Kevin Katz [j], Allison J. McGeer [c,f,k], J. Dayre McNally [h], David C. Richardson [l], Susan E. Richardson [f,m], Laura C. Rosella [a,b,c], Andrew Simor [f,n], Marek Smieja [o], George Zahariadis [p,q], Aaron Campigotto [m,p], Jonathan B. Gubbay [b,f,m]

[a] ICES, Toronto, ON, Canada
[b] Public Health Ontario, Toronto, ON, Canada
[c] Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[d] Department of Family & Community Medicine, University of Toronto, Toronto, ON, Canada
[e] University Health Network, Toronto, ON, Canada
[f] Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada
[g] Kaiser Permanente Washington Health Research Institute, Seattle, WA, United States
[h] Children's Hospital of Eastern Ontario, Ottawa, ON, Canada
[i] Department of Pathology and Laboratory Medicine, University of Ottawa, Ottawa, ON, Canada
[j] North York General Hospital, Toronto, ON, Canada
[k] Sinai Health System, Toronto, ON, Canada
[l] William Osler Health System, Brampton, ON, Canada
[m] Hospital for Sick Children, Toronto, ON, Canada
[n] Sunnybrook Health Sciences Centre, Toronto, ON, Canada
[o] McMaster University, Hamilton, ON, Canada
[p] London Health Sciences Centre, London, ON, Canada
[q] Newfoundland & Labrador Public Health Laboratory, St. John's, NF&L, Canada

ARTICLE INFO

Article history:
Received 30 October 2018
Received in revised form 16 May 2019
Accepted 7 June 2019
Available online 17 June 2019

Keywords:
Influenza
Human
Epidemiology
Influenza Vaccines
Data Linkage
Bias

ABSTRACT

Background: Linking data on laboratory specimens collected during clinical practice with health administrative data permits highly powered vaccine effectiveness (VE) studies to be conducted at relatively low cost, but bias from using convenience samples is a concern. We evaluated the validity of using such data for estimating VE.

Methods: We created the Flu and Other Respiratory Viruses Research (FOREVER) Cohort by linking individual-level data on respiratory virus laboratory tests, hospitalizations, emergency department visits, and physician services. For community-dwelling adults aged > 65 years, we assessed the presence and magnitude of information and selection biases, generated VE estimates under various conditions, and compared our VE estimates with those from other studies.

Results: We included 65,648 unique testing episodes obtained from 54,434 individuals during the 2010–11 to 2015–16 influenza seasons. To examine information bias, we found the proportion testing positive for influenza for patients with unknown interval from illness onset to specimen collection was more similar to patients for whom illness onset date was ≤ 7 days before specimen collection than to patients for whom illness onset was > 7 days before specimen collection. To assess the presence of selection bias, we found the likelihood of influenza testing was comparable between vaccinated and unvaccinated individuals, although the adjusted odds ratios were significantly greater than 1 for some healthcare settings and during some influenza seasons. Over 6 seasons, VE estimates ranged between 36% (95%CI, 27–44%) in 2010–11 and 5% (95%CI, −2, 11%) in 2014–15. VE estimates were similar under a range of conditions, but were consistently higher when accounting for misclassification of vaccination status through a

* Corresponding author at: ICES, G1 06, 2075 Bayview Avenue, Toronto, ON M4N 3M5, Canada.
E-mail address: jeff.kwong@utoronto.ca (J.C. Kwong).

https://doi.org/10.1016/j.vaccine.2019.06.011
0264-410X/

quantitative sensitivity analysis. VE estimates from the FOREVER Cohort were comparable to those from other studies.
Conclusions: Routinely collected laboratory and health administrative data contained in the FOREVER Cohort can be used to estimate influenza VE in community-dwelling older adults.

1. Background

Influenza vaccines are the cornerstone for preventing influenza infections and associated complications, but due to frequent changes in circulating viruses and vaccine components, ongoing monitoring of influenza vaccine effectiveness (VE) is needed [1]. Over the past decade, several networks around the world have created platforms to evaluate VE by prospectively enrolling patients who fulfill pre-specified case definitions for acute respiratory illness (ARI) and testing them for influenza using nucleic acid amplification-based methods [2–5]. These studies employ the test-negative design, which is conceptually similar to a nested case-control study by comparing test-positive "cases" with test-negative "controls" derived from a cohort of patients who are laboratory-tested for influenza [6–9]. While most of these networks are limited to outpatients, a growing number are recruiting inpatients [10–13]. However, many of these studies suffer from restricted statistical power due to the limited number of patients they are able to recruit, leading to wide confidence intervals for VE estimates. This is exacerbated for high-risk groups, who comprise relatively small numbers in these studies but may be those for whom there is the greatest interest in determining VE due to their disproportionate contribution to the overall burden of influenza.

Our team previously investigated the feasibility of applying the test-negative design to routinely collected laboratory and health administrative data to evaluate VE. In a pilot study, we linked laboratory data from Public Health Ontario (PHO) to health administrative data to measure VE against laboratory-confirmed influenza hospitalizations among community-dwelling adults aged > 65 years during the 2010–11 influenza season [14]. This was one of the first studies to assess VE against a serious yet specific outcome (i.e., laboratory-confirmed influenza hospitalization) among older adults, and with a sample size of 2230 subjects, including 569 test-positive cases, remains one of the largest single-season studies to date for this age group and outcome.

Given the increasing capacity to link individual-level records across multiple databases in the "Big Data" era, combining routinely collected laboratory and health administrative data holds tremendous promise by permitting highly powered studies to be conducted at relatively low cost. However, concerns have been raised that VE analyses based on a convenience sample of clinical diagnostic tests could be biased [15]. Therefore, we sought to confirm the appropriateness and validity of using routinely collected laboratory and health administrative data for estimating VE.

2. Methods

2.1. Study population and setting

The Canadian province of Ontario (population 14 million in 2016) has a universal single-payer health system. We created the Flu and Other Respiratory Viruses Research (FOREVER) Cohort by linking respiratory virus test results obtained from 11 PHO laboratories and nine academic hospital laboratories to an extensive array of population-based health databases housed at ICES using a combination of deterministic and probabilistic data linkage (overall linkage proportion = 97.8%). Although the FOREVER Cohort includes results of all respiratory virus testing conducted by the participating laboratories between 1 May 2009 and 30 June 2016, we used data from the 2010–11 to 2015–16 influenza seasons (defined in Table S1) for this study to avoid the 2009 A/H1N1 pandemic because it was epidemiologically dissimilar from seasonal influenza epidemics, and because of heightened data quality concerns regarding the measurement of influenza vaccination during that season. Specifically, a large proportion of pandemic vaccine doses were administered through public health (and are therefore unavailable in Ontario health administrative data), and of the vaccines administered in physician offices, it is not possible to distinguish between pandemic and seasonal vaccines in the physician billing claims data due to the use of a common vaccine billing code. We restricted the cohort to community-dwelling adults aged > 65 - years because of the substantial disease burden from influenza and the greater validity (i.e., higher sensitivity and positive predictive value) of the influenza vaccination data in this population [16]. Moreover, since medications are publicly funded for older adults starting at age 65 years in Ontario, we estimated VE for those aged > 65 years to allow for a 1-year look-back period to identify prescriptions that are included in the definition of certain comorbidities.

At the specimen level (n = 153,531), after excluding duplicates (n = 11,404), non-insured individuals (n = 333), records with data quality issues (n = 154), individuals with unknown postal code (n = 52), and non-Ontario residents (n = 51), we had 141,537 unique respiratory specimens from individuals aged > 65 years. Combining the results of all specimens collected for the same individual on the same day into unique testing episodes (n = 137,426), and after excluding those from long-term care residents (n = 37,846), testing episodes outside of influenza season periods (n = 28,410), and specimens collected from same-day surgeries or unknown settings (n = 5522), the final analytic cohort included 65,648 testing episodes.

3. Data sources and definitions

3.1. Laboratory data

We included individuals with respiratory specimens submitted to PHO laboratories for testing for influenza as part of routine clinical care (87%), outbreak investigations (3%), or research (10%), and individuals with respiratory specimens submitted to the hospital microbiology laboratories (each serving one or more hospitals) for testing for influenza as part of clinical care (some were also forwarded to PHO for confirmation or additional testing). Of these, 9% of patients were tested for influenza only, and 91% were tested for influenza and at least one other respiratory virus. Influenza A subtype information was available for 49% of influenza A-positive specimens. Testing modalities employed included monoplex and multiplex polymerase chain reaction (PCR), direct fluorescent antibody staining, viral culture, and enzyme immunoassay rapid antigen tests (Tables S2 and S3). Specimens are frequently tested by more than one method. Overall, 79% of individuals had specimens that were tested by molecular (PCR) assays, with the proportion of individuals with specimens analyzed using this technology

remaining steady over time, ranging between 77% and 81% during 2010–11 to 2015–16. All network participants are fully accredited clinical laboratories.

3.2. Health administrative data

We used provincial health administrative data to measure exposures (i.e., influenza vaccination), outcomes (e.g., hospitalization), and covariates (e.g., medical conditions). To ascertain influenza vaccination status, we used physician and (starting in 2012, when a policy change permitted pharmacists to administer influenza vaccines) [17] pharmacist billing claims, contained in the Ontario Health Insurance Plan (OHIP) database and the Ontario Drug Benefits database, respectively. We defined those who received influenza vaccination $\geq$ 14 days prior to specimen collection as vaccinated, and excluded patients who received a vaccine < 14 days prior to specimen collection. We used data from the Canadian Institute for Health Information's Discharge Abstract Database, National Ambulatory Care Reporting System database, and Same-Day Surgery database, and the OHIP database to identify hospitalizations (including intensive care admissions), emergency department (ED) visits, same-day surgeries, and physician office visits, respectively. We applied validated algorithms to these databases to identify medical conditions that increase the risk of influenza-related complications (defined in Table S4) [18–33]. We determined age, sex, and location of residence (for rural/urban status and neighbourhood income) from the Registered Persons Database, which includes all individuals eligible for health insurance (essentially the entire Ontario population).

We identified all healthcare encounters associated with a specimen on the date of specimen collection. When individuals had records indicating multiple interactions on the same day (e.g., a physician office visit, an ED visit, and a hospitalization), we designated the setting of specimen collection to the most clinically intensive setting by applying the following hierarchy: intensive care unit > hospital ward > ED > physician office.

We identified encounters for ARIs using diagnostic codes (defined in Table S5) recorded in the above databases as a proxy for the pre-specified case definitions used in prospective test-negative studies.

3.3. Statistical analysis

3.3.1. Testing for bias when using routinely collected specimens

Concerns about the validity of VE estimates based on routinely collected clinical specimens relate to the potential for bias that could arise from using these specimens. Broadly, the use of clinical specimens could result in *information bias* if key data are missing for the specimens, and *selection bias* if the vaccine/outcome association in patients chosen for testing is different from the association in the broader population of ill patients.

3.3.2. Information bias

The process of laboratory testing for influenza does not have perfect sensitivity and specificity. One concern is the potential for false negative results when the interval between illness onset and specimen collection is too long. Patients with a long (>7 day) interval between onset and specimen collection may no longer be shedding detectable virus when their specimen is collected, leading to negative laboratory test results even when influenza was the cause of the illness [34,35]. Including false negative specimens will bias VE estimates toward the null. Test-negative studies typically restrict eligibility to patients enrolled within 7 days of illness onset to mitigate against this bias [2]. Applying this criterion to the FOREVER Cohort is challenging because illness onset date was documented for only 10% of testing episodes in the FOREVER Cohort used for this analysis (but was within 7 days of specimen collection for 89% of those with illness onset date documented).

To test for this bias, we indirectly assessed the extent of outcome misclassification for patients with unknown interval between illness onset and specimen collection by comparing the proportion testing positive for influenza in that group versus those with known illness onset date. Details are provided in the Supplementary Text.

3.3.3. Selection bias

Selection bias can arise when using clinical specimens if the probability of laboratory testing differs between vaccinated and unvaccinated patients [36]. For example, clinicians might discount influenza as a possible cause of an ARI for patients who are known to be vaccinated, leading to under-representation of vaccinated patients in the tested population. If unvaccinated patients are more likely to be tested, VE estimates will be biased away from the null.

To test for the presence of this form of selection bias in the FOREVER Cohort, we compared the proportion tested for influenza among vaccinated and unvaccinated individuals and determined the association between influenza vaccination and subsequent influenza testing. Details are provided in the Supplementary Text.

3.3.4. Estimating influenza VE under various conditions

Most prospective test-negative studies have the following criteria: (1) enrolment of either outpatients or inpatients; (2) use of PCR methods to test for influenza; and (3) restriction to specimens collected from individuals with ARI symptoms within 7 days of symptom onset. Since a frequent limitation of prospective test-negative studies is inadequate sample size, leading to imprecise VE estimates particularly for stratified analyses (e.g., by age group, by influenza subtype), our goal was to identify the criteria that could be relaxed in order to maximize sample size while maintaining validity. Furthermore, it may not always be possible to apply these criteria to routinely collected clinical specimens (i.e., in other databases). Thus, we explored the impact of modifying the criteria outlined above by generating VE estimates under various conditions. We used multivariable logistic regression to produce adjusted VE estimates, where $VE = (1 - \text{adjusted OR}) \times 100\%$, controlling for age, sex, presence of a medical condition, and calendar time.

For each influenza season, we estimated VE for individuals with specimens collected in the following settings: (1) inpatients and outpatients (combining specimens obtained from EDs and physician offices due to small numbers); (2) inpatients only; and (3) outpatients only. We also generated estimates using specimens tested by any laboratory method (to maximize sample size) and only using specimens tested by PCR. Including specimens tested by non-PCR methods is justifiable because the use of laboratory tests with reduced sensitivity but high specificity introduces minimal bias when estimating VE [2,37], and all other methods used by the participating laboratories have high specificity.

Next, we applied three restrictions to emulate the typical inclusion criteria of test-negative studies and determine their impact. First, we restricted the analysis to ARI-coded encounters. Second, we restricted the analysis to non-outbreak-related specimens. Third, we restricted the analysis to patients who were documented to have had a specimen collected $\leq$ 7 days of illness onset. We applied these restrictions both separately and cumulatively. Further details are provided in the Supplementary Text.

Since some Ontarians receive influenza vaccines through public health or workplaces, their immunization events are not captured in provincial health administrative databases (unless patients report them to their physician) and consequently they are misclassified as unvaccinated in these analyses. We previously validated physician billing claims for influenza vaccines against self-reported influenza vaccination status and found they have a sensi-

tivity of 69% and a specificity of 90% [16]. We incorporated these parameters in a quantitative sensitivity analysis that corrects for non-differential misclassification of a binary variable (i.e., seasonal influenza vaccination status) through probabilistic Monte Carlo simulation methods [38]. We estimated VE using this quantitative sensitivity analysis to assess the impact of correcting for misclassification of vaccination status. Further details are included in the Supplementary Text.

3.4. Comparing FOREVER Cohort VE estimates with other test-negative studies

Using the same method to calculate VE as described above, we generated influenza type/subtype-specific VE estimates for each influenza season and healthcare setting, correcting for misclassification of vaccination status. We compared these estimates with published estimates identified from an environmental scan of the literature. Details are provided in the Supplementary Text.

An overview of the analyses conducted can be found in Table 1.

4. Results

The study cohort included 65,648 unique testing episodes with ≥ 1 respiratory specimen collected, obtained from 54,434 individuals, with 6.8% of individuals in the cohort tested in more than one season. Compared to the Ontario population of adults aged > 65 years, those in the FOREVER Cohort were older, were more likely to live in lower-income neighbourhoods and urban areas, had more comorbidities, and received more healthcare services and prescription medications (Table 2).

Table 1
Summary of statistical analyses conducted.

Analysis	Purpose & Rationale	Methods	Restricted analyses
Testing for information bias	To investigate the impact of missing illness onset date on the misclassification of case status. Prospective TND VE studies collect specimens $\leq$ 7 days of illness onset. However, illness onset information is often missing with routinely collected specimens. Including individuals with a longer interval between illness onset and specimen collection could bias VE estimates toward the null because of false-negative controls.	Individuals with specimens tested using PCR by Public Health Ontario, the only laboratory in the FOREVER Cohort that documents illness onset date (for a subset of specimens). Compared characteristics and the adjusted proportions testing positive for influenza between patients with: • Unknown interval between illness onset and specimen collection • Illness onset $\leq$ 7 days before specimen collection • Illness onset > 7 days before specimen collection	1. Excluded those with specimens collected for research purposes, which deliberately collects specimens $\leq$ 7 days of onset and are not routinely collected specimens. 2. Restricted to those with specimens collected during an ARI-coded healthcare encounter to observe the robustness of the comparisons when including only those presenting with symptoms that would justify the collection of a respiratory specimen. 3. Applied both of the above restrictions.
Testing for selection bias	To investigate the presence of preferential testing of unvaccinated patients for routinely collected specimens. Prospective TND VE studies collect specimens from patients meeting eligibility criteria. In contrast, with clinical specimens, physician's knowledge of their patient's influenza vaccination status might influence collection and testing. If those unvaccinated are more likely to be tested, this will bias VE away from null.	Identified healthcare encounters with a respiratory specimen from the FOREVER Cohort collected from all encounters during the influenza seasons from the source population (community-dwelling adults aged > 65 years). Used logistic regression to determine odds of testing between unvaccinated and vaccinated individuals, stratified by influenza season.	1. Restricted healthcare encounters by setting of specimen collection and presence of ARI code: a. ARI-coded hospitalizations b. ARI-coded ED visits c. ARI-coded physician office visits d. Non-ARI-coded hospitalizations e. Non-ARI-coded ED visits f. Non-ARI-coded physician office visits 2. Adjusted for misclassification of vaccination status.
Estimating influenza VE under various conditions	To estimate influenza VE using routinely collected specimens in the FOREVER cohort under conditions that are comparable to prospective TND VE studies. Prospective TND VE studies use specimens collected from either inpatients or outpatients, who meet specific eligibility criteria (ARI case definition, symptom onset $\leq$ 7 days), and are tested by PCR.	Used logistic regression to calculate VE, where VE=(1–adjusted OR)x100%, controlling for age, sex, presence of a medical condition, and calendar time, stratified by influenza season.	1. Restricted by setting of specimen collection: a. Inpatients (hospitalizations) b. Outpatients (ED and physician office visits) and tested by any method of test, or PCR only. 2. Among inpatients, applied the following restrictions independently and cumulatively: a. ARI-coded encounters b. Excluding specimens collected for outbreak investigations c. Specimens collected $\leq$ 7 days of illness onset and applied each restriction's counterfactual condition independently. 3. Adjusted for misclassification of vaccination status.
Comparing FOREVER Cohort VE estimates with other test-negative studies	To calculate influenza type/subtype-specific VE estimates for each influenza season and healthcare setting using routinely collected specimens in the FOREVER cohort and compare with VE estimates from prospective TND VE studies.	Used logistic regression to calculate VE, where VE = (1 − adjusted OR) × 100%, controlling for age, sex, presence of a medical condition, and calendar time, stratified by influenza season and healthcare setting. Test-positive cases are restricted to those who have the type- (e.g., Influenza A or B) or subtype-specific (e.g., H1N1, H3N2) influenza outcome.	1. Restricted by setting of specimen collection: a. Inpatients (hospitalizations) b. Outpatients (ED and physician office visits) and adjusted for misclassification of vaccination status.

TND = test-negative design; VE = vaccine effectiveness; PCR = polymerase chain reaction; ARI = acute respiratory illness; ED = emergency department; OR = odds ratio.

Table 2
Selected characteristics of community-dwelling adults aged > 65 years in the FOREVER cohort.

Characteristic	N (%)						
	Respiratory virus season						Ontario general population, 2016 (N = 2,120,261)
	2010–11 (N = 6243)	2011–12 (N = 3630)	2012–13 (N = 10,945)	2013–14 (N = 11,270)	2014–15 (N = 16,591)	2015–16 (N = 9,984)	
Age group							
66–75 years	2146 (34)	1296 (36)	3725 (34)	4135 (37)	5344 (32)	3820 (38)	1,244,662 (59)
76–85 years	2558 (41)	1458 (40)	4300 (39)	4368 (39)	6449 (39)	3725 (37)	643,851 (30)
≥86 years	1539 (25)	876 (24)	2920 (27)	2767 (25)	4798 (29)	2439 (24)	231,748 (11)
Male sex	3006 (48)	1751 (48)	5238 (48)	5414 (48)	7671 (46)	4890 (49)	973,407 (46)
Neighbourhood income quintile							
1 (lowest)	1442 (23)	801 (22)	2436 (22)	2515 (22)	3671 (22)	2297 (23)	374,489 (18)
2	1384 (22)	823 (23)	2347 (21)	2366 (21)	3483 (21)	2062 (21)	418,877 (20)
3	1103 (18)	713 (20)	2035 (19)	2188 (19)	3173 (19)	1839 (18)	416,134 (20)
4	1076 (17)	605 (17)	1936 (18)	2010 (18)	3002 (18)	1838 (18)	447,342 (21)
5 (highest)	1192 (19)	671 (18)	2110 (19)	2136 (19)	3158 (19)	1893 (19)	455,955 (22)
Rural residence	476 (8)	239 (7)	957 (9)	874 (8)	1597 (10)	929 (9)	285,924 (13)
Medical conditions[1]							
Cardiovascular disease[2]	3964 (63)	2350 (65)	6998 (64)	7165 (64)	10,577 (64)	6450 (65)	580,913 (27)
Chronic obstructive pulmonary disease	3182 (51)	1811 (50)	5645 (52)	5750 (51)	8207 (49)	5301 (53)	425,191 (20)
Diabetes	2578 (41)	1618 (45)	4607 (42)	4764 (42)	7143 (43)	4394 (44)	641,081 (30)
Cancer	1795 (29)	1098 (30)	3135 (29)	3313 (29)	4808 (29)	3075 (31)	350,152 (17)
Asthma	1725 (28)	1076 (30)	3017 (28)	3192 (28)	4501 (27)	2792 (28)	269,778 (13)
Anemia	1501 (24)	911 (25)	2598 (24)	2744 (24)	3905 (24)	2432 (24)	189,283 (9)
Chronic kidney disease	1259 (20)	824 (23)	2345 (21)	2501 (22)	3692 (22)	2340 (23)	146,036 (7)
Dementia/Frailty	1173 (19)	622 (17)	2098 (19)	2166 (19)	3608 (22)	1837 (18)	108,581 (5)
Immunocompromise	832 (13)	465 (13)	1519 (14)	1,571 (14)	2269 (14)	1583 (16)	95,145 (4)
Any of the above medical conditions	5,968 (96)	3,465 (95)	10,425 (95)	10,723 (95)	15,766 (95)	9560 (96)	1,440,310 (68)
Hospitalizations, past 3 yrs (mean ± SD)	1.49 +/− 2.12	1.59 +/− 2.26	1.56 +/− 2.19	1.62 +/− 2.25	1.54 +/− 2.14	1.68 +/− 2.3	0.36 +/− 0.89
Outpatient visits in past yr (mean ± SD)	14.2 +/− 10.6	15.1 +/− 11.2	13.7 +/− 10.6	14.3 +/− 11.0	13.9 +/− 10.8	14.7 +/− 11.5	8.11 +/− 7.76
Prescriptions in past yr (mean ± SD)	16.7 +/− 9.66	16.7 +/− 9.27	16.2 +/− 9.13	16.7 +/− 9.37	16.5 +/− 9.25	16.5 +/− 9.25	8.07 +/− 6.83

[1] See Table S4 for definitions.
[2] Includes congestive heart failure, ischemic heart disease, arrhythmias, acute ischemic stroke, and transient ischemic attack.

4.1. Testing for bias when using routinely collected specimens

4.1.1. Information bias

Based on standardized differences, patients with an unknown interval between illness onset and specimen collection tended to be more similar to those with an interval of ≤ 7 days than those with an interval of > 7 days in both the primary analysis and sensitivity analyses (Table S6).

As anticipated, the adjusted proportion testing positive for influenza appeared higher when illness onset was ≤ 7 days before specimen collection than when the interval was > 7 days, for all healthcare settings (Table S7). The adjusted proportion positive for patients with unknown interval from illness onset to specimen collection (32%; 95%CI, 32–33%) was more similar to patients for whom illness onset was documented to be ≤ 7 days before specimen collection (29%; 95%CI, 28–30%) (difference = −3%; 95%CI, −4%, −1%) than to patients for whom illness onset was > 7 days before specimen collection (20%; 95%CI, 16–24%) (difference = −12%; 95%CI, −14%, −6%). With some exceptions, findings were comparable for all healthcare settings and sensitivity analyses.

4.1.2. Selection bias

Among institutions and physicians that submitted respiratory specimens for testing, the proportions of individuals tested for influenza were higher for hospitalizations than ED visits and physician office visits, and were higher for ARI-coded healthcare encounters compared to non-ARI-coded encounters (Table S8). For most influenza seasons, the proportions tested were comparable between vaccinated and unvaccinated individuals irrespective of the healthcare setting and the presence of ARI diagnostic codes. However, some of the aORs were significantly higher than 1.00, indicating that during certain seasons and in certain healthcare settings, unvaccinated individuals were more likely to be tested for influenza than vaccinated individuals. The magnitudes of the aORs appeared to be highest for patients tested in physician offices (range = 0.79–1.60), followed by EDs (range = 0.90–1.49), then hospitals (range = 0.90–1.14). When we corrected for misclassification of vaccination status, the aORs tended to move further from the null, particularly for non-ARI-coded healthcare encounters compared to ARI-coded healthcare encounters, but this occurred for a minority of the estimates.

4.2. Estimating influenza VE under various conditions

Over the 6 influenza seasons, adjusted VE estimates against any influenza for inpatients and outpatients combined ranged from a high of 36% (95%CI, 27–44%) in 2010–11 to a low of 5% (95%CI, –2, 11%) in 2014–15 (Fig. 1, Table S9). Comparing inpatients and outpatients, VE estimates were similar and/or had highly overlapping confidence intervals, with the latter due to less precise estimates for outpatients. VE estimates were also similar regardless of the testing method(s) used. The numbers of patients included in these analyses are presented in Tables S10A and S10B.

Among inpatients (we excluded outpatients due to the small numbers and potential concerns about selection bias), VE estimates were similar regardless of the restrictions, but they became less precise as more restrictions were applied (Fig. 2, Table S9). VE estimates based on patients with specimens collected during non-ARI-coded hospitalizations and those without illness onset date documented were also similar. Due to low numbers of patients with outbreak-related specimens and specimens collected > 7 days after illness onset, the VE estimates were very imprecise and were not reported. Sensitivity analyses for the VE estimates accounting for misclassification of vaccination status were consistently higher than the original estimates. For example, in 2015–16, the estimate

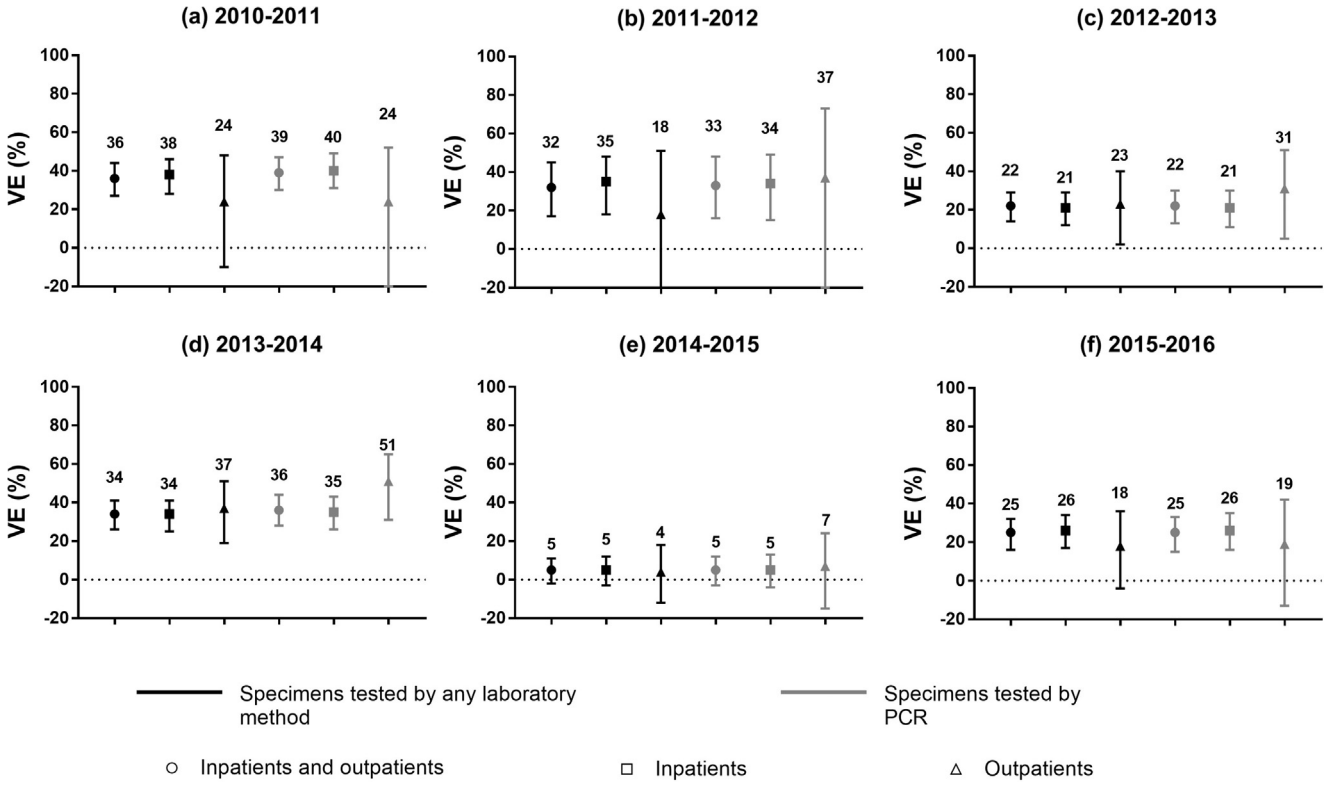

Fig. 1. Forest plot of influenza vaccine effectiveness estimates for adults aged > 65 years using specimens tested by any laboratory method and only specimens tested by polymerase chain reaction, comparing inpatients and outpatients combined, inpatients only, and outpatients only, for the 2010–11 (panel a), 2011–12 (panel b), 2012–13 (panel c), 2013–14 (panel d), 2014–15 (panel e), and 2015–16 (panel f) influenza seasons. VE = vaccine effectiveness; PCR = polymerase chain reaction.

of VE increased from 26% (95%CI, 16–32%) to 50% (95%CI, 42–58%) after correcting for misclassification of vaccination status.

4.3. Comparing FOREVER Cohort VE estimates with other test-negative studies

When stratified by influenza season, healthcare setting, and influenza type/subtype, VE estimates for older adults generated using the FOREVER Cohort (corrected for misclassification of vaccination status) were comparable to those from other studies (Table S11). In particular, when comparing the FOREVER Cohort estimates to those from Canadian prospective test-negative VE networks (i.e., Sentinel Practitioner Surveillance Network [SPSN; outpatients] and Serious Outcomes Surveillance Network [SOS; inpatients], both of which collect data from several Canadian provinces, including Ontario), the point estimates were similar (i.e., within 10%) and/or the 95%CIs overlapped substantially for most seasons and types/subtypes. For 2014–15, VE against H3N2 appeared to be higher for the FOREVER Cohort than SOS [12], but the estimates against both H3N2 and B were similar between the FOREVER Cohort and SPSN [39].

5. Discussion

We explored the validity of using the routinely collected laboratory and health administrative data contained in the FOREVER Cohort for estimating influenza VE by testing for the presence of selected information and selection biases. We found that use of specimens with an unknown interval from illness onset to specimen collection is unlikely to bias VE estimates due to false-negative results. We also found the proportions of patients tested for influenza to be comparable between vaccinated and unvaccinated individuals, but aORs were significantly greater than 1.00 in some instances, and this was more pronounced when we corrected for misclassification of vaccination status, suggesting that some selection bias may be present under certain circumstances. Both of these biases seemed to be less evident for inpatients, and most of the FOREVER Cohort specimens are obtained from this healthcare setting, based on the algorithm used to assign the healthcare setting for specimen collection. VE estimates were similar irrespective of setting and other conditions intended to emulate prospective test-negative studies, but the quantitative sensitivity analysis that corrected for misclassification of vaccination status led to consistently higher VE estimates. VE estimates generated using the FOREVER Cohort were comparable to those from other studies for most seasons, settings, and types/subtypes, considering the heterogeneity of published estimates and the wide confidence intervals often encountered; differences between studies could be attributed to numerous factors, including variations in patient age distribution, presence of comorbidities, vaccine composition, circulating influenza viruses, prior population exposure to both viruses and vaccines, and variables included in the multivariable adjustment. Indeed, VE estimates from the FOREVER Cohort were rarely the outlier among published estimates. Taken together, these results suggest that the FOREVER Cohort can be used to validly estimate VE for older adults, particularly for inpatients, and these estimates could complement those generated by prospective studies.

The main strength of the FOREVER Cohort is that it represents a relatively inexpensive and efficient way to assemble a large sample of individuals who have been tested for influenza and other respiratory viruses, and to collect data on a wide range of covariates. With the very high linkage rate, we expect minimal bias in the cohort arising from differential linkage of tested patients. The population-based ICES data holdings permit identification of large

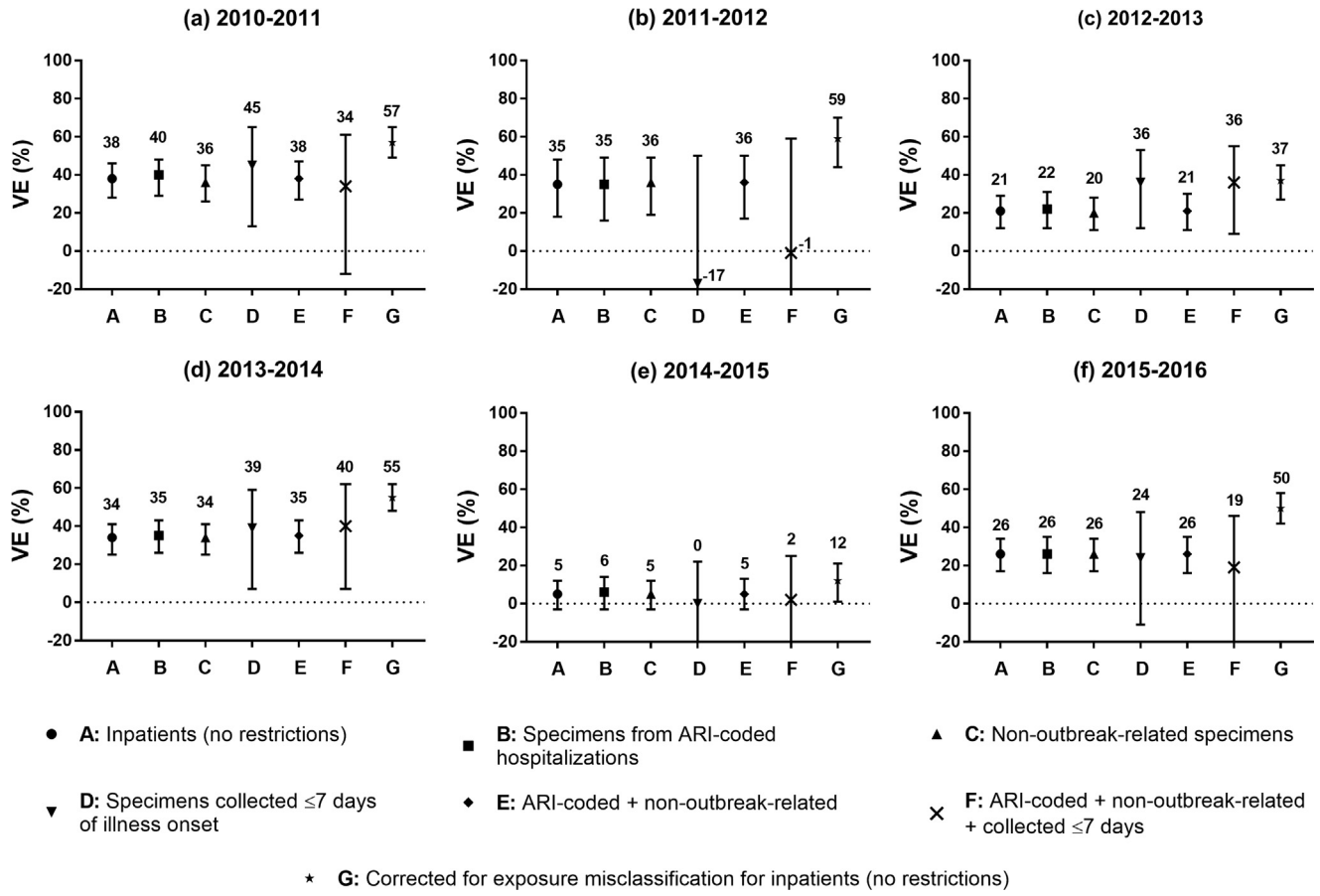

Fig. 2. Forest plot of influenza vaccine effectiveness estimates for adults aged > 65 years using specimens obtained from inpatients and tested by any laboratory test method, estimated under various conditions, and corrected for exposure misclassification, for the 2010–11 (panel a), 2011–12 (panel b), 2012–13 (panel c), 2013–14 (panel d), 2014–15 (panel e), and 2015–16 (panel f) influenza seasons. VE = vaccine effectiveness; ARI = acute respiratory illness.

numbers of people with certain medical conditions (e.g., diabetes, cancer, chronic obstructive pulmonary disease), which will facilitate studies to assess VE in specific high-risk patient populations. Further, most specimens were tested using molecular (PCR) assays, offering optimal sensitivity and specificity.

The FOREVER Cohort also has some weaknesses. First, clinical information related to the acute illness (e.g., illness onset date, presenting symptoms) is often unavailable, preventing the use of a standardized case definition as done in most prospective test-negative studies. Instead, we used diagnostic codes for ARIs as a proxy and found that VE estimates were similar for ARI-coded and non-ARI-coded hospitalizations. Second, the laboratory testing data represent the metaphorical tip of the iceberg, capturing more severe infections while missing a far larger number of milder infections that do not result in testing, but this is also the case with prospective test-negative studies. Nevertheless, determining VE estimates for preventing more severe illness is policy-relevant. Third, testing assays and protocols varied across the contributing laboratory network, but are becoming increasingly uniform over time. All testing methods used are fully validated and have well-defined performance characteristics, as required by provincial laboratory licensing and accreditation bodies, and provincial guidance on patient testing supports uniform practice. Fourth, we must rely on physician and pharmacist billing claims to ascertain vaccination status, and although we are able to perform sensitivity analyses to correct for misclassification of vaccination status, the sensitivity and specificity parameters used for these sensitivity analyses were obtained from a validation study of 2007–2009 data from respondents to the Canadian Community Health Survey that included only physician billing claims, because pharmacists were not yet authorized to administer influenza vaccines. Although these data have imperfect sensitivity due to the availability of influenza vaccines in settings other than physician offices and pharmacies, a recent simulation study demonstrated that reduced specificity of the vaccination status measure has a greater impact on VE estimates than reduced sensitivity [40]. Fifth, we are missing information on potential confounders that are usually unavailable in administrative data (e.g., smoking status, body mass index). However, a study that used a causal approach to develop a parsimonious model for test-negative studies found that only age, calendar time, immunocompromising comorbid conditions, year, and study site (where applicable) were required to produce theoretically unbiased VE estimates [41]. Sixth, the delays in the availability of the administrative data and in obtaining and linking the laboratory data preclude timely estimation of VE, which make these data less useful for rapid assessments of current season VE. Although approaches to reduce this delay are actively being explored, the estimates that these data can produce for specific high-risk populations need not necessarily be timely, since they will involve combining data from multiple influenza seasons in order to have a sufficiently large sample. Finally, use of data from the FOREVER Cohort for test-negative studies is subject to the same potential biases and limitations as traditional test-negative studies and other observational studies, and we were only able to test for the presence of two types of bias in this study.

In summary, we created the FOREVER Cohort by linking routinely collected laboratory and health administrative data at the individual level, and we explored the validity of using these data

for estimating influenza VE for community-dwelling older adults. We encourage others who have access to similar data to evaluate their data for estimating VE and, if valid, to use them to better characterize VE for populations and outcomes that are difficult to assess in prospective test-negative studies.

Funding/Support

This work was supported by an operating grant from the Canadian Institutes of Health Research (CIHR MOP 130568). JCK 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. This study was also supported by Public Health Ontario (PHO) and ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). 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.

Disclaimers

The opinions, results, and conclusions reported in this paper are those of the authors and are independent from the funding sources. Parts of this material are based on data and/or information compiled and provided by the Canadian Institute for Health Information (CIHI) and by Cancer Care Ontario (CCO). However, the analyses, conclusions, opinions, and statement expressed herein are those of the authors, and not necessarily those of CIHI or CCO. No endorsement by ICES, PHO, MOHLTC, CIHI, or CCO is intended or should be inferred.

Declaration of Competing Interest

MLJ has received research funds from Sanofi Pasteur to conduct surveillance for respiratory syncytial virus. AJM has received research funds from GSK and Sanofi Pasteur. MS has received research grants from Janssen Canada for respiratory virus clinical trials. JBG has received research grants from GSK and Hoffmann-LaRoche for antiviral resistance studies, and from Pfizer Inc. to conduct microbiological surveillance of *Streptococcus pneumoniae*. All other authors report no conflicts.

Acknowledgements

We thank IMS Brogan Inc. for use of their Drug Information Database.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.vaccine.2019.06.011.

References

[1] Barr IG, Russell C, Besselaar TG, et al. WHO recommendations for the viruses used in the 2013–2014 Northern hemisphere influenza vaccine: epidemiology, antigenic and genetic characteristics of influenza A(H1N1)pdm09, A(H3N2) and B influenza viruses collected from October 2012 to January 2013. Vaccine 2014;32(37):4713–25.
[2] Sullivan SG, Feng S, Cowling BJ. Potential of the test-negative design for measuring influenza vaccine effectiveness: a systematic review. Exp Rev Vacc 2014;13(12):1571–91.
[3] Skowronski DM, Masaro C, Kwindt TL, et al. Estimating vaccine effectiveness against laboratory-confirmed influenza using a sentinel physician network: results from the 2005–2006 season of dual A and B vaccine mismatch in Canada. Vaccine 2007;25(15):2842–51.
[4] Ohmit SE, Thompson MG, Petrie JG, et al. Influenza vaccine effectiveness in the 2011–2012 season: protection against each circulating virus and the effect of prior vaccination on estimates. Clin Infect Dis 2014;58(3):319–27.
[5] Kissling E, Valenciano M, Falcao J, et al. "I-MOVE" towards monitoring seasonal and pandemic influenza vaccine effectiveness: lessons learnt from a pilot multi-centric case-control study in Europe. Euro Surveill 2009;14(44):2008–9.
[6] De Serres G, Skowronski DM, Wu XW, Ambrose CS. The test-negative design: validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. Euro Surveill 2013;18(37):20585.
[7] Foppa IM, Haber M, Ferdinands JM, Shay DK. The case test-negative design for studies of the effectiveness of influenza vaccine. Vaccine 2013;31(30):3104–9.
[8] Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. Vaccine 2013;31(17):2165–8.
[9] 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(5):345–53.
[10] Blyth CC, Macartney KK, Hewagama S, et al. Influenza epidemiology, vaccine coverage and vaccine effectiveness in children admitted to sentinel Australian hospitals in 2014: the influenza complications alert network (FluCAN). Euro Surveill 2016;21(30).
[11] Huang QS, Turner N, Baker MG, et al. Southern hemisphere influenza and vaccine effectiveness research and surveillance. Influenza Other Respir Viruses 2015;9(4):179–90.
[12] McNeil SA, Andrew MK, Ye L, et al. Interim estimates of 2014/15 influenza vaccine effectiveness in preventing laboratory-confirmed influenza-related hospitalisation from the serious outcomes surveillance network of the Canadian immunization research network, January 2015. Euro Surveill 2015;20(5):21024.
[13] Rondy M, Castilla J, Launay O, et al. Moderate influenza vaccine effectiveness against hospitalisation with A(H3N2) and A(H1N1) influenza in 2013–14: results from the InNHOVE network. Hum Vaccin Immunother 2016;12(5):1217–24.
[14] 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(6):820–7.
[15] 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] Schwartz KL, Jembere N, Campitelli MA, Buchan SA, Chung H, Kwong JC. Using physician billing claims from the Ontario health insurance plan to determine individual influenza vaccination status: an updated validation study. CMAJ Open 2016;4(3):E463–70.
[17] Buchan SA, Rosella LC, Finkelstein M, et al. Impact of pharmacist administration of influenza vaccines on uptake in Canada. CMAJ 2017;189(4):E146–52.
[18] Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying patients with physician-diagnosed asthma in health administrative databases. Can Respir J 2009;16(6):183–8.
[19] Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying individuals with physcian diagnosed COPD in health administrative databases. COPD 2009;6(5):388–94.
[20] Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. Diabetes Care 2002;25(3):512–6.
[21] Hall S, Schulze K, Groome P, Mackillop W, Holowaty E. Using cancer registry data for survival studies: the example of the Ontario cancer registry. J Clin Epidemiol 2006;59(1):67–76.
[22] 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.
[23] Quinn RR, Laupacis A, Austin PC, et al. Using administrative datasets to study outcomes in dialysis patients: a validation study. Med Care 2010;48(8):745–50.
[24] Lam NN, McArthur E, Kim SJ, Knoll GA. Validation of kidney transplantation using administrative data. Can J Kidney Health Dis 2015;2:20.
[25] Antoniou T, Zagorski B, Loutfy MR, Strike C, Glazier RH. Validation of case-finding algorithms derived from administrative data for identifying adults living with human immunodeficiency virus infection. PLoS ONE 2011;6(6):e21748.
[26] 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 Alzheimers Dis 2016;54(1):337–49.
[27] Urquhart R, Giguere AMC, Lawson B, et al. Rules to identify persons with frailty in administrative health databases. Can J Aging 2017;36(4):514–21.
[28] Schultz SE, Rothwell DM, Chen Z, Tu K. Identifying cases of congestive heart failure from administrative data: a validation study using primary care patient records. Chronic Dis Inj Can 2013;33(3):160–6.
[29] The Canadian Stroke Strategy. Canadian Stroke Strategy Core Performance Indicator Update 2010. The Canadian Stroke Strategy 2010 [cited 2016 Dec 14];Available from: URL: http://strokebestpractices.ca/wp-content/uploads/2010/12/Stroke_Core_ENG.pdf.
[30] Gurevich Y, McFarlane A, Morris K, Jokovic A, Peterson GM, Webster GK. 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(7):e249–53.
[31] 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(9):867–73.
[32] Juurlink D, Preyra C, Croxford R, et al. Canadian Institute for Health Information Discharge Abstract Database: A Validation Study. Institute for Clinical Evaluative Sciences; 2006.
[33] 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(25):2408–15.
[34] Carrat F, Vergu E, Ferguson NM, et al. Time lines of infection and disease in human influenza: a review of volunteer challenge studies. Am J Epidemiol 2008;167(7):775–85.
[35] Lau LL, Cowling BJ, Fang VJ, et al. Viral shedding and clinical illness in naturally acquired influenza virus infections. J Infect Dis 2010;201(10):1509–16.
[36] Ferdinands JM, Belongia EA, Nwasike C, Shay DK. Influenza vaccination status is not associated with influenza testing among children: implications for observational studies of vaccine effectiveness. Vaccine 2011;29(10):1935–40.
[37] Jackson ML, Rothman KJ. Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness. Vaccine 2015;33(11):1313–6.
[38] Fox MP, Lash TL, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. Int J Epidemiol 2005;34(6):1370–6.
[39] 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(1):21–32.
[40] Jackson ML. Use of self-reported vaccination status can bias vaccine effectiveness estimates from test-negative studies. Vaccine: X 2019;1:100003.
[41] Lane CR, Carville KS, Pierse N, Kelly HA. Seasonal influenza vaccine effectiveness estimates: development of a parsimonious case test negative model using a causal approach. Vaccine 2016;34(8):1070–6.