Utility of neuraminidase inhibitor dispensing data as a tool for influenza surveillance

Jesse Papenburg, MD

Department of Epidemiology, Biostatistics and Occupational Health

McGill University

Montreal, Quebec

August 2014

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of

Master of Science (MSc.)

TABLE OF CONTENTS

ABSTRACT

Surveillance performed using routinely collected electronic data offers advantages that include a short reporting delay and a low acquisition cost. Monitoring of neuraminidase inhibitor (NI) dispensing in community pharmacies has emerged as a possible automated information source for influenza surveillance. However, little is known about the utility of these data for monitoring influenza activity. Therefore, we aimed to evaluate the timeliness, correlation, and predictive accuracy of community pharmacy NI dispensing in relation to laboratory-confirmed influenza activity in Quebec, Canada, during 2010-2013. Our secondary objective was to compare the characteristics of NI dispensing to those of visits for influenza-like illness (ILI) in emergency departments (ED), a commonly used source of surveillance data.

Provincial weekly counts of positive influenza laboratory tests were used as a reference measure for the level of influenza circulation. We applied ARIMA models to account for seasonality and computed cross-correlation functions to measure the strengths of association and lead-lag-relationships of NI dispensing and ILI ED visits to our reference indicator. Finally, using an ARIMA model, we evaluated the ability of NI dispensing and ILI ED visits to predict laboratory–confirmed influenza.

NI dispensing was significantly correlated (R=0.68) with influenza activity with no lag; the earliest statistically significant correlation occurred with a lead-time of 1 week. The maximal correlation of ILI ED visits was not as strong (R=0.50), but occurred with a lead-time of 1 week. Both NI dispensing and ILI ED visits were significant predictors of laboratory-confirmed influenza in a multivariable model; the predictive potential was greatest when NI counts were lagged to precede laboratory surveillance by two weeks.

We conclude that NI dispensing data can provide timely and valuable information for the surveillance of influenza at the provincial level.

RÉSUMÉ

La surveillance qui s'appuie sur des données électroniques recueillies et enregistrées en routine offre des avantages tels qu'un court retard de déclaration ainsi qu'un faible coût d'acquisition. La surveillance des ventes au détail de médicaments sur prescription contre l'influenza, les l'inhibiteurs de la neuraminidase (NI), a émergé comme une source possible d'information automatisée pour la vigie sanitaire de la grippe. Toutefois, les caractéristiques de la performance de ces données comme objet de surveillance ne sont pas bien connues. Dès lors, nous avons cherché à évaluer les données de distribution des NI dans les pharmacies communautaires comme un nouvel outil de surveillance de l'influenza, de par leur relation d'ordre temporelle (décalage), de leur corrélation et de leur capacité prédictive, en comparaison à l'activité grippale confirmée en laboratoire, au Québec, Canada, de 2010 à 2013. Notre objectif secondaire était de comparer ces caractéristiques à celles de la surveillance des visites pour syndrome d'allure grippal (SAG) inscrites aux d'urgences.

Les données hebdomadaires provinciaux du nombre de tests de laboratoire positifs pour l'influenza ont été utilisées comme mesure de référence pour le niveau d'activité grippale. Nous avons appliqué la méthodologie de modélisation ARIMA pour tenir compte de la saisonnalité et de l'autocorrélation. Nous avons ensuite calculé les fonctions de contre-corrélation pour mesurer les forces d'association et explorer les relations temporelles entre la distribution des NI et les visites SAG avec notre mesure de référence. Enfin, nous avons évalué la valeur prédictive de la distribution des NI et des visites SAG dans le montage d'un modèle ARIMA pour les comptes d'influenza confirmés en laboratoire.

La distribution des NI était significativement corrélée (R = 0,68) avec l'activité grippale au temps de latence zéro; la première corrélation statistiquement significative a eu lieu avec un décalage anticipatoire d'une semaine. La corrélation maximale des visites SAG n'était pas aussi forte (R = 0,50), mais a culminé une semaine plus tôt que les distributions NI. Tant la distribution des NI et les visites SAG à l'urgence étaient des variables prédictives significatives dans un modèle multivarié de cas confirmés en laboratoire; le potentiel prédictif du modèle était maximal lorsque les distributions NI ont été décalées pour précéder la surveillance en laboratoire de deux semaines.

Ainsi, nous concluons que les données de distribution NI peuvent fournir des informations utiles et en temps opportun pour la surveillance de la grippe à l'échelle provinciale.

I dedicate this work to my wife, Patricia. Her love and support are priceless; they elevate all aspects of my life.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. David Buckeridge for his mentorship, guidance and ever-pleasant mood. I would particularly like to express my gratitude to him for his flexibility and understanding as I faced challenges over the past two years. I am also grateful to my thesis committee member, Dr. Gaston De Serres, for asking difficult and insightful questions.

The McGill Surveillance Lab is a true reflection of its leader; it is composed of caring, bright, witty, academically diverse and curious researchers and researchers-to-be. Spending time with this group was rewarding, humbling, and mostly just plain fun. I especially thank Dr. Katia Charland for her kindness and availability in showing me the ropes of ARIMA modelling.

I thank IMS Brogan for providing the neuraminidase inhibitor dispensing data from their Weekly CompuScript database, free of charge.

I am very grateful to the *Institut national de santé publique du Québec* (INSPQ) and the *Laboratoire de santé publique du Québec* (LSPQ) for granting us permission to use their surveillance data for this thesis project. Specifically, I thank Drs. Michel Couilllard and Hugues Charest of the LSPQ for the laboratory surveillance data. Dr Couillard was the initiator of sentinel laboratory surveillance in Québec and, at the moment of his sudden passing in 2013, the associate-director of the LSPQ. I also thank Dr Monique Douville-Fradet, from the *Direction des risques biologiques et de la santé au travail*, INSPQ, for kindly providing the RQSUCH data on ED visits for ILI.

Finally, I thank Dr Patricia Fontela, for her thoughtful reviews of early drafts of this thesis and of its manuscript.

CONTRIBUTION OF AUTHORS

I developed the original research question and objectives for this thesis in collaboration with my thesis supervisor, Dr. David Buckeridge. With the guidance of my supervisor, I selected the design and the methods used in the study. I obtained the data used in this thesis and conducted all data processing and statistical analyses; I was responsible for the first interpretation of the results. After discussing the results with my supervisor and co-authors, I wrote the manuscript, which was revised and edited by Dr. Buckeridge as well as the co-authors. Finally, I wrote all chapters of this MSc thesis.

Dr. Buckeridge, as my primary supervisor, was involved in supervising all stages of this research. Dr. De Serres, as thesis committee member, provided feedback on the study design and the interpretation of the results, and made editorial revisions to the manuscript. Dr. Katia Charland provided advice on the choice and execution of the statistical methods, the interpretation of the results, and made editorial revisions to the manuscript.

LIST OF ABBREVIATIONS

AICc: corrected Akaike information criterion

AR: autoregressive

ARIMA: autoregressive integrated moving average

CCF: cross-correlation function

CDC: United States Centers for Disease Control

CI: confidence interval

ED: emergency department

ILI: influenza-like illness

INSPQ: *Institut national de santé publique du Québec*

LSPQ: *Laboratoire de santé publique du Québec*

MA: moving average

OR: odds ratio

PHAC: Public Health Agency of Canada

RQSUCH: *Relevé quotidien de la situation à l'urgence et au centre hospitalier*

RSV: respiratory syncytial virus

RT-PCR: reverse-transcription polymerase chain reaction

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 - INTRODUCTION

1.1 Background

Surveillance is a cornerstone of public health practice [1]. It is a continuous and dynamic process that involves the collection, analysis, interpretation and dissemination of health data with the final objective of using the results to prevent and control disease [2]. The best recognized use of public health surveillance data is the monitoring of trends for the detection of epidemics of communicable diseases [3, 4].

Among infectious diseases requiring public health monitoring, influenza figures prominently due to the large burden of this disease and its ability to evolve and escape population immunity over time [5]. During a typical seasonal epidemic in Canada, influenza causes 35 hospital admissions per 100,000 persons [6], and 3,500 (95%CI, 3,200 - 3,700) deaths [7]. Influenza prevention and control programs aim to reduce the overall burden of disease – but especially among sub-populations that are at highest risk of serious disease – through yearly vaccination campaigns and the targeted use of antivirals [8, 9]. To achieve this aim effectively, public health agencies require up-to-date and accurate influenza surveillance information. However, timeliness and validity vary between data sources [10, 11].

Traditional data sources for influenza surveillance, such as influenza laboratory tests and cases of influenza-related illnesses seen by sentinel physicians, are typically associated with a reporting delay of as much as 1-2 weeks [12]. Consequently, public health researchers have sought novel electronic data sources that could complement traditional sources and serve as a leading indicator for influenza activity by providing earlier information for rapid outbreak detection and near real-time situational awareness [10, 13]. Examples of potential non-traditional

data streams for influenza surveillance include: absenteeism reports [14, 15], over-the-counter and prescription drug sales [16-18], online activity monitoring [19, 20] and health advice calls [21, 22].

1.2 Rationale

It is estimated that there are over 2 million clinic visits per year for influenza in the United States [23]. Of these, approximately 20% receive a prescription for antiviral treatment. In Canada, the neuraminidase inhibitor (NI) class of medications (oseltamivir and zanamivir) is the only recommended first-line therapy for influenza since 2006 because of widespread resistance to adamantanes [8]. Prescription drug sales databases are potentially easy and fast to access because prescribed drugs are subject to electronic adjudication by insurance companies at the time of dispensing. Therefore, the monitoring of NI dispensing data from community pharmacies has emerged as a possible automated data source for surveillance of influenza, and has been used as an influenza indicator by the Public Heath Agency of Canada (PHAC) since 2012-13 [24]. However, there is currently little information about the performance characteristics of these data as a tool for influenza surveillance. Furthermore, no evaluation of antiviral dispensing to date has controlled for autocorrelation (i.e., the lack of independence between data points in a time series), a requirement for valid inferences from seasonal health data [25, 26].

Because influenza activity in the outpatient setting tends to occur before an increased incidence of more severe disease [27-29], ***we hypothesized that changes in the weekly volume of outpatient antiviral prescriptions might precede changes in weekly counts of positive influenza tests and that NI dispensing could serve as an early indicator of epidemic influenza activity***.

1.3 Objectives

1. The primary objective of this thesis was to evaluate retail pharmacy NI dispensing data as a novel automated influenza surveillance tool in Quebec, Canada, during 2010-2013. In particular, we assessed the timeliness, correlation, and predictive accuracy of NI dispensing data in relation to a reference standard for influenza activity, laboratory-confirmed cases of influenza.

2. Our secondary objective was to compare the characteristics of the NI dispensing data to those of emergency department (ED) visits for influenza-like illness (ILI), a currently-employed method for influenza surveillance.

CHAPTER 2 - LITERATURE REVIEW

2.1 Influenza

2.1.1 Burden of disease

In temperate climates, respiratory viruses circulate with a marked seasonality, with peak activity during the winter months and significant temporal overlap between different viruses [30, 31]. Influenza viruses (influenza type A [sub-types A/H1N1 and A/H3N2] and type B) cause seasonal outbreaks (also referred to as seasonal epidemics) sometime between November and April of each year in North America [32, 33]. However, the exact timing of an outbreak is variable and difficult to predict, as are its overall severity, peak intensity, vaccine field effectiveness and the distribution of its circulating strains [34].

During a typical seasonal epidemic, influenza causes symptomatic infection in approximately 5-30% of the population, with the highest rates of infection in children aged 5–9 years [35, 36]. The manifestations vary widely and may include upper respiratory illness (cold-like symptoms, pharyngitis, acute otitis media or sinusitis) community-acquired pneumonia, exacerbation of underlying medical conditions (e.g., chronic obstructive pulmonary disease, asthma, and congestive cardiac failure), fulminant respiratory disease, and death [5]. The likelihood of each of these outcomes is determined by many factors, including patient age, pre-existing immunity (through prior infection or vaccination), the presence of comorbidities or immune suppression and characteristics of the virus itself.

The burden of influenza on a population is primarily attributable to a combination of the clinical severity of illness and the number of persons infected [34, 37]. These factors can vary between seasonal epidemics as a function of population immunity as well as circulating strain

virulence and transmissibility [38, 39]. Recent Canadian studies have estimated that, nationally, seasonal influenza is responsible for 35 (95% CI, 31–39) respiratory admissions per 100,000 persons in an average year [6], and that approximately 3,500 (95%CI, 3,200 - 3,700) deaths are attributable to influenza annually [7]. The risks of complicated disease causing serious illness, hospitalization or death are highest in infants (age <1 year), the elderly (age ≥65 years), and persons with underlying medical conditions [9, 40-45]. In North America, seasonal influenza hospitalisation rates in infants and the elderly are approximately 2.5 – 5 times that of the general population [6, 40]. In industrialised countries, approximately 90% of seasonal influenza-associated deaths occur among patients ≥65 years old [46]. Worldwide, annual epidemics are estimated to result in about 3 to 5 million cases of severe illness, and about 250,000 to 500,000 deaths [47].

In contrast to seasonal epidemics, in which circulating strains have slowly evolved so as to partially escape population immunity, influenza pandemics are caused by the global spread of a novel influenza A virus to which the majority of the population has no prior immunity [5]. Because the characteristics of the new virus are not well known early in the course of a pandemic, they are necessarily unpredictable, with attack rates and disease severity that may be similar or considerably greater than during a seasonal outbreak [37]. The age distribution of the population affected and the timing of peak activity can also differ from a typical season.

2.1.2 Diagnosis: clinical and laboratory

It is difficult to accurately diagnose influenza in clinical practice; laboratory testing is required for definitive diagnosis [48, 49]. The clinical syndrome typically associated with influenza, influenza-like illness (ILI), has been defined differently, but definitions tend to be similar across most researchers and surveillance systems. In Canada, PHAC defines ILI as an

"acute onset of respiratory illness with fever and cough and with one or more of the following - sore throat, arthralgia, myalgia, or prostration which is likely due to influenza" [50], whereas the United States Centers for Disease Control and Prevention's (CDC) definition is "fever (temperature of 100°F [37.8°C] or greater) and a cough and/or a sore throat without a *known* cause other than influenza" [51]. However, other respiratory viruses that temporally co-circulate with influenza, such as respiratory syncytial virus (RSV), also frequently cause ILI [52]. Therefore, ILI is less specific than laboratory confirmed influenza and estimates of the positive predictive value (PPV) of ILI during periods of influenza activity have varied. In predominantly young healthy adult outpatients, the PPV of ILI for influenza infection can range from 77% to 87%, depending on variations of the ILI definition [53, 54]. Yet the specificity of ILI can be low, as influenza virus was the cause of ILI (PHAC definition) in only 652 of 1501 (43%) participants that presented to a Canadian community-based sentinel clinic surveillance system in 2012-13 [55]. Because of the wide range of symptoms caused by influenza, estimates of the sensitivity of ILI in outpatients have generally centred around 60-80% [53, 54, 56, 57]. Clinical predictors of influenza do not perform as well in young children (<5 years old) [58] and in the elderly [59]. In the former, the incidence of other respiratory viruses, especially RSV, that cause fever and cough in both the outpatient and hospital settings is much higher than that of influenza [60]. Consequently, the PPV of ILI (CDC definition) has been reported to be as low as 20% (95% CI, 17%–23%) among children aged 6–59 months [61]. In patients >60 years old with influenza, fever is frequently absent; therefore, the sensitivity of ILI may be as low as 30% in this population [59].

Laboratory diagnosis of influenza is based on the identification of the virus in a patient's respiratory secretions by one of three methods [62]:

1) ***Molecular assays*** such as reverse-transcription polymerase chain reaction (RT-PCR) are considered the reference standard due to their very high analytical and clinical sensitivity and specificity [49]. Although these assays may require less than 2 hours of analytical time, turn-around time for results may be much longer because specimens may need to be sent to specialized laboratories and testing may be performed in batches due to cost considerations.

2) ***Direct detection of viral antigen*** by immunofluorescence or rapid immunoassays. Immunofluorescence testing is fast (analytical time, ~1h) but requires considerable technical skill, is more subjective and is also less sensitive (80-90%) compared to PCR. Rapid immunoassays are the fastest and simplest method and could potentially be performed at the site of care; however, their sensitivity (highly variable: 40-85%) and specificity (>90-95%) are the poorest of all techniques, especially when used outside of the pediatric population [63].

3) ***Virus isolation in culture*** used to be the gold standard method and is still required for full phenotypic characterization of a strain's antigenic properties and antiviral susceptibility. However, since the advent of RT-PCR and antigen detection techniques, culture has fallen out of favor because of its lower sensitivity (80-90% vs. RT-PCR) and long delays (>48h) to produce results [48].

Despite the difficulties in establishing a clinical diagnosis, the vast majority of influenza diagnoses will not be confirmed microbiologically, principally for two reasons [49, 57, 64]. First, most test results are not available in a timely manner for management decisions in the outpatient setting. Rapid immunoassays could provide results within 20 minutes, but are rarely used at the point of care in clinics and EDs in Quebec and Canada because of concerns regarding accuracy, costs and quality control outside of the laboratory setting [63, 65]. Second, during an epidemic, the presence of a clinical predictor such as ILI is considered to produce a high enough likelihood of true infection that confirmation is usually not required to guide management of outpatients

[49]. The Infectious Diseases Society of America clinical practice guidelines for the management of seasonal influenza provide testing indications based on: patient age, signs and symptoms, time since onset of symptoms, immune status, presence of high risk comorbidities, disease severity (outpatient vs. hospitalized), and the presence of circulating influenza in the community [64]. The net result is that the overwhelming majority of influenza tests are performed on hospitalized patients or patients that are likely to be hospitalized for their current illness.

2.1.3 Influenza prevention and control: vaccination

Immunization is the cornerstone of influenza prevention. In Canada, annual influenza immunisation campaigns preferentially target categories of individuals at high-risk of complications; however, the vaccine is also available and recommended to the general public [9]. The antigenic characteristics of current and emerging influenza virus strains provide the basis for selecting the three strains included in each year's trivalent vaccine (one representative strain of each of: A/H1N1, A/H3N2 and B). Vaccination should be performed yearly for two reasons. First, influenza viruses evolve over time and the contents of the vaccine are re-evaluated yearly to reflect circulating strains. Second, vaccine-induced immunity wanes. Even if circulating strains have not changed, protective antibody levels may not last two influenza seasons [66]. In North America, vaccines are available for distribution in the late Fall, just prior to the onset of the seasonal epidemic.

Jefferson et al from the Cochrane Collaboration reviewed pediatric randomized clinical trial (RCT) data and found that the efficacy of influenza immunisation against laboratory-confirmed disease was 59% (95% CI, 41% - 71%) for inactivated vaccines in children >6 months and 80% (95% CI, 68% - 87%) for live attenuated vaccines in children >2 years [67]. A separate Cochrane review of RCTs in healthy adults reported overall efficacy of inactivated vaccines in preventing

confirmed influenza of 60% (95% CI, 53% - 66%) [68]. Efficacy is thought to be lower in the elderly; however, "given the heterogeneous nature of the vaccines tested (monovalent, trivalent, live, or inactivated aerosol vaccines), setting, follow up and outcome definition, no firm conclusions can be drawn from this body of evidence" according to the 2010 Cochrane publication on the subject [69].

Vaccine field effectiveness is in large part a function of the antigenic match between the strains selected for inclusion in the vaccine and those that actually circulated. Since 2004, estimates of vaccine effectiveness against medically-attended, laboratory-confirmed influenza in Canada have been produced using a test-negative case-control design embedded within a sentinel surveillance network of several hundred community-based practitioners from 5 provinces, including Quebec [70]. Component-specific results from the three seasons that correspond to the study period of this thesis (2010-11, 2011-12 and 2012-13) are presented in Table 2-1 [55, 71, 72]. During the course of our study, influenza vaccine effectiveness varied considerably from season to season (overall and by vaccine strain) and across vaccine components during a season.

Table 2-1. Overall and component-specific vaccine effectiveness estimates against medically-attended, laboratory-confirmed influenza from a Canadian sentinel surveillance network: 2010-11, 2011-12 and 2012-13.

Vaccine component	Vaccine effectiveness, %; (95% CI)		
	2010-11	**2011-12**	**2012-13**
Overall	36 (17–51)	58 (44–69)	52 (38-64)
A/H1N1	60 (19–80)	80 (54–91)	60 (21-80)
A/H3N2	35 (11–53)	63 (36–78)	44 (24-59)
B	28 (−8 to 52)	45 (21–62)	68 (47-81)

CI, confidence interval

Data from references [55, 71, 72].

2.1.4 Influenza prevention and control: antiviral therapy

Influenza is the only respiratory virus with commercially available specific antiviral therapy. Since 2000, the influenza neuraminidase inhibitors (NI) oseltamivir (Tamiflu®, Hoffman-La Roche, Limited, Mississauga, ON, Canada) and zanamivir (Relenza®, GlaxoSmithKline Inc., Mississauga, ON, Canada) have been licensed in Canada for the treatment of influenza infection [8]. Resistance to the adamantane class of antivirals is widespread and their use is no longer recommended as a first-line agent for the treatment or prevention of influenza since 2006.

A recent Cochrane review of published and unpublished RCT data reported that early (within 48 hours of symptom onset) treatment of outpatients with oseltamivir reduces symptom

duration by 17 hours (95% CI, 8 to 25 hours) [73]. Another systematic review and meta-analysis by independent investigators concluded that oseltamivir also reduces the risk of lower respiratory tract complications requiring antibiotic treatment by 37% (95% CI, 18%-52%) [74]. Given the relatively modest benefits for outpatients, Canadian and United States guidelines recommend that antiviral treatment preferentially target persons with suspected or confirmed influenza who are at higher risk for influenza complications because of age or underlying medical conditions [8, 75]. Regarding serious outcomes, observational studies performed among hospitalised patients in Canada suggest that NI treatment, particularly earlier NI treatment, significantly reduces the odds of intensive care unit admission or death [76-78]. A multinational cohort study by Muthuri et al found that, among patients admitted for A/H1N1 infection since 2009, when comparing NI treatment vs. no treatment, early vs. later treatment and early vs. no treatment, all significantly reduced mortality odds [79]. Based on this evidence, treatment of any person with confirmed or suspected influenza who requires hospitalization is recommended, even if the patient presents more than 48 hours after illness onset [8, 75]. Finally, NIs are licensed for both pre-exposure and post-exposure prophylaxis for high-risk patients; however, since the 2009 A/H1N1 pandemic, early treatment is preferred over prophylaxis due to concerns regarding emergence of drug resistance in Canada [8, 80].

2.2 Surveillance of influenza

2.2.1 Overview of influenza surveillance

Surveillance has been described as a continuous and dynamic process that involves the systematic collection, analysis, interpretation and dissemination of data with the final objective of using the results to improve health through the prevention and control of disease [2]. Information

obtained though surveillance can guide immediate public health action, inform longer-term program planning, and stimulate the formulation of research questions.

The best recognized use of public health surveillance data is the monitoring of trends relating to communicable diseases for the detection of outbreaks, i.e., increases in incidence above the expected or background rate for the disease [3, 4, 81]. Infectious disease surveillance by public health agencies has traditionally depended on voluntary or mandatory reporting of cases by physicians, hospitals and laboratories [3] Longstanding traditional data sources for influenza surveillance include: 1) medically-attended influenza-related illnesses in sentinel ambulatory care networks; 2) positive influenza laboratory tests from sentinel diagnostic laboratories; 3) hospital admissions for pneumonia and influenza (P & I) as recorded in hospital discharge summaries; and 4) deaths due to P&I from death certificates [12]. However, these mostly manual approaches to surveillance can be insensitive, inflexible, and slow. For laboratory surveillance, the inherent delays associated with laboratory processing and testing (although modern influenza diagnostic techniques offer shorter delays compared with traditional viral culture), the manual submission of reports, and subsequent data analysis and dissemination may total 2 weeks or more [82, 83]. Reports form ambulatory clinics are also delayed by manual data entry. Adjudication and compilation of hospital discharge summaries and death certificates can take weeks to months. Consequently, recent research has focused on identifying non-traditional, automated data sources for influenza surveillance that can provide an early indication of influenza outbreaks and near real-time situational awareness of influenza activity for clinicians and public health officials [10, 13].

2.2.2 Syndromic surveillance for influenza

Syndromic surveillance is a rapidly evolving field within public health practice. It involves the monitoring of pre-diagnostic data associated with the disease under surveillance, such as clinical syndromes (e.g, ILI), health indicators of different actions persons might take (e.g., visit an ED), or consequences they might suffer (e.g., absence from work) [84]. Syndromic surveillance systems typically seek to use existing electronic health data for near-real time monitoring involving automated methods of data collection, transmission and analysis [85].

For the monitoring of influenza, syndromic information sources include ED chief complaints [86, 87], outpatient clinic visits [27, 29], billing data for medical services [28, 88], school or work absenteeism reports [14, 15], over-the-counter and prescription medication sales [16-18], online activity monitoring [19, 20], and emergency medical system (911) or health advice calls [21, 22].

These pre-diagnostic data sources can be timelier than traditional data sources if data are collected by an automated system to minimize reporting delays. Furthermore, the population captured by novel surveillance streams may differ from the populations monitored by traditional data sources and therefore display different epidemic dynamics. For example, several studies have demonstrated that data for milder influenza-related illness can provide earlier signals for seasonal influenza outbreaks compared with measures of more severe disease [87]. In Australia, Zheng et al observed that monitoring time series of ED visits clinically diagnosed with influenza provided 3-18 days earlier warning compared with surveillance of laboratory-confirmed influenza, as the latter disproportionately represents disease in more severe or hospitalised cases [89]. Using physician billing data, Chan et al. found that outpatient clinic visits for ILI increased in frequency up to two weeks prior to ED visits for ILI and hospital admissions for P&I in Montreal, Quebec [28]. Compared with P&I mortality data, Brownstein et al reported that

ambulatory visits for acute respiratory illness in Massachusetts displayed a lead-time of approximately 4 weeks [27].

Although syndromic surveillance systems are promising, novel, non-specific data sources, systems using data not initially collected for the purpose of surveillance should be evaluated carefully. Their accuracy must be validated and their utility assessed in comparison to established and specific sources of information [81, 90]. Several attributes serve as hallmarks of the utility of a communicable disease surveillance method or system: *sensitivity* to identify both individual cases and outbreaks, *simplicity* of structure and ease of operation, *quality data* (completeness and validity), *positive predictive value* (proportion of reported cases that actually have the disease under surveillance), *representativeness* (ability to accurately describe the occurrence of disease over time, place and person), *stability* (data are reliably available when needed), *flexibility* to adapt to changing information needs or operating conditions, *acceptability* (willingness of persons and organizations to participate in the surveillance system), and *timeliness* (the early identification of trends and outbreaks) [91].

2.2.3 Data streams currently in use for influenza surveillance in Quebec and Canada

In Canada, national influenza surveillance is performed by PHAC through the Centre for Immunization and Respiratory Infectious Diseases and the National Microbiology Laboratory. PHAC publishes a weekly FluWatch report for the dissemination of influenza surveillance information [24]. The FluWatch program includes six routinely collected indicators of influenza activity:

1. **Laboratory surveillance**: sentinel laboratories report the total number of influenza tests performed and the total number of tests positive for influenza (by virus type and sub-type, when typing information available).

2. **Antigenic characterisation and antiviral resistance testing** for circulating influenza viruses.

3. **Sentinel ILI primary care consultation rates**: sentinel physicians report the total number of patients seen for any reason and the total number of patients meeting the PHAC definition for ILI for one clinic day each week.

4. **Regional influenza activity levels**: provincial and territorial representatives provide weekly assessments of regional influenza activity and the number of outbreaks of influenza or ILI in schools, hospitals and residential institutions.

5. **Severe outcomes surveillance**: pediatric and adult influenza-associated hospital admission, intensive-care unit admission and mortality data are monitored through hospital-based surveillance and provincial/territorial reporting directly to PHAC.

6. **Pharmacy surveillance**: antiviral dispensing data are provided by Rx Canada Inc. and sourced from over 3,000 stores major retail drug chains nationwide. Data provided include the number of new NI prescriptions and the total number of new prescriptions dispensed.

In addition, FluWatch assesses international influenza activity and monitors reports of cases of emerging respiratory pathogens such as influenza viruses of avian or swine origin and other respiratory viruses, e.g, the middle-eastern respiratory syndrome coronavirus.

In the province of Quebec, laboratory surveillance for respiratory viruses is performed by the *Laboratoire de santé publique du Québec* (LSPQ), the provincial reference public health laboratory that is part of the *Institut national de santé publique du Québec* (INSPQ) [92]. Aggregate weekly counts of positive test results (by influenza type and subtype, when available) and the total number of tests performed for influenza are reported by sentinel laboratories to the LSPQ. The INSPQ disseminates these results, stratified by age group and health region, through weekly updates and monthly "*Flash Grippe*" summaries [93]. Moreover, the INSPQ monitors

and reports on the following data streams in *Flash Grippe*: ED visits for ILI, pediatric hospital admissions for P&I, calls to the provincial health advice hotlines (- and *Info-Social*) for ILI, and the number of influenza outbreaks in long-term care facilities.

2.3 Evaluations of the utility of influenza antiviral prescription drug dispensing data for the surveillance of influenza activity

We reviewed the studies that assessed the utility of influenza antiviral dispensing data from community-based retail pharmacies for the surveillance of influenza activity through a comparison with an established reference indicator for influenza activity. We identified five such studies (Table 2–2).

Table 2-2. Studies assessing the utility of influenza antiviral dispensing as an influenza surveillance method.

Author, year	Study period	Location	Methodology	Antiviral dispensing data	Primary comparator	Lead-lag relation	Account for autocorr.	Yearly analysis	Primary outcome
Yoshida, 2009 [94]	2004-2006	Osaka, Japan	Correlation analysis	NI + Ad	Confirmed cases reported by sentinel physicians	N/P	N/P	N/P	*R*=0.954
Sugawara, 2012 [95]	2009-2011	Japan	Correlation analysis	NI	Confirmed cases reported by sentinel physicians	N/P	N/P	Yes	2009-10, *R*=0.992; 2010-11, *R*=0.972
Greene, 2012 [96]	2000-2010	US	Correlation analysis	NI + Ad	Laboratory surveillance	N/P	N/P	N/P	*R* ranged from 0.34-0.72, across jurisdictions
Patwhardhan, 2012 [97]	2007	US	Correlation analysis	NI + Ad	CDC ILI surveillance network	N/P	N/P	N/A	*R*=0.92
Aramini, 2013 [98]	2009 (second pandemic wave)	Ontario, Canada	Poisson regression	NI	Laboratory surveillance	Maximal significance at lag 0	N/P	N/A	Significance in a model predicting A/H1N1 lab cases: *P*-value <0.001

Ad, adamantanes

Autocorr., autocorrelation

ILI, influenza-like illness

N/A, not applicable

N/P, not performed

NI, neuraminidase inhibitors

R, Pearson correlation coefficient

US, United States

To evaluate the association between NI dispensing counts and established indicators of influenza activity, studies identified in this review estimated 1) strength of correlation, 2) timeliness, or 3) predictive value.

Correlation analysis, frequently used to measure the association between two time series [10, 25, 28, 99], was employed by four of the five studies [94-97]. Yoshida et al observed a very strong correlation (R=0.954) between antiviral dispensing (NI and adamantanes) and sentinel physician reports of laboratory-confirmed influenza in Osaka, Japan during a two-year period (2004-2006) [94]. At the national level in Japan, using a similar comparator, Sugawara et al reported that NI dispensing produced correlations of 0.992 and 0.972 during the 2009 A/H1N1 pandemic and the following season (2010-11), respectively [95]. However, there was modest variability between prefectures and the lowest regional correlation was 0.689. In the United States, using the Vaccine Safety Datalink Project database, Greene et al examined NI and adamantane prescriptions in relation to laboratory surveillance data over 10 years (2000-2010) [96]. They reported their correlation analyses stratified by 8 different medical care organizations; no pooled estimate was produced. The correlations of weekly antiviral dispensings with the proportion of tests positive for influenza showed variability, ranging from 0.34-0.72. Finally, Patwardhan et al compared antiviral drug sales (NI and adamantanes) to CDC Outpatient Influenza-like Illness Surveillance Network (ILInet) data, observing a correlation of 0.92 in 2007 [97]. They also compared antiviral dispensing to another non-traditional surveillance method, Google Flu Trends, which attempts to provide estimates of influenza activity based on Internet search data [19]. The correlation for five years' aggregate data (2007–2011) was 0.92. For each of the five years between 2007 and 2011, correlations were 0.85, 0.92, 0.91, 0.88, and 0.87 respectively. However, Google Flu Trends is not considered to be a reliable comparator as several validation studies have reported that estimates of influenza activity based on this data

source may deviate significantly from influenza patterns demonstrated by traditional surveillance systems [12, 100, 101].

Taken as a whole, antiviral dispensing data appear to be moderately to strongly correlated with established influenza surveillance methods. Unfortunately, the results of all four of the studies to date are likely biased because none accounted for the autocorrelation that is inherent to seasonal health data. Empirical correlations (estimated from the raw data) are highly prone to bias when the data are temporally autocorrelated (i.e., the value on any given day is correlated with values on previous days). Because of the autocorrelation within each individual series, the empirical correlation of two unrelated time series can be spuriously but significantly high due to chance alone [102] or due to the confounding effect of a seasonal covariate [25, 103].

The evaluation of the timeliness of a novel data source generally aims to assess if it could serve as a leading indicator for influenza activity, i.e., could provide earlier information for rapid outbreak detection and situational awareness [10, 13]. There is no well-established definition of timeliness or one preferred method for its quantification [10]. A frequently encountered method to assess timeliness of data is an extension of simple correlation analysis, the cross-correlation function (CCF), which measures the correlation between two time series that have been lagged by various units of time. Lead-time can then be defined as the lag at which the peak correlation in the CCF occurs or as the earliest lag at which a statistically significant correlation occurs [10, 25, 28]. Unfortunately, none of the reviewed studies that performed correlation analyses explored lead-lag relationships between data sources or addressed timeliness in any way. Regarding the CCF, in addition to the aforementioned reasons for which it is important to account for autocorrelation, it has also been demonstrated that, without appropriate filtering (i.e., "pre-whitening" or removing systematic patterns in the data), long-scale phenomena (over months to years) such as seasonality tend to overwhelm the CCF, obscuring the short-scale fluctuations

(over weeks) that are more relevant to the surveillance of influenza outbreaks [13].

In the context of a Poisson regression model predicting laboratory cases, Aramini et al explored the lead-lag relationship between the weekly volume of NI prescriptions and the weekly numbers of positive laboratory tests of A/H1N1 in Ontario, Canada, during the second wave of the outbreak of the 2009 pandemic [98]. They found "a statistically significant relationship between weekly influenza A(H1N1) case counts and antiviral prescriptions at the local health authority level ($p<0.001$). Statistical significance was greatest when influenza A(H1N1) cases counts were not lagged by time". No further details regarding their modeling strategy were provided; however, it appears that they did not account for autocorrelation in their analysis. In generalized linear models (such as Poisson regression), one of the fundamental assumptions is that the observations of a variable are independent and identically distributed. However, we have already established that observations in seasonal data are autocorrelated and, consequently, not independent of each other. Furthermore, the mean weekly values and their variances will vary over time with seasonality. Therefore, it is well understood that, when using generalized linear models to study outcomes related to influenza, seasonal variation and secular trends in the data must be removed [40, 104, 105].

Finally, it is notable that influenza outbreaks can show important year-to-year variability in terms of peak activity, intensity, timing, duration, disease severity, vaccine effectiveness and distribution of circulating viruses [34]. All of these variables will affect the healthcare seeking behavior of patients and the propensity of clinicians to test and treat for influenza. Therefore, if possible when assessing a novel data source, a year-by-year analysis should be performed in addition to aggregated estimates, as results may not always be consistent across influenza seasons [28]. In our review, only the Sugawara et al study of national data from Japan provided estimates stratified by study year.

In conclusion, there is little information available on the utility as a surveillance method of NI dispensing data, a currently used indicator of influenza activity in Canada. Reported empirical correlations with traditional data sources appear promising; however, valid estimates are lacking because no study to date has accounted for autocorrelation. The timeliness and predictive accuracy of NI dispensing has also yet to be properly characterized. Furthermore, little is known about the representativeness of the results from year to year. Finally, no study has used more than one established comparator to place NI dispensing in the common context of monitoring multiple data streams, such as laboratory surveillance and ED visits for ILI.

CHAPTER 3 - METHODS

3.1 Overview, study setting and study design

For this thesis, we evaluated the utility of NI dispensing in retail pharmacies as a data source for influenza surveillance by assessing three key attributes of a surveillance method: timeliness, correlation and predictive accuracy [11].

This is an ecological time-series study; we analysed only aggregate measures [106]. We compared the weekly time series of counts of NI dispensing and ED ILI visits to the weekly time series of positive influenza laboratory tests (as a reference measure of influenza circulation).

The study period covered three years, including three seasonal influenza epidemics, from July 4, 2010 to June 29, 2013, in Quebec, Canada. All data were obtained from province-wide databases. The province of Quebec is situated in the central region of Canada. It is the country's largest province, and the second most populated. In 2013, its population was estimated to be 8,155,300 inhabitants [107].

We obtained permission from the LSPQ, the INSPQ and IMS Brogan Canada, to use the data for this thesis. In addition, the Institutional Review Board at McGill University approved this project.

3.2 Sources of data

3.2.1 Provincial sentinel laboratory surveillance

The INSPQ performs year-round laboratory-based surveillance for influenza, thereby contributing data to PHAC's FluWatch program and to the World Health Organization's Global Influenza Surveillance and Response System [92, 108]. The LSPQ collects weekly information

on results of testing for influenza from each of the over 40 sentinel virology laboratories across the Province of Quebec. Weekly counts of positive tests for influenza (incident cases) and weekly counts of the total number of diagnostic tests performed for influenza (culture, antigen detection [immunofluorescence or enzyme immunoassay] RT-PCR) are reported to the LSPQ by manual entry (by a laboratory technologist or a clerk) on a web-based platform within 72 hours of the end of the previous week. Weekly percentages of positive tests can therefore be calculated. When influenza is identified, laboratories report virus type (A or B); however, less than 15% have the capacity to distinguish influenza A subtype (A/H1N1 or A/H3N2).

Although the virology laboratories are located in hospital centres, specimens may originate from patients of any age and from various clinical settings, such as community or hospital outpatient clinics, EDs, and acute care or long-term care inpatient wards. Nevertheless, based on published indications for testing, it is expected that the vast majority of results originate from hospitalised patients or from patients at risk of severe outcomes [49, 64].

The surveillance program is passive (testing is performed at treating physicians' discretion and cases are not actively sought). Laboratory participation is voluntary; however, all 18 *régions socio-sanitaires* are represented. A survey conducted in 2013 by the LSPQ found that a total of 61 laboratories perform diagnostic testing for influenza in Quebec [109]. By season, the number of hospitals providing data was: 44 in 2010-11 44 (72% of the total number of virology laboratories); 45 in 2011-12 (74%); 46 in 2012-13 (75%).

Sentinel hospitals also provide the LSPQ with a sample of each season's influenza isolates for further characterization to estimate the proportion of each influenza subtype circulating that year, their antigenic similarity to the annual vaccine strain, and their antiviral susceptibility profiles.

3.2.2 NI dispensing

IMS Brogan Canada, Inc., a commercial provider of information services and technology for the healthcare industry, provided, free of charge, aggregate weekly counts of oseltamivir and zanamivir prescriptions dispensed in Quebec. These data were obtained from their proprietary Canadian Weekly CompuScript drug use database. The Weekly CompuScript database collects prescription volume dispensed by Canadian retail (i.e., non-hospital) pharmacies to outpatients on a weekly aggregate basis. Information was available for approximately 60% of pharmacies in the province of Quebec. The number of pharmacies providing weekly data ranged from 1,071 to 1,080 over the study period. According to IMS Brogan, for Weekly Compuscript, each supplier and their data are verified on a weekly basis to ensure that they are within the standards set for quality control (based on consistency with previous weeks). If variances are noticed, then they are examined to determine if they are explainable or require further verification with the supplier.

3.2.3 ED visits for ILI

Aggregate weekly counts of visits for ILI to a Quebec ED were obtained from the Daily Report on the Situation in Emergency Departments and Hospitals (*Relevé quotidien de la situation à l'urgence et au centre hospitalier* – RQSUCH) database. Each acute care hospital in Quebec is required to report the daily number of patients registering to the ED who present with the chief complaint of ILI, defined as fever and cough, as well as the total number of patients registering to the ED for any reason. These counts are entered manually by nurses or clerks in each hospital and are transmitted daily to local public health authorities and the Ministry of Health for healthcare utilization surveillance. However, the manner in which the ILI definition is applied, by whom (healthcare professional or support staff), and from which source, is left to the discretion of each ED. This system has been in operation since May 2008, with 100% coverage

of Quebec EDs since May 2009. During our study, data completeness (assessed daily, for each hospital) was 99.83%.

3.3 Choice of reference indicator

We used the laboratory counts as our reference indicator for the level of influenza circulation because laboratory testing is the most specific method for the detection of influenza. Laboratory surveillance is commonly used as both a data source and an outcome measure in influenza forecasting [82].

3.4 Outcomes

1) *Timeliness.* There is no standard metric to evaluate the timeliness of a data source for surveillance purposes [10]. For this thesis, we assessed timeliness, or lead-time, by exploring lead/lag relationships with our reference indicator, the influenza laboratory surveillance data stream, by cross-correlation analysis and by fitting Box-Jenkins transfer function models (see Data analysis section). A data source was considered timely (i.e., a leading indicator) if it demonstrated statistically significant cross-correlations at lags ≥ 0 i.e., when it was lagged relative to the reference series so as to precede it.

2) *Correlation*. The Pearson product-moment correlation coefficient (R) was used to measure correlation during cross-correlation analysis [99]. R is a measure of the linear correlation (dependence) between two variables, giving a value between +1 and −1 inclusive, where 1 is total positive correlation, 0 is no correlation, and −1 is total negative correlation.

3) *Predictive accuracy* was evaluated by fitting multivariable transfer function models to the laboratory-confirmed influenza time series. We defined the best model as the model with the lowest corrected Akaike's Information Criterion (AICc); an input time series was considered to be a useful predictor if its inclusion lowered the AICc.

3.5 Data analysis

Analyses were performed using R version 2.14 (www.r-project.org) and SAS version 9.3 (SAS Institute, Inc., Cary NC). A two-sided *P* value of <0.05 was considered statistically significant.

3.5.1 Removal of autocorrelation through ARIMA modeling

As discussed in section 2.3, to obtain valid estimates of timeliness and correlation when comparing time series, it is imperative to account for seasonal variation and long-term trends, thereby removing the autocorrelation structure in the data and controlling for the "non-independence" of events.

To overcome the possible bias in estimates due to autocorrelation, one can transform (or filter) the series under consideration by fitting an appropriate statistical model, so that the residuals are a series of independent, identically distributed random observations. This process was termed "pre-whitening" by Box and Jenkins [26].

Different statistical models can be used for pre-whitening. For instance, seasonality and trend can be modeled within a generalized linear model regression approach (e.g., Poisson, quasi-Poisson, or negative binomial model) by incorporating polynomial terms, autoregressive terms or smoothing splines. However, classical regression may sometimes be insufficient for explaining

all of the dynamics of a time series [110]. The autoregressive integrated moving average (ARIMA) approach, first popularized by Box and Jenkins, was specifically designed to address the issue of modeling the stochastic dependence of consecutive data [102]. An ARIMA model predicts a current value Y_t in a response (output) series by a linear combination of the p previous observations $Y_{t-1}, \ldots, Y_{t-p}$, a linear combination of the q previous errors or "random shocks" $a_{t-1}, \ldots a_{t-q}$ and a constant term. Here, p and q represent the order of the autoregressive (AR) and moving average (MA) components, respectively. Mathematically, this can be generalized to:

$$Y_t = \text{Constant} + \varphi_1(Y_{t-1}) + \ldots + \varphi_p(Y_{t-p}) + a_t - \theta_1(a_{t-1}) - \ldots - \theta_q(a_{t-q})$$

where $\{\ldots Y_{t-1}, Y_t, Y_{t+1}, \ldots\}$ is a series of observations at equally spaced time intervals, $\{\ldots a_{t-1}, a_t, a_{t+1}, \ldots\}$ is a white noise series of independent and identically distributed random variables whose distribution is approximately normal with mean zero and variance σ^2, and $\varphi_1,\ldots,\varphi_p$ and $\theta_1,\ldots,\theta_q$ are parameters to be estimated from the data.

Box-Jenkins modeling should only be applied to time series that is stationary, i.e., the series has a constant mean and variance over time. If the variance is related to the mean, then a variance-stabilizing transformation has to be applied, such as log-transformation. Stabilizing the mean can be achieved by differencing the data: subtracting the value at each time point by the value at the previous time point to detrend the series. The parameter d is used to refer to the number of differences, resulting in an ARIMA (p,d,q) model.

The standard approach to ARIMA modeling has three steps, which may be iterative.

1) *Model identification*, wherein the orders of the AR (p) and MA (q) terms are chosen. The simplest model that fits the data should be selected.

2) *Estimation* of the model parameters. Values of the AR and MA coefficients that provide the best fit are determined using computational algorithms such maximum likelihood estimation.

3) *Diagnostic checking* of the models is then conducted by examining their residuals. The residuals should be independent of each other and resemble a white noise process; that is, no significant autocorrelations between residuals should be detected and they should display normality. If the model is inadequate, steps 1 to 3 are repeated to identify another potential model.

When discriminating between adequate models, the AIC can used to assess the goodness of fit of each model [110].

For any statistical model, the AIC can be calculated as:

$$AIC = 2k - 2\ln(L)$$

where k is the number of parameters in the model, and L is the maximized value of the likelihood function for the model.

However, it has been suggested that, in the context of ARIMA modeling, the AICc is a preferred criterion, because it has a greater penalty for extra parameters and thereby reduces the risk of over-fitting, i.e. of tailoring the fit too closely to the particular numbers observed. [111].

$$AICc = AIC + (2k(k+1))/(n-k-1)$$

where n denotes the sample size (i.e., the number of time periods).

3.5.2 Cross-correlation analysis

The association between two time series can be quantified by the CCF, which determines the correlation (Pearson's R) between the two series over a range of time lags [10, 25, 28]. As described above, we differenced each time series to render it stationary and then applied ARIMA

models to remove autocorrelation [26, 102]. We pre-whitened the NI dispensing data by using the standard approach of identification, estimation and diagnostic checking of the residuals. Models were fit using maximum likelihood estimation. The AICc guided selection among adequate models, taking into account both goodness of fit and parsimony. Then, we applied the final ARIMA model to the laboratory surveillance reference time series. The CCF was subsequently computed between the residuals of both pre-whitened series. We identified the lags at which the maximum and earliest statistically significant correlations occurred. We repeated this process for the ED ILI data.

Because seasonal influenza epidemics can show important year-to-year variability in terms of their intensity, duration, timing and distribution of circulating viruses [34], we also performed the same CCF analyses separately for each of the three outbreaks to evaluate if our results would be consistent across influenza seasons.

3.5.3 Transfer function models

Transfer functions estimate the dynamic linear relation between an input and an output series. A transfer function model is therefore an ARIMA model that predicts a value in the output time series as a linear combination of its own past values, past errors, *and current and past values of other time series*.

We assessed the value of NI dispensing data for the prediction of influenza activity using transfer function models [26, 102, 105]. First, we differenced, log-transformed and pre-whitened each series as previously described. Then, bivariate and multivariate Box-Jenkins transfer function models were developed to describe the relationship between the pre-whitened output series (laboratory surveillance data) and each of the pre-whitened input series (NI dispensing and ED ILI visits). Only positive lags with a significant cross-correlation were considered. Log-

transformations of the series were considered when model assumptions were not easily met by models constructed from the untransformed series. The AICc of the ARIMA model of the output series was used as a yardstick when comparing the informative potential of different transfer function models. An input series was considered to be a useful predictor if its inclusion lowered the AICc.

3.5.4 Sensitivity analyses

To measure influenza activity, surveillance systems frequently monitor the *proportion* of the number of positive influenza tests over the total number of influenza tests performed [112, 113]. Applying a denominator is thought to account for the fact that health care seeking behavior and physician propensity to test may be influenced by factors unrelated to the level of influenza circulation. For instance, during holidays or during severe weather, patients may be less likely to consult for influenza infection; proportions may be insensitive to this effect and thus, compared to counts, better reflect the prevalence of circulating virus. Holidays (and the accompanying delay in test turnaround time), the presence of other circulating viruses, and results of recent influenza surveillance reports may also affect a physician's likelihood to test for influenza. However, we did not use proportions in our primary analyses; we chose to use laboratory counts and ILI counts. The primary reason for this choice was that we did not have access to an appropriate denominator for NI dispensing (e.g., total number of all drugs dispensed). Using a denominator to account for time-varying healthcare utilization phenomena in our reference series but not in our NI data might bias our results. The use of proportions to measure influenza activity also has the disadvantage that the denominator is influenced by the presence or absence of other cyclical diseases that are not related to influenza. For instance, when monitoring proportions of ED visits for ILI, if there is a gastroenteritis outbreak during the winter, then the total number of

ED visits will be inflated and will bias towards a relative underestimation of influenza activity. Similarly, during peak RSV season, physicians will perform more influenza tests (RSV may present like influenza) thereby lowering the proportion of positive tests for influenza, even if the prevalence of influenza is in reality unchanged.

Nevertheless, to assess if our results were robust to the use of proportions as opposed to counts, we reran our CCF and prediction model analyses using the weekly proportion of positive influenza tests (weekly count of positive laboratory influenza tests / weekly count of influenza tests performed) as the reference time series and the weekly proportion of ILI ED visits (weekly count of ILI ED visits / weekly count of total ED visits) instead of ILI ED counts.

CHAPTER 4 – THE ACCURACY AND TIMELINESS OF NEURAMINIDASE INHIBITOR DISPENSING DATA FOR PREDICTING LABORATORY-CONFIRMED INFLUENZA

4.1. Preamble

There is much interest in identifying syndromic surveillance data sources that could serve as a leading indicator for influenza activity, i.e., could provide earlier information for rapid outbreak detection and near real-time situational awareness. Pharmacy surveillance of outpatient prescription sales of the neuraminidase inhibitor (NI) class of influenza antivirals is currently being used as indicator of influenza activity by the Public Health Agency of Canada. However, to date, studies evaluating the utility of antiviral dispensing have not controlled for autocorrelation. Therefore, previous estimates of the timeliness and correlation of antiviral dispensing in relation to established surveillance methods are likely to be biased.

In this manuscript, we evaluated the timeliness, correlation, and predictive accuracy of community pharmacy NI dispensing in relation to laboratory-confirmed influenza activity in Quebec, Canada, during 2010-2013. We did so by first applying ARIMA modelling to the data to control for autocorrelation, and then computing cross-correlations of the residuals of both series across various lags to assess timeliness and correlation. We also compared these results to those obtained with another commonly used influenza indicator, visits to hospital emergency departments for influenza-like illness (ILI). Finally, we assessed the predictive accuracy of both NI dispensing and ILI visits in modelling laboratory counts of influenza.

This manuscript has been formatted for submission to *Clinical Infectious Diseases*.

4.2. MANUSCRIPT

The accuracy and timeliness of neuraminidase inhibitor dispensing data for predicting laboratory-confirmed influenza

Jesse Papenburg MD [1, 2, 3], Katia M. Charland, PhD [3], Gaston De Serres MD PhD [4] , David L. Buckeridge MD PhD [1, 3]

[1] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada.

[2] Division of Pediatric Infectious Diseases, Departments of Pediatrics and Microbiology, Montreal Children's Hospital, Montreal, QC, Canada.

[3] Clinical and Health Informatics Research Group, McGill University, Montreal, QC, Canada.

[4] Institut national de santé publique du Québec, Quebec City, QC, Canada

MeSH keywords : influenza; public health surveillance; antiviral drugs; community pharmacies; multivariate analysis

Running title: Antiviral data for influenza surveillance

Corresponding author: Dr. David Buckeridge

Clinical and Health Informatics Research Group, McGill University

1140 Pine Ave West, Montreal, Quebec, H3A 1A3, Canada

Tel.: 514-934-1934 ext. 32991; Fax: 514-843-1551

E-mail: david.buckeridge@mcgill.ca

Alternate corresponding author: Dr. Jesse Papenburg

Montreal Children's Hospital

2300 Tupper St., Room C-1246, Montreal, Quebec, Canada, H3H 1P3

Tel.: (514) 412-4485; Fax: (514) 412-4494

E-mail: jesse.papenburg@mail.mcgill.ca

Brief summary: The correlation, timeliness and predictive accuracy of NI dispensing data in relation to laboratory-confirmed influenza activity suggest that this readily available data stream could be used as a leading indicator for outbreak detection and situational awareness during seasonal epidemics.

ABSTRACT

Background: Neuraminidase inhibitor (NI) dispensing in community pharmacies has emerged as a possible automated data source for influenza surveillance. However, little is known about the utility of these data for influenza surveillance. We aimed to evaluate the timeliness, correlation, and predictive accuracy of community pharmacy NI dispensing in relation to laboratory-confirmed influenza activity in Quebec, Canada, during 2010-2013. Our secondary objective was to compare these characteristics to those of surveillance for influenza-like illness (ILI) in emergency departments (ED), a commonly used source of surveillance data.

Methods: Provincial weekly counts of positive influenza laboratory tests were used as a reference measure for the level of influenza circulation. We applied ARIMA models to account for serial correlation. We computed cross-correlations to measure the strengths of association and lead-lag-relationships of NI dispensing and ILI ED visits with our reference indicator. Finally, we evaluated the predictive value of NI dispensing and ILI ED visits in fitting an ARIMA model for laboratory–confirmed influenza.

Results: NI dispensing was significantly correlated (R=0.68) with influenza activity at lag 0 and the earliest statistically significant correlation occurred with a lead-time of 1 week. The maximal correlation of ILI ED visits was not as strong (R=0.50), but peaked one week earlier. Both NI dispensing and ILI ED visits were significant predictor variables in a multivariable model of laboratory-confirmed cases; the predictive potential was maximal with NI counts lagged to precede laboratory surveillance by two weeks.

Conclusions: NI dispensing data provides timely and valuable information for influenza surveillance.

Influenza surveillance is of public health importance. One of its primary goals is to provide public health officials and clinicians with early detection and situational awareness of influenza activity by determining the timing, location and degree of influenza circulation and associated diseases [1, 2]. Traditional data sources for influenza surveillance such as monitoring of positive influenza laboratory tests and medically attended influenza-related illnesses are typically associated with a reporting delay of as much as 1-2 weeks [3]. Consequently, public health researchers have sought novel electronic data sources that could provide timely information for rapid influenza outbreak detection [4, 5].

It is estimated that there are over 2 million clinic visits per year for influenza in the United States [6]. Of these, approximately 20% receive a prescription for antiviral treatment. In Canada, the neuraminidase inhibitor (NI) class of medications (oseltamivir and zanamavir) is the only recommended empiric therapy for influenza since 2006 because of widespread resistance to adamantanes [7]. Monitoring NI dispensings from community pharmacy drug sales databases has recently emerged as a possible automated source of timely information regarding influenza [8].

Influenza activity in the outpatient setting tends to occur before an increased incidence of more severe disease [9-11]. Consequently, we hypothesized that changes in the weekly volume of outpatient antiviral prescriptions might precede changes in weekly counts of positive influenza tests and that NI dispensing could serve as an early indicator of epidemic influenza activity. Therefore, we aimed to evaluate retail pharmacy NI dispensing data as a novel automated influenza surveillance tool in Quebec, Canada, by assessing its timeliness, correlation, and predictive accuracy in relation to laboratory-confirmed influenza activity. Our secondary objective was to compare the above characteristics of the NI dispensing data to those of an established influenza surveillance data source, emergency department (ED) visits for influenza-like illness (ILI).

METHODS

Overview and study design

In this ecological study, we compared the weekly time series of counts of NI dispensing and ED ILI visits to the weekly time series of positive influenza laboratory tests (as a reference measure of influenza circulation) over three years, for the period of July 4, 2010 to June 29, 2013, in the province of Quebec, Canada. We evaluated three key performance characteristics of the NI dispensing and ED ILI data streams for influenza surveillance: timeliness, correlation and predictive accuracy [12].

Ethics approval was granted by the McGill University Faculty of Medicine Institutional Review Board

Outcomes

Timeliness, or lead-time, was assessed by exploring lead/lag relationships with the influenza laboratory surveillance data stream by cross-correlation function (CCF) analysis and by fitting Box-Jenkins transfer function models (see Data analysis section). A data source was considered timely if it demonstrated statistically significant cross-correlations at lags ≥ 0 i.e., when it was lagged relative to the reference series so as to precede it. Strength of correlation (Pearson's *R*) [13] was measured by the greatest significant cross-correlation in the CCF. Finally, predictive accuracy was evaluated by fitting multivariable transfer function models to the laboratory-confirmed influenza time series. We defined the best model as the model with the lowest corrected Akaike Information Criterion (AICc).

Sources of data

Provincial sentinel laboratory surveillance

The *Institut national de santé publique du Québec* performs laboratory-based surveillance for influenza year-round [14]. Aggregate weekly counts of laboratory-confirmed influenza A or B detection (positive culture, antigen detection [immunofluorescence or enzyme immunoassay] or polymerase chain reaction results) and the number of tests performed in participating hospitals are collated and disseminated publicly. The surveillance program is passive (testing is performed at treating physicians' discretion; cases are not actively sought) and laboratory participation is voluntary. However, all 18 Quebec health regions are represented and the number of hospitals providing data was stable throughout the study period (44 in 2010-11; 45 in 2011-12; 46 in 2012-13). A sample of each season's influenza isolates are characterised to estimate the proportion of each influenza subtype circulating that year and their antigenic similarity to the annual vaccine strain.

We used the laboratory counts as our reference indicator for the level of influenza circulation because, due to the inaccuracy of clinical diagnostic criteria, laboratory testing is required to confirm influenza infection [15, 16]. Virological surveillance is therefore commonly used as both a data source and an outcome measure in influenza forecasting [17].

NI dispensing

Aggregate weekly counts of NI prescriptions dispensed to outpatients in Quebec retail (non-hospital) pharmacies were obtained from IMS Brogan's Canadian Weekly CompuScript proprietary drug use database. The number of pharmacies providing weekly data ranged from 1,071 to 1,080 (>60% of all Quebec retail pharmacies) over the study period.

ED visits for ILI

Aggregate weekly counts of visits for ILI to a Quebec ED were obtained from the Daily Report on the Situation in Emergency Departments and Hospitals (*Relevé quotidien de la situation à l'urgence et au centre hospitalier* – RQSUCH) database. Each acute care hospital in Quebec is required to report the daily number of patients registering to the ED who present with the chief complaint of ILI, defined as fever and cough, as well as the total number of patients registering to the ED for any reason. These counts are entered manually by nurses or clerks in each hospital and are transmitted daily to local public health authorities and the Ministry of Health for healthcare utilisation surveillance. During our study, data completeness was 99.83%.

Data analysis

Analyses were performed using R version 2.14 (www.r-project.org) and SAS version 9.3 (SAS Institute, Inc., Cary NC).

CCF: analysis of timeliness and correlation

The CCF is a frequently used measure of the association between two time series; it quantifies their correlation over a range of time lags [4, 10, 18]. However, the empirical CCF (estimated from the raw data) is highly prone to bias, especially when one or both data series exhibit temporal autocorrelation (i.e., the value on any given day is correlated with values on previous days). Because of the autocorrelation within each individual series, the empirical CCF of two unrelated time series can display significantly high but spurious correlations due to chance alone [19] or due to the confounding effect of a seasonal covariate [18, 20]. Long-scale phenomena (over months to years), such as the seasonality within a series, tend to overwhelm the CCF, obscuring the short-scale fluctuations (over weeks) that are more relevant to the

surveillance of influenza outbreaks [21]. In order to overcome these problems, series are filtered (i.e., "pre-whitened") to remove the effects of systematic patterns and long-term trends, such as seasonality, on the empirical CCF [21]. To pre-whiten the series we fit an autoregressive integrated moving average (ARIMA) model to the NI dispensing data by using the standard approach of identification, estimation and diagnostic checking of the residuals, thereby transforming the NI series into a white noise series with no autocorrelation [19, 22]. In the case of several models passing all diagnostic tests, the corrected Akaike's Information Criterion (AICc) guided model selection [23]. The best fitting ARIMA model was then applied to the laboratory surveillance reference time series, and the CCF of the two residual series was estimated. From this CCF, we identified the lags at which the maximum and earliest statistically significant correlations occurred. We did the same for the ED ILI series to estimate the CCF the ED ILI series and the laboratory-confirmed series and then produced the CCF of the two residual series. A two-sided *P* value of <0.05 was considered statistically significant. Because influenza outbreaks can show important year-to-year variability in terms of their intensity, duration, timing and distribution of circulating viruses [24], we also performed the same CCF analyses separately for each of the three outbreaks to evaluate if our results would be consistent across influenza seasons.

Transfer function models.

We assessed the value of NI dispensing data for the prediction of influenza activity using ARIMA modelling [19, 22, 25]. First, we pre-whitened each series as previously described. Then, bivariate and multivariate Box-Jenkins transfer function models were developed to describe the relationship between the pre-whitened output series (laboratory surveillance data) and each of the pre-whitened input series (NI dispensing and ED ILI visits). Only positive lags with a significant

cross-correlation were considered. Log-transformations of the series were considered when model assumptions were not easily met by models constructed from the untransformed series. The AICc of the ARIMA model of the output series was used as a yardstick when comparing the informative potential of different transfer function models. An input series was considered to be a useful predictor if its inclusion lowered the AICc.

Sensitivity analyses

Health care seeking behaviour and physician propensity to perform laboratory testing for influenza vary over time. Applying a denominator to laboratory and ILI surveillance counts to account for such phenomena may modify estimates of influenza prevalence [1, 2]. Therefore, to assess if our results were robust to the use of a proportion as opposed to counts, we reran our CCF and prediction model analyses using the weekly proportion of positive influenza tests (weekly count of positive laboratory influenza tests / weekly count of influenza tests performed) as the reference time series and the weekly proportion of ILI ED visits (weekly count of ILI ED visits / weekly count of total ED visits) instead of ILI ED counts. No denominator for NI dispensing was available in our dataset.

RESULTS

NI dispensing

There were 21,066 NI prescriptions dispensed during the study period (5,550 in 2010-11; 3,347 in 2011-12; 12,169 in 2012-13). Of these dispensings, 20,999 (99.7%) were for oseltamivir.

Description of the 2010-11, 2011-12 and 2012-13 influenza seasons

The time series of weekly laboratory confirmed influenza cases, NI retail pharmacy dispensing and ED ILI visits for the period of July 4, 2010 to June 29, 2013 are presented in Figure 4-1. The influenza types and sub-types circulating in Quebec during each season of the study period are described in Table 4-1. The 2010-11 and 2012-13 influenza seasons were both characterised by a predominance of A/H3N2 (87% and 78% of strains, respectively). However, the intensity of the 2012-13 epidemic was remarkable, with peak weekly counts that were more than twice as high as in 2010-11, and four times greater than during the relatively mild 2011-12 season. In 2011-12, the outbreak was briefer and peaked later (March 2013), with concomitant circulation of both influenza A and B in roughly equal proportions (45% and 55%, respectively).

Description of the NI and ILI series in relation to the influenza series

Visually, the NI series closely tracked the laboratory influenza series (Figure 1). For the ILI ED series, although its seasonality was similar to that of the laboratory cases, there was a considerable volume of ILI presenting to Quebec EDs year-round, even when little or no influenza was circulating in the province. Moreover, in 2011-12, the peak in weekly ILI ED visits occurred in early January, before the seasonal influenza epidemic had even begun.

CCF: correlation and timeliness

In the overall analysis, the CCFs for the NI series with the laboratory series demonstrated that NI dispensing temporally coincided with (maximal correlation at lag 0) and was strongly correlated with (correlation of 0.68) laboratory-confirmed influenza activity (Table 4-2). Assessing the CCF of these two series separately for each of the three seasons, these observations were consistent over the entire study, with maximal correlations of ≥ 0.5 occurring at lag 0 and

the earliest significant correlation at lag 1 in each individual season. The magnitude of the peak correlation with influenza activity was stronger for the NI series than the ILI series overall (0.68 vs. 0.50) and in two of the three seasons (0.50 vs. 0.67 in 2010-11; 0.54 vs. 0.33 in 2011-12; 0.73 vs. 0.61 in 2012-13). The timeliness of the ILI series showed mild variability in the year-to-year analysis, with peak correlations and earliest significant correlations occurring at either lags 0 or 1.

Prediction models

Unlike the estimation of the CCF, for the transfer function models the final results are based on an analysis of log-transformed series since model residuals of the transformed series more easily passed diagnostic tests for model assumptions. Using NI dispensing or ILI ED visits as input series in separate single-input Box-Jenkins transfer function models improved the fit of a predictive model for weekly counts of laboratory-confirmed influenza cases (Table 4-3). In a multivariable model, both series were significant predictors of influenza counts. Including NI dispensing data at a lag of 2 weeks in this model optimised fit.

Sensitivity analyses

Repeating the analyses using the proportion of positive influenza tests as the reference time series and the proportion of ILI ED visits instead of ILI ED counts had little effect on the magnitude and direction of the associations that we observed.

DISCUSSION

We found that NI dispensing from retail pharmacies was timely and strongly correlated with laboratory-confirmed influenza activity during the same week over three non-pandemic seasons. This association was observed after filtering to correct for autocorrelation, such as

seasonality, and is therefore a feature of the short-scale relationship between the two data streams [21]. NI dispensings were also a significant predictor of laboratory-confirmed influenza activity in a multivariable model. The model's predictive potential was maximal when the log-transformed NI time series was lagged to precede the log-transformed laboratory surveillance data by two weeks.

Traditionally, systems for monitoring influenza activity have relied on reports from diagnostic virology laboratories as their primary source of information. Such laboratory surveillance data are highly specific: all cases reported are confirmed influenza infections. However, in Quebec, as in many jurisdictions [17, 26], sentinel laboratories must first manually submit weekly data on a web portal and there is typically a delay of 1-2 weeks before the results are published. Because all prescriptions in Quebec are subject to electronic adjudication at the time of dispensing, the monitoring of NI sales represents a potentially feasible, less laborious, inexpensive, and automated surveillance method.

Previous studies in Quebec and elsewhere have shown that ambulatory care or ED-based syndromic surveillance data for acute respiratory illness or ILI can provide earlier signals for seasonal influenza outbreaks compared with measures of more severe disease, such as hospitalizations or mortality due to pneumonia and influenza [9-11]. Because community pharmacy NI dispensing represents milder infections treated as outpatients, we hypothesized that the NI data would lead the laboratory surveillance series, as the latter is primarily representative of patients hospitalised with severe disease [27]. Conversely, if clinicians' knowledge of current levels of influenza circulation, based on laboratory surveillance reports, very strongly influences antiviral treatment, NI dispensing would lag behind this reference indicator. In our CCF analysis, however, the maximal cross-correlation between the NI and laboratory surveillance data was at lag zero and we also observed a significant cross-correlation at a lead-time of one week.

Moreover, when included in a regression model, the greatest predictive value of the NI series was with a lag of two weeks. Taken together, our observations suggest that, in addition to being a promptly available data source, NI prescriptions may also serve as an early indicator of epidemic influenza activity, especially when used for forecasting.

Earlier studies have suggested that prescription drug dispensing might offer timely information regarding influenza activity, without necessarily quantifying a lead-lag relationship to an established reference time series. Using a space-time permutation scan statistic, Greene et al compared the performance of 10 types of electronic clinical data, including antiviral dispensing (NIs and adamantanes), for the detection of clusters of illness related to influenza (though not necessarily laboratory-confirmed) during the 2007-08 influenza season in Northern California. [28]. Antiviral dispensing provided the earliest signal for one of the clusters, detected two of the four events, and produced no false alarms. During the second wave of the 2009 pandemic in Ontario, Canada, Aramini et al found that A/H1N1 counts were associated with NI prescriptions in a Poisson regression analysis; statistical significance was greatest when the series were not lagged by time ($p < 0.001$) [31].

To date, assessments of the correlation between antiviral prescribing and influenza activity, including the Ontario study, have not accounted for autocorrelation and seasonality. Therefore, these estimates are likely to be biased. Furthermore, in contrast to our study, none of the prior work on NI dispensing compared it to more than one traditional data source, which is necessary to understand its usefulness in context with currently used methods. In Japan, local and national-level dispensing data demonstrated empirical correlations (R) of > 0.95 with ILI sentinel surveillance data [26, 30]. In the national study, correlations during the 2009 A/H1N1 pandemic and the following season (2010-11) were very similar ($R = 0.992$ and 0.972, respectively) [26]. However, there was modest variability between prefectures and the lowest regional correlation

was 0.689. In the United States Vaccine Safety Datalink Project, differences were observed between 8 different medical care organizations. The empirical correlations of weekly antiviral dispensings with the proportion of tests positive for influenza ranged from 0.34-0.72 [8].

As with all forms of syndromic surveillance, antiviral dispensings cannot be as specific an indicator of influenza activity as laboratory data. Several factors may contribute to "false-positives", i.e., occurrences of NI dispensing that do not represent an incident case of influenza infection. First, while we expect that the majority of prescriptions were for treatment of acute illness, our data did not allow us to assess indication. Prophylaxis has been estimated to be the indication for <10% of antiviral dispensing in the 2000-2010 Vaccine Safety Datalink Project study [8]. That proportion is probably even smaller in our Quebec data; since the 2009 A/H1N1 pandemic, early treatment is preferred over prophylaxis due to concerns regarding emergence of drug resistance [7, 32]. Although personal stockpiling of antivirals is discouraged [33], evidence of such activity has been reported when spikes in NI sales coincided with media coverage of highly pathogenic H5N1 influenza, but not with other markers of influenza activity, such as laboratory and ILI surveillance [34]. Significant amounts of stockpiling might therefore trigger a false alarm for an influenza outbreak. However, if the monitoring of NI dispensing is performed as part of a multistream surveillance program, it offers the opportunity for public health officials to recognize inappropriate prescribing and intervene to reduce the risk of lack of availability of treatment for those that need it most [33, 35].

Outpatient antiviral treatment of influenza is rarely based on laboratory testing, in large part because test results are not available during the patient encounter. It has been estimated that only 3-6% of outpatients treated with antivirals in the United States were tested for influenza [8]. Since NIs are therefore being prescribed empirically based on a clinical syndrome [7], it is intriguing that our NI dispensing series was clearly more specific than our ED ILI data and that

trends in NI dispensing and laboratory surveillance data were so closely associated, even after correcting for seasonality. We believe that prescribers' prior knowledge of circulating levels of influenza contributes to, but cannot fully account for this phenomenon. Laboratory surveillance results take >1 week before being published and our NI data coincided with or led laboratory surveillance by 1-2 weeks. Fiejté et al reported that, during the second wave of the 2009 A/H1N1 pandemic, in Utrecht, Netherlands, among patients prescribed oseltamivir in the community setting, those in whom the decision to treat was in accordance with national guidelines more frequently started their course of therapy compared to those in whom the treatment decision was deemed inappropriate (97.4% vs. 55.9%, P <.001). Thus, it appears that patients' behaviour after the medical visit affects NI dispensing counts and that those with more severe symptoms or those more likely to be infected with influenza may also be more likely to fill their prescription.

Among the limitations of our study, we note that our data did not allow for age-stratified analyses. Patient age influences influenza transmission dynamics and, consequently, data timeliness. Studies using laboratory-based [36] and syndromic surveillance [9, 10] have demonstrated that data from children offer the earliest lead-times. Identifying the age groups that provide the timeliest data might further improve the utility of monitoring antiviral dispensings. We were also unable to determine if NI data timeliness and correlation with influenza activity are consistent geographically across Quebec. Therefore, despite the fact that our data were collected province-wide, results may not be valid for all 18 health regions. Another caveat is that NI dispensing data may not have the same performance characteristics during a pandemic period when disease severity, age distribution, timing within the calendar year, media coverage, medical-seeking behaviour, and physician propensity to test and treat can differ greatly from seasonal influenza. Furthermore, indications for NI treatment and NI prescription rates vary significantly across countries [37]. Consequently, our results may not be applicable in

jurisdictions without similar access to prescription drugs. Finally, estimating true influenza incidence from laboratory or ILI surveillance remains a challenge [1, 25, 38]. Therefore, any reference indicator used to assess NI dispensing will be imperfect and results may vary based on the choice of comparator.

We took several measures to ensure the validity of our results. Chiefly, we pre-whitened the time series prior to estimating correlations between them, removing autocorrelation and thereby reducing the possibility of biased measures of association [18]. Also, by filtering the effect of seasonality, we focused on the short scale features of the lead-lag relationships to produce estimates applicable to the detection of rapidly evolving influenza outbreaks [21]. Furthermore, the choice of ARIMA methodology for building prediction models is well suited for shorter-term forecasting, as it places greater weight upon recent past values [22]. Finally, year-by-year analyses demonstrated that our estimates were stable to variations in epidemic season timing, duration, overall severity, peak intensity and antigenic characterisation of the predominant strains.

In summary, the correlation, timeliness and predictive ability of NI dispensing data in relation to laboratory-confirmed influenza activity that we report here suggest that this readily available data stream could act as a leading indicator for outbreak detection. Monitoring NI dispensing, especially in parallel to traditional sources of surveillance data, should increase public health practitioners' situational awareness of influenza activity thereby facilitating timely interventions and resource management.

Acknowledgements

We thank Patricia Fontela, MD, PhD for thoughtful review of this manuscript; Hugues Charest, PhD and Michel Couillard, PhD for kindly providing the laboratory surveillance data; Monique Douville-Fradet, MD, MHSc for kindly providing RQSUCH data; IMS brogan for generously providing the antiviral dispensing data.

Potential conflicts of interest:

GDS has received research funds from GlaxoSmithKline and Sanofi-Pasteur for unrelated studies.

REFERENCES

1. Thompson WW, Comanor L, Shay DK. Epidemiology of seasonal influenza: use of surveillance data and statistical models to estimate the burden of disease. J Infect Dis **2006**; 194 Suppl 2: S82-91.
2. World Health Organisation. WHO Global Epidemiological Surveillance Standards for Influenza. http://www.who.int/influenza/resources/documents/influenza_surveillance_manual/en/index.html. Last accessed: 20 July, 2014. **2012**.
3. Nsoesie EO, Brownstein JS, Ramakrishnan N, Marathe MV. A systematic review of studies on forecasting the dynamics of influenza outbreaks. Influenza Other Respir Viruses **2014**; 8(3): 309-16.
4. Dailey L, Watkins RE, Plant AJ. Timeliness of data sources used for influenza surveillance. J Am Med Inform Assoc **2007**; 14(5): 626-31.
5. Shaman J, Karspeck A, Yang W, Tamerius J, Lipsitch M. Real-time influenza forecasts during the 2012-2013 season. Nat Commun **2013**; 4: 2837.
6. Hsieh YH, Chen KF, Gaydos CA, Rothman RE, Kelen GD. Antiviral prescriptions to U.S. ambulatory care visits with a diagnosis of influenza before and after high level of adamantane resistance 2005-06 season. PLoS One **2010**; 5(1): e8945.
7. Aoki FY, Allen UD, Stiver HG, Evans GA. The use of antiviral drugs for influenza: Guidance for practitioners 2012/2013. Can J Infect Dis Med Microbiol **2012**; 23(4): e79-e92.

8. Greene SK, Shay DK, Yin R, et al. Patterns in influenza antiviral medication use before and during the 2009 H1N1 pandemic, Vaccine Safety Datalink Project, 2000-2010. Influenza Other Respir Viruses **2012**; 6(6): e143-51.

9. Brownstein JS, Kleinman KP, Mandl KD. Identifying pediatric age groups for influenza vaccination using a real-time regional surveillance system. Am J Epidemiol **2005**; 162(7): 686-93.

10. Chan EH, Tamblyn R, Charland KM, Buckeridge DL. Outpatient physician billing data for age and setting specific syndromic surveillance of influenza-like illnesses. J Biomed Inform **2011**; 44(2): 221-8.

11. Miller B, Kassenborg H, Dunsmuir W, et al. Syndromic surveillance for influenzalike illness in ambulatory care network. Emerg Infect Dis **2004**; 10(10): 1806-11.

12. Lombardo JS, Buckeridge DL, eds. Disease surveillance: a public health informatics approach. 1 ed. Hoboken, NJ: John Wiley & Sons, Inc. , **2007**.

13. Taylor R. Interpretation of the correlation coefficient J Diagn Med Sonogr **1990**; 1(1): 35-9.

14. Institut national de santé publique. Surveillance de l'influenza et des infections respiratoires virales. http://www.inspq.qc.ca/influenza/surveillance-de-l-influenza. Last accessed: 15 June, 2014. **2013**.

15. Zambon MC, Stockton JD, Clewley JP, Fleming DM. Contribution of influenza and respiratory syncytial virus to community cases of influenza-like illness: an observational study. Lancet **2001**; 358(9291): 1410-6.

16. Call SA, Vollenweider MA, Hornung CA, Simel DL, McKinney WP. Does this patient have influenza? JAMA **2005**; 293(8): 987-97.

17. Chretien JP, George D, Shaman J, Chitale RA, McKenzie FE. Influenza forecasting in human populations: a scoping review. PLoS One **2014**; 9(4): e94130.

18. Bowie C, Prothero D. Finding causes of seasonal diseases using time series analysis. Int J Epidemiol **1981**; 10(1): 87-92.

19. Helfenstein U. Box-Jenkins modelling in medical research. Stat Methods Med Res **1996**; 5(1): 3-22.

20. Fisman D. Seasonality of viral infections: mechanisms and unknowns. Clin Microbiol Infect **2012**; 18(10): 946-54.

21. Bloom RM, Buckeridge DL, Cheng KE. Finding leading indicators for disease outbreaks: filtering, cross-correlation, and caveats. J Am Med Inform Assoc **2007**; 14(1): 76-85.

22. Box G, Jenkins G, Reinsel G. Time series analysis: forecasting and control. Fourth ed. Hoboken, NJ: John Wiley & Sons, **2008**.

23. Brockwell PJ, Davis RA. Time series: theory and methods. Second ed. New York: Springer, **1991**.

24. Bresee J, Hayden FG. Epidemic influenza--responding to the expected but unpredictable. N Engl J Med **2013**; 368(7): 589-92.

25. Gilca R, De Serres G, Skowronski D, Boivin G, Buckeridge DL. The need for validation of statistical methods for estimating respiratory virus-attributable hospitalization. Am J Epidemiol **2009**; 170(7): 925-36.

26. Sugawara T, Ohkusa Y, Ibuka Y, Kawanohara H, Taniguchi K, Okabe N. Real-time prescription surveillance and its application to monitoring seasonal influenza activity in Japan. J Med Internet Res **2012**; 14(1): e14.

27. Centers for Disease Control and Prevention. Guidance for clinicians on the use of RT-PCR and other molecular assays for diagnosis of influenza virus infection.

http://www.cdc.gov/flu/professionals/diagnosis/molecular-assays.htm Last accessed: 20 July, 2014. **2013**.

28. Greene SK, Kulldorff M, Huang J, et al. Timely detection of localized excess influenza activity in Northern California across patient care, prescription, and laboratory data. Stat Med **2011**; 30(5): 549-59.

29. Borders-Hemphill V, Mosholder A. U.S. utilization patterns of influenza antiviral medications during the 2009 H1N1 influenza pandemic. . Influenza Other Respir Viruses **2012**; 6(6): e129-33.

30. Yoshida M, Matsui T, Ohkusa Y, et al. Seasonal influenza surveillance using prescription data for anti-influenza medications. Jpn J Infect Dis **2009**; 62(3): 233-5.

31. Aramini JJ, Muchall PK, Pollari F. Value of pharmacy-based influenza surveillance - Ontario, Canada, 2009. MMWR Morb Mortal Wkly Rep **2013**; 62(20): 401-4.

32. Baz M, Abed Y, Papenburg J, Bouhy X, Hamelin ME, Boivin G. Emergence of oseltamivir-resistant pandemic H1N1 virus during prophylaxis. N Engl J Med **2009**; 361(23): 2296-7.

33. Brett AS, Zuger A. The run on tamiflu--should physicians prescribe on demand? N Engl J Med **2005**; 353(25): 2636-7.

34. Ortiz JR, Kamimoto L, Aubert RE, et al. Oseltamivir prescribing in pharmacy-benefits database, United States, 2004-2005. Emerg Infect Dis **2008**; 14(8): 1280-3.

35. van den Wijngaard CC, van Pelt W, Nagelkerke NJ, Kretzschmar M, Koopmans MP. Evaluation of syndromic surveillance in the Netherlands: its added value and recommendations for implementation. Euro Surveill **2011**; 16(9).

36. Schanzer D, Vachon J, Pelletier L. Age-specific differences in influenza A epidemic curves: do children drive the spread of influenza epidemics? Am J Epidemiol **2011**; 174(1): 109-17.

37. Kramarz P, Monnet D, Nicoll A, Yilmaz C, Ciancio B. Use of oseltamivir in 12 European countries between 2002 and 2007--lack of association with the appearance of oseltamivir-resistant influenza A(H1N1) viruses. Euro Surveill **2009**; 14(5).

38. Thompson WW, Ridenhour BL, Barile JP, Shay DK. Time-series analyses of count data to estimate the burden of seasonal infectious diseases. Epidemiology **2012**; 23(6): 839-42; discussion 43-4.

Table 4-1. Circulating influenza strains, by season, from 2010 to 2013 in Quebec, Canada.

Season	A, H1N1 subtype (%)[a]	A, H3N2 subtype (%)[a]	B, Yamagata lineage (%)[a]	B, Victoria lineage (%)[a]
2010-11	A/California/07/2009 (2)	**A/Perth/16/2009 (87)**	*B/Wisconsin/01/2010 (0.4)	B/Brisbane/60/2008 (10.6)
2011-12	A/California/07/2009 (24)	A/Perth/16/2009 (21)	*B/Wisconsin/01/2010 (6)	**B/Brisbane/60/2008 (49)**
2012-13	A/California/07/2009 (5)	**A/Victoria/361/2011 (78)**	B/Wisconsin/01/2010 (16)	*B/Brisbane/60/2008 (1)

[a] Estimated percentage of total circulating strains for that year, based on genetic and antigenic characterisation of a sample of provincial surveillance viral isolates from throughout the season.

* Strains not included in that year's trivalent seasonal influenza vaccine

Bold face indicates the predominant circulating strain for that season

Table 4-2. Correlation coefficients in the cross-correlation functions between the time series of neuraminidase inhibitor prescription dispensing, acute-care hospital emergency department visits for influenza-like illness, and a common reference time series of laboratory confirmed influenza cases in Quebec, Canada, 2010 to 2013.

Time series	Peak correlation		Earliest statistically significant correlation	
	Lag (weeks)	Correlation	Lag (weeks)	Correlation
Overall (2010-2013)				
NI	0	0.68	1	0.22
ILI ED	1	0.50	1	0.50
2010-2011				
NI	0	0.50	1	0.34
ILI ED	0	0.67	1	0.32
2011-2012				
NI	0	0.54	1	0.44
ILI ED	0	0.35	0	0.35
2012-2013				
NI	0	0.73	1	0.22
ILI ED	1	0.61	1	0.61

NI, neuraminidase inhibitor

ILI, influenza-like illness

ED, Emergency department

Table 4-3. Fitted ARIMA model and Box-Jenkins transfer function models for the prediction of weekly cases of laboratory-confirmed influenza infection.

Time series	Form (p, d, q) [a]	Lag (weeks) [b]	*P* value [c]	AICc
Laboratory influenza cases	4,1,0	NA	NA	270.3
Single input models				
NI dispensing	4,1,0	+2	0.020	263.1
ED visits for ILI	4,1,0	0	0.001	262.6
Multivariable model (two input series)				
NI dispensing	0,1,4	+2	0.030	255.6
ED visits for ILI	0,1,4	0	0.001	

[a] Where p represents the order of the autoregressive term, d represents the order of differencing, and q represents the order of the moving average term

[b] The positive lag (i.e., lead-time, when the input series is lagged to precede the reference indicator output series) with a statistically significant cross-correlation that minimized the AICc, and optimised model fit

[c] The *P* value of the coefficient for the input series

ARIMA, autoregressive integrated moving average

AICc, Akaike's Information Criterion corrected

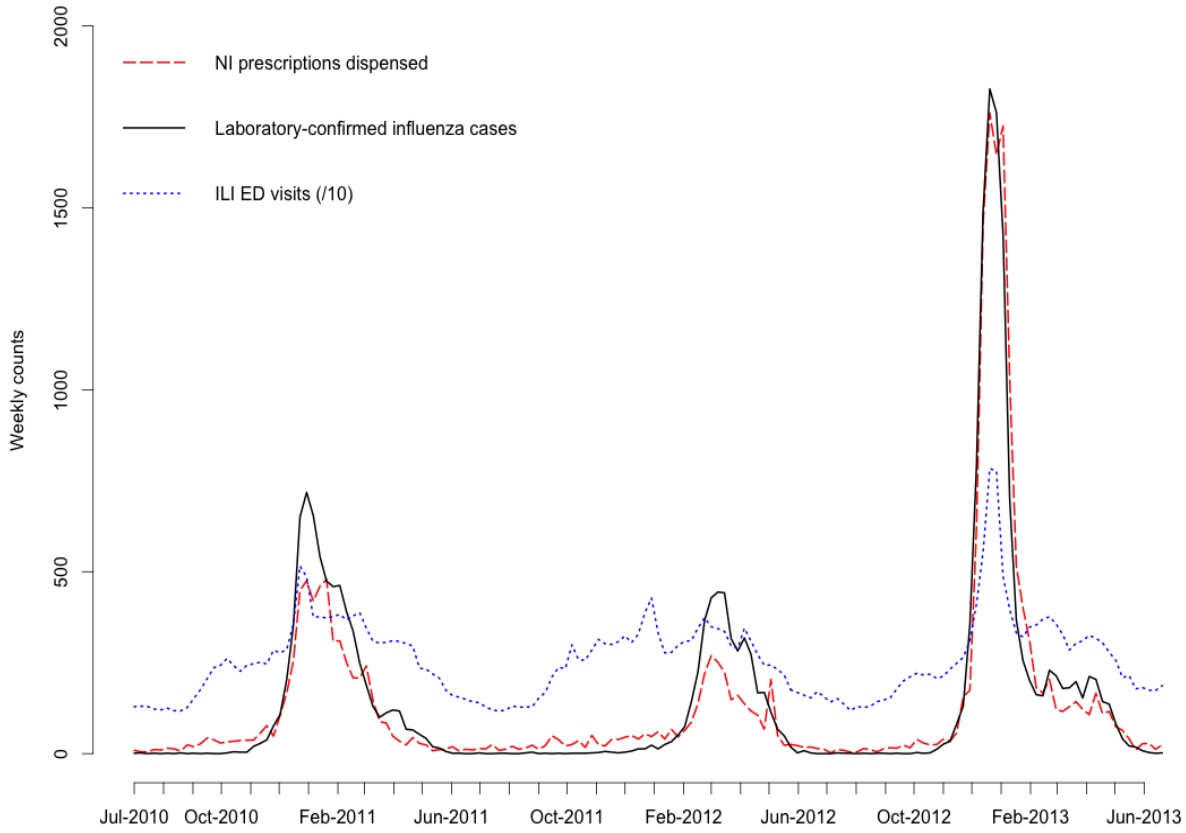
NA, not applicable

NI, neuraminidase inhibitor

ILI, influenza-like illness

ED, Emergency department

Figure 4-1. Time series plots of the weekly counts of neuraminidase inhibitor (NI) prescriptions dispensed, acute-care hospital emergency department (ED) visits for influenza-like illness (ILI), and laboratory confirmed cases of influenza in Quebec, Canada, 2010 to 2013.



CHAPTER 5 - SUMMARY AND CONCLUSIONS

In this dissertation, we evaluated the utility of monitoring NI dispensing in community retail pharmacies as a surveillance method for influenza at the provincial level, using time series methodology to account for autocorrelation. We compared NI dispensing to a reference standard for influenza activity, laboratory surveillance of influenza cases. The correlation, timeliness and predictive accuracy of NI dispensing data in relation to laboratory-confirmed influenza activity suggest that this readily available data stream could be used as a leading indicator for outbreak detection and situational awareness during seasonal epidemics. These observations were consistent over time, despite the different characteristics (timing, duration, peak intensity, vaccine effectiveness, and distribution of influenza subtypes) displayed by each of the three influenza seasons studied. To place our results in context with current influenza surveillance practice, which typically monitors several information sources in parallel, we performed our evaluation of NI dispensing data alongside an identical analysis of ED visits for ILI, a commonly used syndromic surveillance data stream.

Our results, their comparison with the literature, and their limitations were already discussed in the manuscript (chapter 4). In this chapter, a few additional issues will be addressed, in particular with regards to some of the implications of our findings for public health practice.

Analysis of aggregate weekly data allowed us to estimate the timeliness, strength of association and predictive accuracy of NI dispensing in relation to laboratory surveillance for influenza. However, several other important attributes for surveillance [11, 81, 91], introduced in section 2.2.2 of this thesis, could not be evaluated in the context of our study.

- Estimating the *sensitivity* and *positive and negative predictive values* of NI prescriptions for detecting individual cases of influenza would require diagnostic accuracy assessment of individual-level data against a reference standard diagnostic method. Regarding the *sensitivity* and *predictive values* for detecting influenza outbreaks, an potential approach would be to employ aberration detection algorithms [114] (e.g., cumulative sum [CUSUM], exponentially-weighted moving average or scan statistics methods) that generate alerts when a certain threshold is exceeded.
- The *simplicity* of NI surveillance was not formally assessed in this study, but we expect that it would be high, as very little data is necessary to establish that the health-related event being monitored has occurred; also, data collection and management should be straightforward because the data are available electronically.
- Monitoring prescription drug data may require public health agencies to establish partnerships with the owners of prescription sales data (health maintenance organisations, commercial providers of information services [e.g., IMS Brogan], or large retail pharmacy chains). The willingness of private partners to participate in such a program (*acceptability*) might be challenged by confidentiality issues, resource constraints, and lack of financial incentives. However, in jurisdictions with universal health care and electronic medical records, the public health agency may have ready access to such data for surveillance purposes.
- Establishing *data quality* would require a verification of the completeness and validity of the NI dispensing data. Although we could not perform such a verification for the Canadian Weekly CompuScript database used in this thesis, IMS Brogan informed us that they have standard operating procedures for quality control and quality assurance.

- To assess the *flexibility* of NI monitoring, we would need to observe how such a surveillance system would respond to a new demand; for instance, the addition of a new medication to the NI class of drugs or a request to analyse daily (instead of weekly) counts.
- We found that the *representativeness* of NI dispensing as an indicator of influenza activity at the provincial level appears to be very good and did not fluctuate over time (three years). However, as discussed in Chapter 4, we could not assess this attribute across age groups or across geographic areas, nor do we know how it would perform during a pandemic.
- The *stability* of an NI surveillance system (i.e., the ability to collect, manage, and provide data properly without failure) is a very practical matter that will vary across databases, data providers, and the influenza monitoring program that the data is being reported to.

Despite the positive results presented in this thesis, it is unlikely that NI dispensing would be used as the sole indicator of influenza activity by a surveillance program. Although NI prescribing usually reflects a high degree of suspicion for influenza infection on the part of prescribing physician, it remains an indirect measure of influenza circulation – like all forms if syndromic surveillance. Behaviours unrelated to true influenza activity, such as stockpiling in the context of increased media coverage relating to potential influenza pandemics, can influence NI dispensing counts [115]. While these types of events might occasionally threaten the validity of NI sales as an indicator of influenza activity, they provide an additional motivation to collect this information. Oseltamivir and zanamivir are limited resources; shortages of oseltamivir have

occurred in North America, especially during peaks of influenza activity [116, 117]. Monitoring NI dispensing could allow public health officials to recognize impending shortages earlier and intervene in a timely manner to reduce the risk of lack of availability of treatment for those that need it most. A Dutch evaluation of multistream syndromic surveillance judged that there was value in following NI prescriptions during the 2009 A/H1N1 pandemic: "early in the pandemic, the reaction of the public to media reports on pandemic influenza was illustrated by sharp elevations in the number of oseltamivir prescriptions. This information was used to urge physicians to exercise restraint in prescribing oseltamivir, in order to decrease the risk of oseltamivir shortage and viral resistance later in the pandemic" [90].

In summary, in this thesis I have presented the first evaluation of the utility of NI dispensing data that used robust time-series analysis methods to account for autocorrelation. I demonstrated that this readily available and easily accessible electronic data source provides timely and accurate information for influenza surveillance. My findings suggest that influenza antiviral dispensing counts could provide early information and act as a leading indicator for outbreak detection. This novel syndromic surveillance method, especially in parallel to traditional sources of surveillance data, should increase public health practitioners' and clinicians' situational awareness of influenza activity, thereby facilitating timely population-level interventions and resource management, and informing testing and treatment decisions at the individual patient level.

REFERENCES

1. Lee LM, Thacker SB, Centers for Disease C, Prevention. The cornerstone of public health practice: public health surveillance, 1961--2011. MMWR Surveill Summ **2011**; 60 Suppl 4: 15-21.
2. Porta M, Greenland S, Last JM, eds. Dictionary of epidemiology. 5 ed. New York, NY: Oxford University Press, **2008**.
3. Langmuir AD. The surveillance of communicable diseases of national importance. N Engl J Med **1963**; 268: 182-92.
4. Thacker SB, Qualters JR, Lee LM, Centers for Disease C, Prevention. Public health surveillance in the United States: evolution and challenges. MMWR Surveill Summ **2012**; 61 Suppl: 3-9.
5. Nicholson KG, Wood JM, Zambon M. Influenza. Lancet **2003**; 362(9397): 1733-45.
6. Schanzer DL, McGeer A, Morris K. Statistical estimates of respiratory admissions attributable to seasonal and pandemic influenza for Canada. Influenza and other respiratory viruses **2013**; 7(5): 799-808.
7. Schanzer DL, Sevenhuysen C, Winchester B, Mersereau T. Estimating influenza deaths in Canada, 1992-2009. PLoS One **2013**; 8(11): e80481.
8. Aoki FY, Allen UD, Stiver HG, Evans GA. The use of antiviral drugs for influenza: Guidance for practitioners 2012/2013. Can J Infect Dis Med Microbiol **2012**; 23(4): e79-e92.
9. Public Health Agency of Canada. National Advisory Committee on Immunization statement on seasonal influenza vaccine for 2014-2015.

http://publications.gc.ca/site/eng/469063/publication.html Last accessed: 8 August, 2014. **2014**.

10. Dailey L, Watkins RE, Plant AJ. Timeliness of data sources used for influenza surveillance. J Am Med Inform Assoc **2007**; 14(5): 626-31.
11. Lombardo JS, Buckeridge DL, eds. Disease surveillance: a public health informatics approach. 1 ed. Hoboken, NJ: John Wiley & Sons, Inc. , **2007**.
12. Nsoesie EO, Brownstein JS, Ramakrishnan N, Marathe MV. A systematic review of studies on forecasting the dynamics of influenza outbreaks. Influenza and other respiratory viruses **2014**; 8(3): 309-16.
13. Bloom RM, Buckeridge DL, Cheng KE. Finding leading indicators for disease outbreaks: filtering, cross-correlation, and caveats. J Am Med Inform Assoc **2007**; 14(1): 76-85.
14. Egger JR, Hoen AG, Brownstein JS, Buckeridge DL, Olson DR, Konty KJ. Usefulness of school absenteeism data for predicting influenza outbreaks, United States. Emerg Infect Dis **2012**; 18(8): 1375-7.
15. Kom Mogto CA, De Serres G, Douville Fradet M, et al. School absenteeism as an adjunct surveillance indicator: experience during the second wave of the 2009 H1N1 pandemic in Quebec, Canada. PLoS One **2012**; 7(3): e34084.
16. Vergu E, Grais RF, Sarter H, et al. Medication sales and syndromic surveillance, France. Emerg Infect Dis **2006**; 12(3): 416-21.
17. Greene SK, Kulldorff M, Huang J, et al. Timely detection of localized excess influenza activity in Northern California across patient care, prescription, and laboratory data. Stat Med **2011**; 30(5): 549-59.
18. Welliver RC, Cherry JD, Boyer KM, et al. Sales of nonprescription cold remedies: a unique method of influenza surveillance. Pediatr Res **1979**; 13(9): 1015-7.

19. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature **2009**; 457(7232): 1012-4.

20. Brownstein JS, Freifeld CC, Chan EH, et al. Information technology and global surveillance of cases of 2009 H1N1 influenza. N Engl J Med **2010**; 362(18): 1731-5.

21. Yih WK, Teates KS, Abrams A, et al. Telephone triage service data for detection of influenza-like illness. PLoS One **2009**; 4(4): e5260.

22. Espino JU, Hogan WR, Wagner MM. Telephone triage: a timely data source for surveillance of influenza-like diseases. AMIA Annu Symp Proc **2003**: 215-9.

23. Hsieh YH, Chen KF, Gaydos CA, Rothman RE, Kelen GD. Antiviral prescriptions to U.S. ambulatory care visits with a diagnosis of influenza before and after high level of adamantane resistance 2005-06 season. PLoS One **2010**; 5(1): e8945.

24. Public Health Agency of Canada. FluWatch. http://www.phac-aspc.gc.ca/fluwatch/index-eng.php Last accessed: 20 July, 2014, **2014**.

25. Bowie C, Prothero D. Finding causes of seasonal diseases using time series analysis. Int J Epidemiol **1981**; 10(1): 87-92.

26. Box G, Jenkins G, Reinsel G. Time series analysis: forecasting and control. Fourth ed. Hoboken, NJ: John Wiley & Sons, **2008**.

27. Brownstein JS, Kleinman KP, Mandl KD. Identifying pediatric age groups for influenza vaccination using a real-time regional surveillance system. Am J Epidemiol **2005**; 162(7): 686-93.

28. Chan EH, Tamblyn R, Charland KM, Buckeridge DL. Outpatient physician billing data for age and setting specific syndromic surveillance of influenza-like illnesses. J Biomed Inform **2011**; 44(2): 221-8.

29. Miller B, Kassenborg H, Dunsmuir W, et al. Syndromic surveillance for influenzalike illness in ambulatory care network. Emerg Infect Dis **2004**; 10(10): 1806-11.

30. Fisman DN. Seasonality of infectious diseases. Annu Rev Public Health **2007**; 28: 127-43.

31. Pica N, Bouvier NM. Ambient temperature and respiratory virus infection. Pediatr Infect Dis J **2014**; 33(3): 311-3.

32. Azziz Baumgartner E, Dao CN, Nasreen S, et al. Seasonality, timing, and climate drivers of influenza activity worldwide. J Infect Dis **2012**; 206(6): 838-46.

33. Schanzer DL, Langley JM, Dummer T, Aziz S. The geographic synchrony of seasonal influenza: a waves across Canada and the United States. PLoS One **2011**; 6(6): e21471.

34. Bresee J, Hayden FG. Epidemic influenza--responding to the expected but unpredictable. N Engl J Med **2013**; 368(7): 589-92.

35. Monto AS, Koopman JS, Longini IM, Jr. Tecumseh study of illness. XIII. Influenza infection and disease, 1976-1981. Am J Epidemiol **1985**; 121(6): 811-22.

36. Monto AS. Epidemiology of viral respiratory infections. Am J Med **2002**; 112 Suppl 6A: 4S-12S.

37. Reed C, Biggerstaff M, Finelli L, et al. Novel framework for assessing epidemiologic effects of influenza epidemics and pandemics. Emerg Infect Dis **2013**; 19(1): 85-91.

38. Viboud C, Nelson MI, Tan Y, Holmes EC. Contrasting the epidemiological and evolutionary dynamics of influenza spatial transmission. Philos Trans R Soc Lond B Biol Sci **2013**; 368(1614): 20120199.

39. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. The genomic and epidemiological dynamics of human influenza A virus. Nature **2008**; 453(7195): 615-9.

40. Zhou H, Thompson WW, Viboud CG, et al. Hospitalizations associated with influenza and respiratory syncytial virus in the United States, 1993-2008. Clin Infect Dis **2012**; 54(10): 1427-36.

41. Quach C, Piche-Walker L, Platt R, Moore D. Risk Factors Associated With Severe Influenza Infections in Childhood: Implication for Vaccine Strategy. Pediatrics **2003**; 112(3): e197-201.

42. Nair H, Brooks WA, Katz M, et al. Global burden of respiratory infections due to seasonal influenza in young children: a systematic review and meta-analysis. Lancet **2011**; 378(9807): 1917-30.

43. Schanzer DL, Langley JM, Tam TW. Role of influenza and other respiratory viruses in admissions of adults to Canadian hospitals. Influenza Other Respi Viruses **2008**; 2(1): 1-8.

44. Schanzer DL, Langley JM, Tam TW. Hospitalization attributable to influenza and other viral respiratory illnesses in Canadian children. Pediatr Infect Dis J **2006**; 25(9): 795-800.

45. 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: f5061.

46. Thompson WW, Shay DK, Weintraub E, et al. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA **2003**; 289(2): 179-86.

47. World Health Organisation. Influenza (Seasonal). http://www.who.int/mediacentre/factsheets/fs211/en/ Last accessed 30 July, 2014. **2014**.

48. Petric M, Comanor L, Petti CA. Role of the laboratory in diagnosis of influenza during seasonal epidemics and potential pandemics. J Infect Dis **2006**; 194 Suppl 2: S98-110.

49. Centers for Disease Control and Prevention. Guidance for clinicians on the use of RT-PCR and other molecular assays for diagnosis of influenza virus infection.

http://www.cdc.gov/flu/professionals/diagnosis/molecular-assays.htm - clinical . Last accessed: 20 July, 2014. **2013**.

50. Public Health Agency of Canada. FluWatch - Definitions for the 2013-14 season. http://www.phac-aspc.gc.ca/fluwatch/13-14/def13-14-eng.php . Last accessed: 31 July, 2014. **2014**.

51. Centers for Disease Control and Prevention. Overview of Influenza Surveillance in the United States. http://www.cdc.gov/flu/weekly/overview.htm. Last accessed: 31 July 2014. **2013**.

52. Zambon MC, Stockton JD, Clewley JP, Fleming DM. Contribution of influenza and respiratory syncytial virus to community cases of influenza-like illness: an observational study. Lancet **2001**; 358(9291): 1410-6.

53. Monto AS, Gravenstein S, Elliott M, Colopy M, Schweinle J. Clinical Signs and Symptoms Predicting Influenza Infection. Arch Intern Med **2000**; 160(21): 3243-7.

54. Boivin G, Hardy I, Tellier G, Maziade J. Predicting influenza infections during epidemics with use of a clinical case definition. Clin Infect Dis **2000**; 31(5): 1166-9.

55. 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(3): e92153.

56. Papenburg J, Baz M, Hamelin ME, et al. Household transmission of the 2009 pandemic A/H1N1 influenza virus: elevated laboratory-confirmed secondary attack rates and evidence of asymptomatic infections. Clin Infect Dis **2010**; 51(9): 1033-41.

57. Call SA, Vollenweider MA, Hornung CA, Simel DL, McKinney WP. Does this patient have influenza? JAMA **2005**; 293(8): 987-97.

58. Ohmit SE, Monto AS. Symptomatic predictors of influenza virus positivity in children during the influenza season. Clin Infect Dis **2006**; 43(5): 564-8.

59. Govaert TM, Dinant GJ, Aretz K, Knottnerus JA. The predictive value of influenza symptomatology in elderly people. Fam Pract **1998**; 15(1): 16-22.

60. Papenburg J, Hamelin ME, Ouhoummane N, et al. Comparison of risk factors for human metapneumovirus and respiratory syncytial virus disease severity in young children. J Infect Dis **2012**; 206(2): 178-89.

61. Conway NT, Wake ZV, Richmond PC, et al. Clinical Predictors of Influenza in Young Children: The Limitations of "Influenza-Like Illness". Journal of the Pediatric Infectious Diseases Society **2013**; 2(1): 21-9.

62. Renaud C, Papenburg J. Chapter 12: Surveillance for respiratory viruses: influenza and RSV. In: Garcia L. Clinical Microbiology and Procedures Handbook, 4th Ed. Washington, DC: American Society for Microbiology Press, **2014**.

63. Chartrand C, Leeflang MM, Minion J, Brewer T, Pai M. Accuracy of rapid influenza diagnostic tests: a meta-analysis. Ann Intern Med **2012**; 156(7): 500-11.

64. Harper SA, Bradley JS, Englund JA, et al. Seasonal influenza in adults and children--diagnosis, treatment, chemoprophylaxis, and institutional outbreak management: clinical practice guidelines of the Infectious Diseases Society of America. Clin Infect Dis **2009**; 48(8): 1003-32.

65. Prendergast C, Papenburg J. Rapid antigen-based testing for respiratory syncytial virus: moving diagnostics from bench to bedside? Future Microbiol **2013**; 8(4): 435-44.

66. World Health Organisation. Recommended composition of influenza virus vaccines for use in the 2012-2013 northern hemisphere influenza season, 2012.

http://www.who.int/influenza/vaccines/virus/recommendations/2012_13_north/en/index.html **2012**.

67. Jefferson T, Rivetti A, Di Pietrantonj C, Demicheli V, Ferroni E. Vaccines for preventing influenza in healthy children. Cochrane Database Syst Rev **2012**; 8: Cd004879.

68. Jefferson T, Di Pietrantonj C, Rivetti A, Bawazeer GA, Al-Ansary LA, Ferroni E. Vaccines for preventing influenza in healthy adults. Cochrane Database Syst Rev **2010**; (7): Cd001269.

69. Jefferson T, Di Pietrantonj C, Al-Ansary LA, Ferroni E, Thorning S, Thomas RE. Vaccines for preventing influenza in the elderly. Cochrane Database Syst Rev **2010**; (2): Cd004876.

70. 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).

71. Skowronski DM, Janjua NZ, Sabaiduc S, et al. Influenza A/Subtype and B/Lineage Effectiveness Estimates for the 2011-2012 Trivalent Vaccine: Cross-Season and Cross-Lineage Protection With Unchanged Vaccine. J Infect Dis **2014**.

72. Skowronski DM, Janjua NZ, De Serres G, et al. A sentinel platform to evaluate influenza vaccine effectiveness and new variant circulation, Canada 2010-2011 season. Clin Infect Dis **2012**; 55(3): 332-42.

73. Jefferson T, Jones MA, Doshi P, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database Syst Rev **2014**; 4: Cd008965.

74. Hernan MA, Lipsitch M. Oseltamivir and risk of lower respiratory tract complications in patients with flu symptoms: a meta-analysis of eleven randomized clinical trials. Clin Infect Dis **2011**; 53(3): 277-9.

75. Centers for Disease Control and Prevention. Guidance on the Use of Influenza Antiviral Agents. http://www.cdc.gov/flu/professionals/antivirals/antiviral-use-influenza.htm. Last accessed: 31 July, 2014. . **2013**.

76. McGeer A, Green KA, Plevneshi A, et al. Antiviral therapy and outcomes of influenza requiring hospitalization in Ontario, Canada. Clin Infect Dis **2007**; 45(12): 1568-75.

77. Gilca R, De Serres G, Boulianne N, et al. Risk factors for hospitalization and severe outcomes of 2009 pandemic H1N1 influenza in Quebec, Canada. Influenza and other respiratory viruses **2011**; 5(4): 247-55.

78. Zarychanski R, Stuart TL, Kumar A, et al. Correlates of severe disease in patients with 2009 pandemic influenza (H1N1) virus infection. CMAJ **2010**; 182(3): 257-64.

79. Muthuri SG, Venkatesan S, Myles PR, et al. Effectiveness of neuraminidase inhibitors in reducing mortality in patients admitted to hospital with influenza A H1N1pdm09 virus infection: a meta-analysis of individual participant data. The lancet Respiratory medicine **2014**; 2(5): 395-404.

80. Baz M, Abed Y, Papenburg J, Bouhy X, Hamelin ME, Boivin G. Emergence of oseltamivir-resistant pandemic H1N1 virus during prophylaxis. N Engl J Med **2009**; 361(23): 2296-7.

81. Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. MMWR Recomm Rep **2004**; 53(Rr-5): 1-11.

82. Chretien JP, George D, Shaman J, Chitale RA, McKenzie FE. Influenza forecasting in human populations: a scoping review. PLoS One **2014**; 9(4): e94130.

83. Jajosky RA, Groseclose SL. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. BMC Public Health **2004**; 4: 29.

84. Buehler JW. Review of the 2003 National Syndromic Surveillance Conference--lessons learned and questions to be answered. MMWR Morb Mortal Wkly Rep **2004**; 53 Suppl: 18-22.

85. Henning KJ. What is syndromic surveillance? MMWR Morb Mortal Wkly Rep **2004**; 53 Suppl: 5-11.

86. Olson DR, Heffernan RT, Paladini M, Konty K, Weiss D, Mostashari F. Monitoring the impact of influenza by age: emergency department fever and respiratory complaint surveillance in New York City. PLoS Med **2007**; 4(8): e247.

87. Hiller KM, Stoneking L, Min A, Rhodes SM. Syndromic surveillance for influenza in the emergency department-A systematic review. PLoS One **2013**; 8(9): e73832.

88. Cadieux G, Buckeridge DL, Jacques A, Libman M, Dendukuri N, Tamblyn R. Accuracy of syndrome definitions based on diagnoses in physician claims. BMC Public Health **2011**; 11: 17.

89. Zheng W, Aitken R, Muscatello DJ, Churches T. Potential for early warning of viral influenza activity in the community by monitoring clinical diagnoses of influenza in hospital emergency departments. BMC Public Health **2007**; 7: 250.

90. van den Wijngaard CC, van Pelt W, Nagelkerke NJ, Kretzschmar M, Koopmans MP. Evaluation of syndromic surveillance in the Netherlands: its added value and recommendations for implementation. Euro Surveill **2011**; 16(9).

91. German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep **2001**; 50(Rr-13): 1-35; quiz CE1-7.

92. Institut national de santé publique. Surveillance de l'influenza et des infections respiratoires virales. http://www.inspq.qc.ca/influenza/surveillance-de-l-influenza. Last accessed: 15 June, 2014. **2013**.

93. Institut national de santé publique. Flash Grippe. http://www.msss.gouv.qc.ca/sujets/prob_sante/influenza/index.php?Flash_influenza. Last accessed: 20 July 2014. **2014**.

94. Yoshida M, Matsui T, Ohkusa Y, et al. Seasonal influenza surveillance using prescription data for anti-influenza medications. Jpn J Infect Dis **2009**; 62(3): 233-5.

95. Sugawara T, Ohkusa Y, Ibuka Y, Kawanohara H, Taniguchi K, Okabe N. Real-time prescription surveillance and its application to monitoring seasonal influenza activity in Japan. J Med Internet Res **2012**; 14(1): e14.

96. Greene SK, Shay DK, Yin R, et al. Patterns in influenza antiviral medication use before and during the 2009 H1N1 pandemic, Vaccine Safety Datalink Project, 2000-2010. Influenza Other Respi Viruses **2012**; 6(6): e143-51.

97. Patwardhan A, Bilkovski R. Comparison: Flu prescription sales data from a retail pharmacy in the US with Google Flu trends and US ILINet (CDC) data as flu activity indicator. PLoS One **2012**; 7(8): e43611.

98. Aramini JJ, Muchall PK, Pollari F. Value of pharmacy-based influenza surveillance - Ontario, Canada, 2009. MMWR Morb Mortal Wkly Rep **2013**; 62(20): 401-4.

99. Taylor R. Interpretation of the correlation coefficient J Diagn Med Sonogr **1990**; 1(1): 35-9.

100. Butler D. When Google got flu wrong. Nature **2013**; 494(7436): 155-6.

101. Cook AR, Chen MI, Pin Lin RT. Internet search limitations and pandemic influenza, Singapore. Emerg Infect Dis **2010**; 16(10): 1647-9.

102. Helfenstein U. Box-Jenkins modelling in medical research. Statistical methods in medical research **1996**; 5(1): 3-22.

103. Fisman D. Seasonality of viral infections: mechanisms and unknowns. Clin Microbiol Infect **2012**; 18(10): 946-54.

104. Thompson WW, Ridenhour BL, Barile JP, Shay DK. Time-series analyses of count data to estimate the burden of seasonal infectious diseases. Epidemiology **2012**; 23(6): 839-42; discussion 43-4.

105. Gilca R, De Serres G, Skowronski D, Boivin G, Buckeridge DL. The need for validation of statistical methods for estimating respiratory virus-attributable hospitalization. Am J Epidemiol **2009**; 170(7): 925-36.

106. Morgenstern H. Ecologic studies. In: Rothman KJ, Greenland S, Lash LL. Modern epidemiology. 3 ed. Philadelphia, PA, USA: Lippincott, Williams & Wilkins, **2008**:511-30.

107. Statistics Canada. Population by sex and age group, by province and territory. http://www40.statcan.gc.ca/l01/cst01/demo31a-eng.htm Last accessed: 3 August, 2014. **2013**.

108. World Health Organisation. Global Influenza Surveillance and Response System (GISRS). http://www.who.int/influenza/gisrs_laboratory/en/ Last accessed: 20 July, 2014. **2013**.

109. Institut national de santé publique. Diagnostic de laboratoire de la grippe – saison 2013-2014.

http://www.inspq.qc.ca/lspq/fichesPDF/annonce_diagnostic_laboratoire_grippe_2013_2014.pdf Last accessed: 26 July, 2014. **2013**.

110. Shumway RH, Stoffer DS. Time-series analysis and its applications. Third ed. New York: Springer, **2011**.

111. Brockwell PJ, Davis RA. Time series: theory and methods. Second ed. New York: Springer, **1991**.

112. World Health Organisation. WHO Global Epidemiological Surveillance Standards for Influenza. http://www.who.int/influenza/resources/documents/influenza_surveillance_manual/en/index.html. Last accessed: 20 July, 2014. **2012**.

113. Thompson WW, Comanor L, Shay DK. Epidemiology of seasonal influenza: use of surveillance data and statistical models to estimate the burden of disease. J Infect Dis **2006**; 194 Suppl 2: S82-91.

114. Cowling BJ, Wong IO, Ho LM, Riley S, Leung GM. Methods for monitoring influenza surveillance data. Int J Epidemiol **2006**; 35(5): 1314-21.

115. Ortiz JR, Kamimoto L, Aubert RE, et al. Oseltamivir prescribing in pharmacy-benefits database, United States, 2004-2005. Emerg Infect Dis **2008**; 14(8): 1280-3.

116. Brett AS, Zuger A. The run on tamiflu--should physicians prescribe on demand? N Engl J Med **2005**; 353(25): 2636-7.

117. Public Health Agency of Canada. News Release -- Government of Canada to address temporary shortage of flu drug oseltamivir (Tamiflu). http://www.phac-aspc.gc.ca/media/nr-rp/2013/2013_0108-eng.php Last accessed: 9 August, 2014. **2013**.