

**Augmenting Surveillance to Minimize the Burden of Norovirus-like Illness
in Ontario: Using Telehealth Data to Detect the Onset of Community
Activity**

by

Stephanie Lesia Hughes

**A Thesis
presented to
The University of Guelph**

**In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Population Medicine**

Guelph, Ontario, Canada

© Stephanie Lesia Hughes, April, 2018

ABSTRACT

AUGMENTING SURVEILLANCE TO MINIMIZE THE BURDEN OF NOROVIRUS-LIKE ILLNESS IN ONTARIO: USING TELEHEALTH DATA TO DETECT THE ONSET OF COMMUNITY ACTIVITY

Stephanie Lesia Hughes
University of Guelph, 2018

Advisor:
Professor A. Papadopoulos

This thesis sets out to describe the essential elements required in the creation of a syndromic surveillance system for human infectious disease. Using this information, this thesis also presents a novel early warning syndromic surveillance system for the winter norovirus season in the province of Ontario, Canada. Telehealth data were utilized as the source of syndromic data alongside laboratory data to confirm the onset of the winter norovirus season. Syndromic methods were selected for this novel surveillance system because they can predict disease outbreaks earlier than laboratory and other traditional surveillance methods and cover a wider scope of the population. Norovirus outbreaks benefit from such methods due to the increased coverage of disease in communities, which often go unreported.

A scoping review was performed to describe the current state of research of the creation of syndromic surveillance systems for human infectious disease. A narrative synthesis built upon the scoping review to describe in detail all of the essential elements required in the creation of a syndromic surveillance system for human infectious disease.

Descriptive analyses were performed on three sources of Ontario norovirus data spanning 2009 to 2014: positive samples submitted for laboratory testing, institutional outbreaks, and calls to telehealth with vomiting as a chief complaint. However, it is important to note that the calls to telehealth (syndromic data source) were only used as a proxy for norovirus. The gender and age

distribution, institutions most commonly affected, total outbreak counts, total positive sample counts, and seasonality of norovirus were determined. Using this information to define the winter norovirus season and out-of-season periods, Shewhart (control) chart methods were employed to create the early warning syndromic surveillance system using the telehealth data as the source of syndromic data. These data were shown to predict the laboratory results by two weeks, and annual early alarm thresholds for the winter norovirus season were devised.

This thesis demonstrates the unique role of telehealth syndromic surveillance for norovirus in Ontario. Although these methods are widely used in other countries, they are not used in Ontario at the present time. The results from this study serve as a proposal for implementation and utilization in Ontario public health. In doing so, this syndromic surveillance system could strengthen current surveillance methods and help determine the true burden of norovirus in Ontario.

ACKNOWLEDGEMENTS

Anyone who knows me well knows that I believe strongly in pursuing our passions in life. We only get one shot here on this earth -- therefore we should make the best of it and be as happy as possible. Spending time on activities we love, with people that truly love us and make us smile, and working in a field that leaves us feeling excited to get out of bed and head to work every morning. I try my best to live by these rules each and every day: watching the Jays, traveling, and taking time to relax; heading north to spend some time with my parents far from the city; getting spoiled by my brother and his husband due to their impressive culinary skills; having dinners or playing squash with friends; all the while absorbing every bit of information possible from this Ph.D. I have loved every minute of these past 3.5 years (no matter how busy), and cannot imagine what my career & life trajectories would have been without it.

To Andy, my advisor – you helped guide, inform, motivate, and teach me during this journey. You are also the reason I ventured on to this path, a fact for which I will be forever grateful. It was your encouragement and suggestion 3.5 years ago that opened my eyes to the world of opportunities possible from a Ph.D. Thank you for all of your help and all of the time you kindly took out of the busyness of life to meet with me, I could not have asked for a better advisor. Your passion for public health and food safety is admirable and extremely contagious!

To Alex, Ian, Scott, and Amy – you have been the most supportive, positive, helpful, and selfless advisory committee. Each of you have your own individual set of remarkable strengths and skills which you kindly shared with me during our time together.

- Alex, thank you for agreeing to collaborate with an unknown, norovirus-obsessed, Canadian Ph.D. candidate from Guelph 3.5 years ago. This study would simply not have been possible without you and your guidance. I am so incredibly fortunate to

have gotten to know you; you have passed on to me countless skills which inspired my degree and will continue to change my life as I begin my career. It is not everyday that one comes across someone with your leadership and selflessness. Oh, and by any chance, has the “drinkable” maple tea grown on you?

- Scott, thank you for being an inspiration to me not only during this Ph.D., but for many years prior. You were one of the first epidemiologists I got to know during my MPH degree, and one of the first experts in foodborne disease. Thank you kindly as well for your introduction to working with the WHO – that was one of my most notable introductions to the field of international foodborne disease. I don't think I would have ended up here, had it not been for you and your influence.
- Ian, thank you for all of the hours you spent teaching me the art of the scoping review! You are truly an expert. I am so appreciative to have learned this valuable skill from you, as I had little exposure prior to this degree. You were always there to meet in person or reply to my many e-mails to help whenever I reached out. I also want to thank you for all of the detailed assistance you provided to the entire thesis; I have learned to be much more meticulous in my work and writing as a result.
- Amy, thank you for your incredible insight into the field of infectious disease dynamics and statistical modeling. You have been a big influence during this process, helping to guide my analyses, recommend relevant conferences, and improve my STATA code. It has always felt like you were in my corner and supported me. Lastly, thank you for all your valuable advice for my personal development moving forward.

Where do I even start regarding the entire Real-time Syndromic Surveillance Team at Public Health England... Roger, Gillian, Paul, Helen (x 2), Sue, Sally, Trudy, Winnie, Liz, Ana

(in addition to Alex!), you've all been so incredibly helpful, supportive, welcoming, and inspirational. You are truly world experts in what you do, and I am so fortunate to have learned from the best. Thank you for welcoming me in Birmingham for my visits, and teaching me everything there is to know about syndromic surveillance. I enjoyed all of our time spent together – both our times at the PHE office as well as out and about in Birmingham at the Christmas market, “proper English pubs,” and Nando's – all very “morish!”

It is absolutely essential that I thank some of my mentors/colleagues here as well: David James Allen, Beth Driscoll, Derek Ritz, Steffen Graether, and Richard Booth. You have all had a direct impact on my personal development over these past 3.5 years, for which I am incredibly grateful. Thank you all for the get-togethers, collaborations, calls, e-mails, updates, and overall support. And thank you for understanding my love of epidemiology & norovirus!

Lastly, but definitely not least, thank you to all of my family and friends. Mom, Dad, Paul, Stu – you all believed in me right from the start. And I know you have all believed in me in everything I've done. I can't express to you how much that has meant to me. It is because of you that I am the positive, happy, and determined person I am today. In addition, thank you to all of my closest friends, you are all wonderful. Not to mention H.E.N. – the most helpful and supportive landlords a person could ever ask for. Each and every one of you has had a life-changing impact on me, my life would not be what it is without you guys. You've all been there to help lift me through rough times and take me out for some fun when needed. Most importantly, you all make me feel supported, loved, and accepted exactly the way I am. (In all my food safety and norovirus-obsessed weirdness.)

STATEMENT OF WORK

Stephanie Lesia Hughes, under the guidance of Dr. Andrew Papadopoulos, designed all projects included in this thesis. She also requested, acquired, and analyzed the datasets used, and wrote this dissertation. Assistance regarding project design, data analyses and interpretation, chapter writing, and dissertation editing was provided by Dr. Andrew Papadopoulos, as well as the advisory committee: Drs. Alex J. Elliot, Amy L. Greer, Ian Young, and Scott A. McEwen. Assistance with data analyses was also provided by Dr. Roger Morbey and William Sears.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
STATEMENT OF WORK	vii
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xii
LIST OF APPENDICES	xiii
CHAPTER 1: Introduction	1
CHAPTER 2: Essential elements of human infectious disease syndromic surveillance systems: A scoping review of research from 2004 to 2016	26
ABSTRACT	26
INTRODUCTION	27
METHODS	28
RESULTS	31
DISCUSSION	33
FIGURES	43
CHAPTER 3: Essential elements of human infectious disease syndromic surveillance systems: A narrative synthesis	45
ABSTRACT	45
INTRODUCTION	46
METHODS	47
RESULTS	48
DISCUSSION	55
TABLES	66
CHAPTER 4: Estimating the burden of norovirus and viral gastroenteritis in Ontario, Canada – 2009-2014	68
ABSTRACT	68
INTRODUCTION	69
METHODS	70
RESULTS	72
DISCUSSION	75
TABLE.....	85

FIGURES	86
CHAPTER 5: Monitoring telehealth vomiting calls as a potential public health early warning system for seasonal norovirus activity in Ontario, Canada	91
ABSTRACT	91
INTRODUCTION	92
METHODS	94
RESULTS	97
DISCUSSION	99
FIGURES	111
CHAPTER 6: Conclusion.....	113
MAJOR FINDINGS.....	114
STRENGTHS.....	116
LIMITATIONS.....	117
FUTURE RESEARCH	118
REFERENCES.....	120
APPENDICES	123

LIST OF TABLES

Table 2.1. Descriptive characteristics of 55 selected articles describing the essential elements of human infectious disease syndromic surveillance systems

Table 3.1. List of all essential elements of syndromic surveillance systems for human infectious disease, organized by category and associated subcategories

Table 3.2. List of data sources utilized in syndromic surveillance systems

Table 4.1. List of acute gastroenteritis chief complaints to Telehealth Ontario and those selected as the “relevant chief complaints” for norovirus and gastroenteritis-related illness

Table 5.1. Pseudo- R^2 values for all negative binomial regression models

Table 5.2. P-values for norovirus laboratory data predictor variable in all negative binomial regression models

Table 5.3. Telehealth vomiting call thresholds and start week for winter norovirus activity; values by standard deviation increment used for thresholds

LIST OF FIGURES

Figure 2.1. Document screening and selection process for scoping review of human infectious disease syndromic surveillance systems

Figure 2.2. Analysis of the various terms used in the 55 documents to refer to “syndromic surveillance” and their frequency, by publication date

Figure 4.1. Percent of institutions affected by viral gastroenteritis and norovirus outbreaks in Ontario, 2009-2014

Figure 4.2. Seasonality of norovirus laboratory submissions and outbreaks in Ontario by month and year, 2009-2014.

Figure 4.3. Number of positive specimens submitted to PHO by virus type and year, 2009-2014 (per 100,000 population).

Figure 4.4. Percent of viral gastroenteritis (VGE) activity attributed to norovirus.

Figure 4.5. Percentage of acute gastroenteritis calls to Telehealth Ontario attributed to relevant chief complaints (vomiting, diarrhea, and vomiting+diarrhea calls) by year, 2009-2014

Figure 5.1. Weekly vomiting telehealth calls (“vomiting” and “vomiting and diarrhea” chief complaints) and laboratory positive norovirus submissions, 17/6/2011 – 30/9/2014

Figure 5.2. Weekly telehealth vomiting calls adjusted to remove long-term trend, with the ‘baseline’ mean number of calls and standard deviations marked, 19/6/2011 – 29/8/2015

LIST OF ABBREVIATIONS

AGE	Acute gastroenteritis
CDC	U.S. Centers for Disease Control and Prevention
ED	Emergency department
GP	General practitioner
iPHIS	integrated Public Health Information System
NHS	U.K. National Health Service
PHO	Public Health Ontario
PHOL	Public Health Ontario Laboratories
PHU	Public health unit
SD	Standard deviation
THO	Telehealth Ontario
VGE	Viral gastroenteritis

LIST OF APPENDICES

Appendix 2.1. Detailed protocol for scoping review

Appendix 2.2. List of all 55 articles selected for inclusion in scoping review

Appendix 4.1. Details of three datasets used in norovirus descriptive statistics analyses

CHAPTER 1

Introduction

The unique characteristics and traits of norovirus infection have made it one of the most successful and impactful communicable diseases in the developed world. With a short incubation period, acute onset, memorable combination of symptoms, numerous modes of transmission, high infectivity, low mortality rate, limited immunity, and fast mutation rate, it has become a notable infectious disease (1).

Virology

Noroviruses comprise a group of genetically diverse viruses which belong to the genus *Norovirus*, in the family *Caliciviridae* (2). This group of viruses are the most commonly implicated cause of acute gastroenteritis (AGE), accounting for roughly 50% of outbreaks and greater than 90% of non-bacterial outbreaks worldwide (3–8).

History and clinical features

Norovirus illness was initially described as the “winter vomiting disease” by the pediatrician J. Zahorsky in 1929 (9). The name was derived from the pronounced, increased activity in winter seasons and symptoms he observed: vomiting, diarrhea (typically non-bloody), nausea, abdominal cramps, fever, and general malaise (8–10). Although vomiting and diarrhea are both common symptoms of the illness, vomiting is somewhat more prevalent in people greater than 1

year of age (8,11,12). The observations of Adler and Zickl (1969) following a large outbreak in Norwalk, Ohio, U.S. in 1968 demonstrated that 91 out of the 99 primary cases exhibited vomiting, while 38 exhibited diarrhea (12). In addition, a study performed by Kaplan et al. (1982) demonstrated that in 38 outbreaks of norovirus disease in the U.S. from 1976-1980, vomiting was present in >50% of cases in 89 outbreaks, and diarrhea in 74 outbreaks (13). The incubation period of the illness is relatively short, with an average incubation period of 1.2 days (range 0.5 – 2.6 days) (14). Symptoms generally last between 1 and 2 days; however, the duration, can be 3-5 days or greater in vulnerable populations such as young children, the elderly, or immunocompromised persons (13,15,16). The onset of symptoms is often acute with little to no prodrome (10,17,18). Individuals often experience the sudden onset of vomiting and/or diarrhea shortly after feeling healthy (10,19,20). The self-limiting nature of norovirus illness and short duration of symptoms discourage many sufferers from seeking physician assistance when ill (21,22). Consequently, there is a lack of sufficient, representative, and/or timely laboratory data available due to the under-reporting of norovirus illness (1,22–25).

The characteristics consistently present in outbreaks of norovirus illness led Kaplan et al. (1982) to develop a set of standards commonly referred to as “Kaplan’s criteria,” which are used to help identify an AGE outbreak as due to norovirus infection: vomiting in greater than 50% of cases, short incubation period (24-48 h) and duration of symptomatic illness (12-60 h), and absence of bacterial aetiologic agent (13). These criteria are particularly useful due to the lack of *in vitro* culturing and diagnostic methods available for *Norovirus* (13,16,24,26).

Burden

Several studies have been conducted in Canada and internationally to estimate the extent of AGE and norovirus under-reporting. Methods in these studies include extrapolating results from laboratory data using under-reporting multipliers to estimate case counts of those who do not present to health care or are based in institutions (22,27). Hall et al. (2011) demonstrated only a small percentage (15%) of AGE-affected cases seek physician assistance, and of those, samples for laboratory confirmation were only requested from 13% (21). Majowicz et al. (2015) performed a study estimating the under-reporting of AGE in the Canadian province of Ontario where they estimated that for every one case reported to the province, there are approximately 313 additional cases that occur in the community, but that are not diagnosed or reported to public health authorities (27). Comparatively, Cowden (2002) found great under-reporting of norovirus illness in the U.K. with approximately 1500 additional community cases for every one reported case; this number decreased to 300 for non-institutional, community-acquired cases by accounting for those residing in hospital and long-term care homes (18). The research performed by Verhoef et al. (2012) supplemented these results by suggesting the burden of norovirus disease in the community was greater than in institutional settings; however, various gaps in data affected analyses (28). These studies highlight the bias towards institutional cases and lack of community data.

Complications

Complications from norovirus illness are more common in vulnerable populations, such as young children, immunocompromised individuals, or the elderly (8,13,29,30). Examples of complications include prolonged illness, severe dehydration, hospitalization, sequelae, and/or death. Although complications can occur in any age group, they are most common among the elderly (≥ 65 years) due to weakened immune systems and a higher frequency of comorbidities (10,13,16,29). Possible sequelae from infection include hypovolemia, hypokalemia, electrolyte imbalance, and renal insufficiency (1,31,32). Additional research has identified possible links between norovirus disease and encephalopathy, and also the incidence of postinfectious irritable bowel syndrome for up to three months following infection (28,33). Death, although rare, may occur and is more likely in vulnerable populations. Mortality estimates for the illness include 0.27 deaths per 100,000 in the U.S. and 0.40 deaths per 100,000 in the Netherlands (28,34).

Outbreaks and detection

Considerable research has been performed on noroviruses and norovirus illness throughout the past few decades; however, its identification is still relatively recent, limiting the amount of information available. The first notable outbreak of the illness occurred in Oct-Nov 1968, when a large number of schoolchildren acquired AGE in Norwalk, Ohio, U.S. A total of 116 of the students and staff (out of a total of 232) were affected with what was identified as the winter vomiting disease due to the unique combination of symptoms (12). This high attack rate led to investigations by the National Communicable Disease Center in the U.S. and studies attempting

to identify the non-bacterial aetiologic agent (12,35). In 1972, the virus was detected for the first time using electron microscopy, and then in 1978 using a radioimmunoassay technique (35,36). For many years, the virus was categorized into a general grouping of viruses based on its morphology (“small round structured viruses”), and its type species was referred to using the location of its initial isolation (“*Norwalk-like virus*”). This was until preliminary research in 1978 and confirmatory research in the 1990s identified it as a member of the *Caliciviridae* family, and the International Committee on Taxonomy of Viruses officially renamed the genus to “*Norovirus*” and type species “*Norwalk virus*” in 2002 (15,16,37,38).

Transmission and control

The unique combination of *Norovirus* attributes has led to its high morbidity and transmissibility in human populations. For one, individuals can spread the virus prior to symptom onset as well as after symptoms have cleared, and not solely during peak symptom activity (11,39,40). Noroviruses exhibit great hardiness and can persist in the environment and under extreme temperatures for weeks or months, complicating mitigation efforts (15,41,42). The virus has a rapid mutation rate, with new epidemic strains afflicting populations every few years (15,43,44). This is in combination with the fact there is incomplete cross-protection and a lack of long-term host immunity following infection, producing a continual supply of susceptible individuals (15,45). There is also significant strain diversity, with five genogroups existing in total (GI, II, and IV affect humans, GIII and V animals such as pigs, cows, and mice) (15,16,46). Lastly, the infectious dose of *Norovirus* is low, with studies showing doses as low as 10-20 virions resulting

in illness, as demonstrated in volunteer challenge studies (15,17,47); an infectious dose as low as this is a major contributing factor to large norovirus outbreaks.

Norovirus is transmitted via aerosolized vomitus and the fecal-oral route (15,16,29). More specifically, spread occurs via contaminated food, water, and fomites, aerosolized droplets, and person-to-person contact (15,17). Outbreaks are often attributed to foodborne transmission resulting from infectious food handlers or environmental contamination of food, leading to large outbreaks originating from a myriad of food sources (1,15,25,48,49). In Canada, it is estimated that 31% of norovirus outbreaks are due to foodborne transmission (23). Due to its many transmission routes and high transmissibility, outbreaks often occur in high-density venues such as cruise ships, camps, retirement homes, hospitals, hotels, and schools (13,50–53).

Diagnosis and under-reporting

Currently, surveillance of norovirus activity is heavily reliant on laboratory data (21). These data report the results of confirmatory tests performed on submitted samples, which provide the most accurate record of illness. A variety of testing methods are used worldwide to confirm the presence of noroviruses in submitted samples; polymerase chain reaction (PCR) is the most common due to its high sensitivity and specificity and low cost, although electron microscopy, immunochromatography tests, and enzyme-linked immunosorbent assays (ELISA) are also performed (54,55). These data are commonly used to assess norovirus illnesses and most other communicable diseases; they have been collected for many decades by countries worldwide and typically collect reportable disease data (illnesses which are mandated by law to be reported to

local Medical Officers of Health) (22,28,56,57). However, laboratory data are heavily biased towards institutional cases and vulnerable populations who seek clinical care (21,22). Samples are most often taken from very young or elderly individuals (e.g. those who are in long-term care homes, daycares, or are hospitalized), or from those who are residing in an institution at the time of an outbreak of norovirus illness (e.g. correctional facility) (21). Therefore, these data do not represent all ages, and only represent a small percentage of all norovirus-infected individuals. Therefore, reported data are not representative of the community picture of norovirus activity.

Another consequence of the under-reporting of norovirus illnesses is the lag in time between the onset of illness and reporting of illness to public health authorities (58). This can result in a delay in detection of changes in disease pattern (i.e. outbreak) and subsequent timely implementation of disease control strategies. The potential effects of these time lags in disease progression have been well-studied for other infectious diseases. Examples from other pathogens have illustrated these effects; for example, a delay in implementing control measures (e.g. alerts to public health authorities, quarantine, and hygiene procedures) by even one week during the Severe Acute Respiratory Syndrome (SARS) epidemics in Hong Kong, Vietnam, Canada, and Singapore was estimated to have potentially increased the sizes of the epidemics 2.6-fold and extended their duration by four weeks (59). In addition, delaying anthrax detection from an ideal 2 days to 4.8 days could potentially double the death count from the illness (60).

Traditional surveillance

Epidemiological surveillance is defined as “the ongoing systematic collection, analysis, and interpretation of health data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know” (61). The goal of disease surveillance is to reduce the spread and impact of infectious diseases in populations by linking public health, health care, and government to control disease occurrences in a timely manner (62). Effective surveillance systems are enabled by stakeholder engagement, collaboration, evaluation, and timely and complete data (63). In the Canadian province of Ontario, disease surveillance is conducted primarily using laboratory testing and confirmation, as well as by the reporting of outbreaks and estimated case counts, which are both directed by Public Health Ontario (PHO) (64–66). Thirty-six autonomous local public health units (PHU) across Ontario independently carry out infectious disease surveillance and investigation (i.e. handling cases, sample submission, data entry and reporting) (63). In addition, the structure and governance of health care and laboratory service providers differ across the province, although the overall services provided are uniform (63). These differences in public health and health care result in varying degrees of sufficient and timely data collected and reported across the province. In the case of norovirus illness, institutional outbreaks are reportable in Ontario, meaning that every outbreak of the illness which occurs in a long-term care home, retirement home, child care centre, or correctional facility must be reported to the local PHU (57). These differences in case management, data collection, data recording, case reporting, and organizational policies in the 36 local PHUs, health care facilities, and laboratory services contribute and complicate the data challenges of norovirus illnesses.

Economic impact

Studies have estimated that roughly 3.4 million Canadians are affected annually, resulting in 4,000-11,000 hospitalizations (23,67). Similarly, research has estimated that roughly 21-23 million Americans contract the illness every year, with 50,000-71,000 hospitalizations occurring (21,68,69). The economic impact of norovirus-associated hospitalization has been estimated at 21 million CAD in Canada, 500 million USD per year in the U.S., and 115 million GBP in the U.K. (30,67). These costs do not address the additional non-healthcare costs due to norovirus, such as work absenteeism and lost productivity, which are more challenging to quantify (10). Lopman et al. (2004) estimated some of these additional costs with a study focussed on gastroenteritis in Avon, England from 2002 to 2003. The authors estimated that ward closures cost hospitals 480,000 GBP per 1,000 beds and lost productivity cost hospitals (i.e. staff absences due to gastroenteritis) 66,000 GBP per 1,000 beds (30). These estimates are typically derived using available surveillance data, extrapolated to population levels. However, these data are impacted by the underreporting of the illness.

Syndromic surveillance

Relying on data which suffer from a myriad of challenges presents significant barriers for surveillance of norovirus disease in terms of identifying cases and outbreaks in a timely manner. To mitigate these issues, the implementation of a novel surveillance system for norovirus to complement current laboratory and outbreak surveillance methods in Ontario could be advantageous. Syndromic surveillance is a novel, underutilized type of surveillance which may

compensate for some of the deficiencies of traditional, passively collected surveillance data. It is defined as the “real-time (or near real-time) collection, analysis, interpretation and dissemination of health-related data to enable the early identification of the impact (or absence of impact) of potential human or veterinary public health threats which require effective public health action” (70). It utilizes pre-existing, pre-diagnostic data sources to detect community and population disease activity and environmental health events earlier than laboratory-based surveillance (70–73). These data sources are not typically collected for surveillance purposes, but can be used as proxy measures for disease activity. Examples of such data sources include emergency department (ED) visits, calls to telehealth services, pharmacy purchases, internet search queries, untreated sewage pathogen counts, visits to general practitioners (GPs), and school absenteeism (70,72,74–77). Due to the fact these data typically exist prior to their use in syndromic surveillance systems, these systems are also well-known for their strong economic values (70).

Certain data sources are timelier and/or more specific than others, depending on what they represent. The timeliest data sources are those which represent the earliest points of medical care and those which can be reported in a timely manner (i.e. ill persons may phone telehealth lines for advice and/or purchase over-the-counter pharmacy products prior to presenting at EDs) (77). Many illnesses present with a standard set of symptoms which can be grouped into syndromes (78). Syndromic surveillance systems monitor these syndromes to identify seasonal trends and increases which may indicate an outbreak.

Initially, syndromic surveillance systems were developed to detect the presence of agents released as part of bioterrorism plots (70,71,78); however, their utility has now been widened to support a range of different public health scenarios. They are effective at determining the seasonal spread of common infectious diseases (e.g. influenza), and providing intelligence on the start, severity, and cessation of activity. They can also be used to detect and determine the size and scope of local outbreaks (71). They are also helpful in monitoring populations and establishing when a potential threat does not exist (i.e. providing reassurance), which is particularly useful during mass gathering events (70). From the long list of benefits, syndromic surveillance systems are particularly known for their ability to detect disease in a timelier manner than laboratory or outbreak reporting methods, therefore enabling more timely public health intervention (71,78). The telehealth syndromic surveillance system built by Public Health England in the U.K. demonstrated that vomiting calls to NHS Direct predicted increases in norovirus activity from traditional surveillance methods with up to four weeks early warning (76). The Google internet query system created by Desai et al. (2012) showed that norovirus-related searches demonstrated a two week early warning ahead of norovirus illness emergency department data (79). Syndromic data tend to be more representative of illness in the community because many syndromic datasets (e.g. Google queries, school absenteeism, and calls to telehealth lines) capture the subset of the population which may not present to health care (75,77). The main drawback to these systems is the non-specific nature of the data analyzed. Because only proxy measures are used to estimate disease activity in populations, one particular illness cannot be ascertained from the data (80).

Syndromic surveillance systems complement and strengthen laboratory and outbreak surveillance systems by integrating novel data sources and techniques; this results in more sufficient and timelier data for use in surveillance, and a better understanding of disease across a broader scope of the population (70,81,82). Together, these surveillance systems work to identify cases which would not normally have been detected with the use of one system independently.

Reasons to improve surveillance

It is especially imperative that surveillance of norovirus disease is improved because of the lack of treatments available for the illness. There is neither a vaccine currently on the market for norovirus, nor are there medical treatments to improve or cure the illness (1,83). The only medical treatments currently available are supportive, remedies such as rehydration therapy or antiemetics, with a norovirus vaccine currently in human tolerability trials (1,16,30). Therefore, improving surveillance activities to monitor norovirus activity in the community would help to identify the onset of norovirus as early as possible and provide intelligence during the outbreaks to enable health care services to prepare for, and manage, outbreaks especially in settings involving vulnerable populations. In light of current vaccine developments, developing surveillance systems to track norovirus activity would also provide valuable “baseline” incidence data, illustrating norovirus burden pre-vaccine introduction and, thus, contributing to the estimation of vaccine effectiveness following its introduction. Because of the lack of concrete treatments, it is essential that surveillance functions to best capture as many cases as possible and reduce the impact in vulnerable populations who are disproportionately affected.

Current surveillance and needs in Ontario, Canada

In Ontario, the Public Health Ontario Laboratories (PHOL) laboratory dataset, and integrated Public Health Information System (iPHIS) outbreak dataset are compiled by PHO to report communicable disease information for public health mitigation strategies. The field of syndromic surveillance is becoming increasingly recognized in Ontario, with various studies being performed on the possible utilization of pre-existing data sources for syndromic purposes (75,84,85). At the present time, the only system which has been employed in routine public health across the province is the Acute Care Enhanced Surveillance syndromic surveillance system, developed by Kingston, Frontenac, and Lennox & Addington Public Health in Kingston, Ontario (86). This is an ED-based syndromic surveillance system which monitors increases in ED admissions to signal the onset of a range of potential public health threats (e.g. communicable disease outbreaks, severe weather events, drug overdoses) (86). Studies have been performed, analyzing the potential utilization of syndromic data in gastroenteritis detection; however, none have been specific to norovirus illness (75,87). The expansion of syndromic surveillance in Ontario could help to improve public health surveillance and disease detection. To accomplish this, additional systems could be developed with predetermined illnesses/syndromes of choice and sources of syndromic data selected *a priori*.

It is essential to understand the key components which are necessary for the design and construction of a syndromic surveillance system for norovirus infection. All required inputs, i.e. technical, legal, logistical, etc., must be understood, learning from international experts in the field of syndromic surveillance. This will help in the design of the system and position it well for

implementation in public health. In addition, it is crucial to observe and analyze norovirus seasonal activity and historical trends in Ontario to customize the syndromic surveillance system. No one system is universal, and each system must fit the specific needs of the population it is meant to serve. Finally, it is necessary to build the syndromic surveillance system for the winter norovirus season in Ontario, keeping the necessary components in mind as well as the patterns of activity seen in the province. This must be a fully comprehensive system which will detect the winter norovirus season earlier than current laboratory and outbreak surveillance methods.

Thesis purpose/objectives

The aims of the studies in this thesis are to:

1. Assess the essential elements of syndromic surveillance systems for human infectious disease, both identifying gaps in research areas and describing existing research.
2. Using a range of data sources, assess norovirus activity in Ontario from 2009-2014.
3. With this knowledge, develop an early warning system, using syndromic surveillance, for winter norovirus activity in Ontario.

The objectives of the studies in this thesis are to:

- Conduct a scoping review on the essential elements of human infectious disease syndromic surveillance systems. This will be a scoping rather than systematic review due to the fact it is still a relatively new area of study (88). This review will be broken into two chapters, one quantitative and one qualitative, to identify current research gaps and

describe the necessary components in starting a syndromic surveillance system from its conception to evaluation.

- Assess the descriptive statistics of norovirus activity in Ontario from 2009 to 2014.
- Develop a novel syndromic surveillance system for the winter norovirus season using systems developed by Public Health England as models.

REFERENCES

1. MacCannell T, Umscheid CA, Agarwal RK, Lee I, Kuntz G, Stevenson KB, et al. Guidelines for the Prevention and Control of Norovirus Gastroenteritis Outbreaks in Healthcare Settings. Centers for Disease Control and Prevention. 2011.
2. Glass RI, Noel J, Ando T, Fankhauser R, Belliot G, Mounts A, et al. The epidemiology of enteric caliciviruses from humans: A reassessment using new diagnostics. *J Infect Dis.* 2000;181(s2):S254–61.
3. Buesa J, Collado B, López-Andújar P, Abu-Mallouh R, Diaz JR, Diaz AG, et al. Molecular epidemiology of Caliciviruses causing outbreaks and sporadic cases of acute gastroenteritis in Spain. *J Clin Microbiol.* 2002;40(8):2854–9.
4. Ike AC, Brockmann SO, Hartelt K, Marschang RE, Contzen M, Oehme RM. Molecular epidemiology of norovirus in outbreaks of gastroenteritis in Southwest Germany from 2001 to 2004. *J Clin Microbiol.* 2006;44(4):1262–7.
5. Lopman BA, Reacher MH, Duijnhoven Y Van, Hanon F-X, Brown D, Koopmans M. Viral gastroenteritis outbreaks in Europe, 1995 – 2000. *Emerg Infect Dis.* 2003;9(1):90–6.
6. Reuter G, Krisztalovics K, Vennema H, Koopmans M, Szucs G. Evidence of the etiological predominance of norovirus in gastroenteritis outbreaks - Emerging new-variant

- and recombinant noroviruses in Hungary. *J Med Virol.* 2005;76(4):598–607.
7. van Duynhoven YTHP, de Jager CM, Kortbeek LM, Vennema H, Koopmans MPG, van Leusden F, et al. A one-year intensified study of outbreaks of gastroenteritis in The Netherlands. *Epidemiol Infect.* 2005;133(1):9–21.
 8. Patel MM, Hall AJ, Vinjé J, Parashar UD. Noroviruses: A comprehensive review. *J Clin Virol.* 2009;44(1):1–8.
 9. Zahorsky J. Hyperemesis Hiemis or the winter vomiting disease. *Arch Pediatr Adolesc Med.* 1929;46:391–5.
 10. Belliot G, Lopman BA, Ambert-Balay K, Pothier P. The burden of norovirus gastroenteritis: An important foodborne and healthcare-related infection. *Clin Microbiol Infect.* 2014;20(8):724–30.
 11. Rockx B, de Wit M, Vennema H, Vinje J, Bruin E De, van Duynhoven Y, et al. Natural history of human Calicivirus infection : A prospective cohort study. *Clin Infect Dis.* 2002;35:246–53.
 12. Adler JL, Zickl R. Winter Vomiting Disease. *J Infect Dis.* 1969;119:668–73.
 13. Kaplan JE, Gary GW, Baron RC, Singh N, Schonberger LB, Feldman R, et al. Epidemiology of Norwalk gastroenteritis and the role of Norwalk virus in outbreaks of acute nonbacterial gastroenteritis. *Ann Intern Med.* 1982;96:756–61.
 14. Lee RM, Lessler J, Lee RA, Rudolph KE, Reich NG, Perl TM, et al. Incubation periods of viral gastroenteritis: A systematic review. *BMC Infect Dis.* 2013;13(1):1.
 15. Glass RI, Parashar UD, Estes MK. Norovirus gastroenteritis. *N Engl J Med.* 2009;361:1776–85.
 16. Atmar RL, Estes MK. The epidemiologic and clinical importance of norovirus infection.

- Gastroenterol Clin North Am. 2006;35(2):275–90.
17. Caul EO. Small round structured viruses: airborne transmission and hospital control. Lancet. 1994;343:1240–2.
 18. Cowden JM. Winter vomiting. BMJ. 2002;324(7332):249–50.
 19. Thornley CN, Emslie NA, Sprott TW, Greening GE, Rapana JP. Recurring norovirus transmission on an airplane. Clin Infect Dis. 2011;53(6):515–20.
 20. Marks PJ, Vipond IB, Carlisle D, Deakin D, Fey RE, Caul EO. Evidence for airborne transmission of Norwalk-like virus (NLV) in a hotel restaurant. Epidemiol Infect. 2000;124:481–7.
 21. Hall AJ, Rosenthal M, Gregoricus N, Greene SA, Ferguson J, Henao OL, et al. Incidence of acute gastroenteritis and role of norovirus, Georgia, USA, 2004-2005. Emerg Infect Dis. 2011;17(8):1381–8.
 22. Bernard H, Werber D, Höhle M. Estimating the under-reporting of norovirus illness in Germany utilizing enhanced awareness of diarrhoea during a large outbreak of Shiga toxin-producing *E. coli* O104:H4 in 2011 - a time series analysis. BMC Infect Dis. 2014;14:116.
 23. Thomas MK, Murray R, Flockhart L, Pintar K, Pollari F, Fazil A, et al. Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. Foodborne Pathog Dis. 2013 Jul;10(7):639–48.
 24. Yen C, Hall AJ. Challenges to estimating norovirus disease burden. J Pediatric Infect Dis Soc. 2013;2(1):61–2.
 25. Martinez A, Dominguez A, Torner N, Ruiz L, Camps N, Barrabeig I, et al. Epidemiology of foodborne Norovirus outbreaks in Catalonia, Spain. BMC Infect Dis. 2008;8:47.

26. Turcios RM, Widdowson M-A, Sulka AC, Mead PS, Glass RI. Reevaluation of epidemiological criteria for identifying outbreaks of acute gastroenteritis due to norovirus: United States, 1998-2000. *Clin Infect Dis.* 2006;42(7):964–9.
27. Majowicz SE, Edge VL, Fazil A, McNab WB, Doré K a., Sockett PN, et al. Estimating the under-reporting rate for infectious gastrointestinal illness in Ontario. *Can J Public Heal.* 2005;96(3):178–81.
28. Verhoef L, Koopmans M, Van Pelt W, Duizer E, Haagsma J, Werber D, et al. The estimated disease burden of norovirus in The Netherlands. *Epidemiol Infect.* 2012;0(May 2012):1–11.
29. Lopman BA, Reacher MH, Vipond IB, Sarangi J, Brown DWG. Clinical manifestation of Norovirus gastroenteritis in health care settings. *Clin Infect Dis.* 2004;39:318–24.
30. Lopman BA, Reacher MH, Vipond IB, Hill D, Perry C, Halladay T, et al. Epidemiology and cost of nosocomial gastroenteritis, Avon, England, 2002-2003. *Emerg Infect Dis.* 2004;10(10):2002–3.
31. Ochi F, Furuno K, Chong PF, Tezuka J, Mizuno Y, Aoki T, et al. Bilateral hydronephrosis due to obstructive ureteral stone associated with norovirus gastroenteritis. *Clin Case Reports.* 2017;5(6):936–8.
32. Zenda T, Kaneko S, Noriki S. Norovirus gastroenteritis accompanied by ischemic colitis: A case report. *Hiroshima J Med Sci.* 2010;59(4):83–5.
33. Marshall JK, Thabane M, Borgaonkar MR, James C. Postinfectious irritable bowel syndrome after a food-borne outbreak of acute gastroenteritis attributed to a viral pathogen. *Clin Gastroenterol Hepatol.* 2007;5(4):457–60.
34. Hall AJ, Curns AT, Clifford McDonald L, Parashar UD, Lopman BA. The roles of

- Clostridium difficile* and norovirus among gastroenteritis-associated deaths in the United States, 1999-2007. Clin Infect Dis. 2012;55(2):216–23.
35. Kapikian AZ, Wyatt RG, Dolin R, Thornhill TS, Kalica AR, Chanock RM. Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. J Virol. 1972;10(5):1075–81.
 36. Greenberg HB, Wyatt RG, Valdesuso J, Kalica AR, London WT, Chanock RM, et al. Solid-phase microtiter radioimmunoassay for detection of the Norwalk strain of acute nonbacterial, epidemic gastroenteritis virus and its antibodies. 1978;2:97–108.
 37. Mayo MA. A summary of taxonomic changes recently approved by ICTV. Arch Virol. 2002;147(8):1655–6.
 38. McSwiggan DA, Cubitt D, Moore W. Calicivirus associated with winter vomiting disease. Lancet. 1978;1:1215.
 39. Atmar RL, Opekun AR, Gilger MA, Estes MK, Crawford SE, Neill FH, et al. Norwalk virus shedding after experimental human infection. Emerg Infect Dis. 2008;14(10):1553–7.
 40. Graham DY, Jiang X, Tanaka T, Opekun AR, Madore HP, Estes MK. Norwalk virus infection of volunteers: New insights based on improved assays. J Infect Dis. 1994;170(1):34–43.
 41. Hoelzer K, Fanaselle W, Pouillot R, Van Doren JM, Dennis S. Virus inactivation on hard surfaces or in suspension by chemical disinfectants: Systematic review and meta-analysis of norovirus surrogates. J Food Prot. 2013;76(6):1006–16.
 42. Li J, Predmore A, Divers E, Lou F. New interventions against human norovirus: Progress, opportunities, and challenges. Annu Rev Food Sci Technol. 2012;3(1):331–52.

43. Karst SM, Baric RS. What is the reservoir of emergent human norovirus strains? *J Virol.* 2015;89(11):5756–9.
44. Bull RA, Hansman GS, Clancy LE, Tanaka MM, Rawlinson WD, White PA. Norovirus recombination in ORF1/ORF2 overlap. *Emerg Infect Dis.* 2005;11(7):1079–85.
45. Parrino TA, Schreiber DS, Trier JS, Kapikian AZ, Blacklow NR. Clinical immunity in acute gastroenteritis caused by Norwalk agent. *N Engl J Med.* 1977;297:86–9.
46. Zheng DP, Ando T, Fankhauser RL, Beard RS, Glass RI, Monroe SS. Norovirus classification and proposed strain nomenclature. *Virology.* 2006;346(2):312–23.
47. Teunis PFM, Moe CL, Liu P, Miller SE, Lindesmith L, Baric RS, et al. Norwalk virus: How infectious is it? *J Med Virol.* 2008;80:1468–76.
48. Becker KM, Moe CL, Southwick KL, MacCormack JN. Transmission of Norwalk virus during football game. *N Engl J Med.* 2000;343:1223–7.
49. Vo TH, Okasha O, Al-Hello H, Polkowska A, Räsänen S, Bojang M, et al. An outbreak of norovirus infections among lunch customers at a restaurant, Tampere, Finland, 2015. *Food Environ Virol.* 2016;8(3):174–9.
50. Love SS, Jiang X, Barrett E, Farkas T, Kelly S. A large hotel outbreak of Norwalk-like virus gastroenteritis among three groups of guests and hotel employees in Virginia. *Epidemiol Infect.* 2002;129(1):127–32.
51. Wikswo ME, Cortes J, Hall AJ, Vaughan G, Howard C, Gregoricus N, et al. Disease transmission and passenger behaviors during a high morbidity norovirus outbreak on a cruise ship, January 2009. *Clin Infect Dis.* 2011;52(9):1116–22.
52. Morbidity and Mortality Weekly Report. Norovirus outbreak in an elementary school—District of Columbia, February 2007. *J Am Med Assoc.* 2008;299(6):627.

53. Luchs A, Morillo SG, Ribeiro CD, Vilanova BC, Calux SJ, de Cassia Compagnoli Carmona R, et al. Gastroenteritis outbreak due to G2P[4] rotavirus and GII norovirus at two correctional facilities in Brazil, 2010. *J Clin Virol.* 2011;51(3):213–4.
54. Rabenau HF, Stürmer M, Buxbaum S, Walczok A, Preiser W, Doerr HW. Laboratory diagnosis of norovirus: Which method is the best? *Intervirology.* 2003;46(4):232–8.
55. Fitzgerald TL, Zammit A, Merritt TD, McLeod C, Landinez LM, White PA, et al. An outbreak of norovirus genogroup II associated with New South Wales oysters. *Commun Dis Intell Q Rep [Internet].* 2014;38(1):E9–15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25409363>
56. Doyle TJ, Glynn MK, Groseclose SL. Completeness of notifiable infectious disease reporting in the United States: An analytical literature review. *Am J Epidemiol.* 2002;155(9):866–74.
57. Toronto Public Health. Communicable Disease Reporting [Internet]. 2015. Available from: http://www.health.state.mn.us/divs/idepc/dtopics/reportable/rule/hipaacomm.html%5Chttps://www.google.ca/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=0ahUKEwi5zsWd9d_RAhUm74MKHcW4CyIQFggrMAI&url=https%3A%2F%2Fwww1.toronto.ca%2FCity%2520Of%2520Toronto%2FTor
58. Ashford DA, Kaiser RM, Bales ME, Shutt K, Patrawalla A, McShan A, et al. Planning against biological terrorism: Lessons from outbreak investigations. *Emerg Infect Dis.* 2003;9(5):515–9.
59. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol.* 2004;160(6):509–16.

60. Wein LM, Craft DL, Kaplan EH. Emergency response to an anthrax attack. *Proc Natl Acad Sci U S A*. 2003;100(7):4346–51.
61. Thacker SB, Berkelman RL. Public health surveillance in the United States. *Epidemiol Rev*. 1988;10:164–90.
62. Centers for Disease Control and Prevention. CDC’s vision for public health surveillance in the 21st century. *MMWR*. 2012;61.
63. Ontario Agency for Health Protection and Promotion (Public Health Ontario). Infectious disease surveillance framework 2014-2019: Better data for better action. Toronto, ON: Queen’s Printer for Ontario; 2014.
64. Peel Public Health. iPHIS data [Internet]. 2018 [cited 2018 Jan 10]. Available from: https://www.peelregion.ca/health/statusdata/DataSources/HSD12_17.asp
65. Public Health Ontario. Laboratory Services [Internet]. 2018 [cited 2018 Jan 10]. Available from: <https://www.publichealthontario.ca/en/ServicesAndTools/LaboratoryServices/Pages/default.aspx>
66. Edge VL. Syndromic surveillance in public health. ProQuest Dissertations and Theses. [Ann Arbor]: University of Guelph (Canada); 2006.
67. Morton VK, Thomas MK, McEwen SA. Estimated hospitalizations attributed to norovirus and rotavirus infection in Canada, 2006 – 2010. *Epidemiol Infect*. 2015;143:3528–37.
68. Mead PS, Slutsker L, Dietz V, McCraig. LF, Bresee JS, Shapiro C, et al. Food-related illness and death in the United States. *Emerg Infect Dis*. 1999;5(5):607–25.
69. Lopman BA, Hall AJ, Curns AT, Parashar UD. Increasing rates of gastroenteritis hospital discharges in US adults and the contribution of norovirus, 1996-2007. *Clin Infect Dis*.

- 2011;52(4):466–74.
70. Triple-S Project. Guidelines for designing and implementing a syndromic surveillance system. 2013.
 71. Henning KJ. What is syndromic surveillance? *MMWR Suppl.* 2004;53(Sept):7–11.
 72. 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):pii=19806.
 73. Leonardi GS, Hajat S, Kovats RS, Smith GE, Cooper DL, Gerard E. Syndromic surveillance use to detect the early effects of heat-waves: An analysis of NHS direct data in England. *Soz Präventiv Med.* 2006;51(4):194–201.
 74. Hellmér M, Paxéus N, Magnus L, Enache L, Arnholm B, Johansson A, et al. Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks. *Appl Environ Microbiol.* 2014;80(21):6771–81.
 75. Edge VL, Pollari F, Ng LK, Michel P, McEwen SA, Wilson JB, et al. Syndromic surveillance of norovirus using over-the-counter sales of medications related to gastrointestinal illness. *Can J Infect Dis Med Microbiol.* 2006 Jul;17(4):235–41.
 76. Loveridge P, Cooper D, Elliot AJ, Harris J, Gray J, Large S, et al. Vomiting calls to NHS Direct provide an early warning of norovirus outbreaks in hospitals. *J Hosp Infect.* 2010;74:385–93.
 77. Pavlin JA. Syndromic surveillance for infectious diseases. In: *Infectious Disease Surveillance: Second Edition.* 2nd ed. Blackwell Publishing Ltd; 2013. p. 470–81.
 78. Yan P, Chen H, Zeng D. Syndromic surveillance systems. In: *Annual Review of Information Science and Technology.* University of Arizona; 2009. p. 425–95.

79. Desai R, Hall AJ, Lopman B a, Shimshoni Y, Rennick M, Efron N, et al. Norovirus disease surveillance using Google Internet query share data. *Clin Infect Dis* [Internet]. 2012 Oct [cited 2014 Nov 10];55(8):e75-8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22715172>
80. Ziemann A. *Syndromic surveillance: Made in Europe*. Maastricht University; 2015.
81. Ziemann A, Rosenkötter N, Garcia-Castrillo Riesgo L, Fischer M, Krämer A, Lippert FK, et al. Meeting the International Health Regulations (2005) surveillance core capacity requirements at the subnational level in Europe: The added value of syndromic surveillance. *BMC Public Health*. 2015;15(1):107.
82. Berger M, Shiau R, Weintraub JM. Review of syndromic surveillance: implications for waterborne disease detection. *J Epidemiol Community Health*. 2006 Jun 19;60(6):543–50.
83. Atmar RL, Bernstein DI, Harro CD, Al-Ibrahim MS, Chen WH, Ferreira J, et al. Norovirus vaccine against experimental human Norwalk virus illness. *N Engl J Med*. 2011;365:2178–87.
84. van Dijk A, McGuinness D, Rolland E, Moore KM. Can Telehealth Ontario respiratory call volume be used as a proxy for emergency department respiratory visit surveillance by public health? *Can J Emerg Med*. 2008;10(1):18–24.
85. Rolland E, Moore KM, Robinson VA, McGuinness D. Using Ontario’s “Telehealth” health telephone helpline as an early-warning system: a study protocol. *BMC Health Serv Res*. 2006 Jan;6:10.
86. KFL&A Public Health. *Acute Care Enhanced Surveillance* [Internet]. 2015. Available from: http://www.kflaphi.ca/wp-content/uploads/ACES_Brochure_for-KFLAPHI.pdf
87. Caudle JM, Van-Dijk A, Rolland E, Moore KM. Telehealth Ontario detection of

gastrointestinal illness outbreaks. *Can J Public Heal.* 2009;100(4):253–7.

88. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005 Feb 1;8(1):19–32.

CHAPTER 2

Essential elements of human infectious disease syndromic surveillance systems: A scoping review of research published from 2004 to 2016

ABSTRACT

Syndromic surveillance is an alternative form of public health surveillance that utilizes pre-diagnostic data to detect outbreaks earlier than conventional surveillance. This scoping review sought to identify and synthesize all existing literature detailing the components of the following themes: the definition, development, implementation, and evaluation of syndromic surveillance systems aimed at monitoring human infectious disease.

The review was conducted following internationally-recommended scoping review guidelines. Four bibliographic databases and 11 websites were searched for peer-reviewed and grey literature. The titles and abstracts were screened for relevance, and full-texts of relevant citations reviewed and characterized.

From a total of 1237 titles and abstracts, 55 relevant documents were identified. Primary research articles were the most commonly identified type of article [23/55 (41.8%)], and articles were most often published in the United States [36/55 (65.5%)]. The development of new syndromic surveillance systems was the most described theme out of the four aforementioned themes; data sources were discussed in 27/55 (49.1%) documents and statistical methods in 21/55 (38.2%) documents. The definition of syndromic surveillance was the least described with characteristics only being described in 6/55 (10.9%) and 3/55 (5.5%) documents.

Over the past decade, syndromic surveillance systems have become integral in public health practice internationally. This review identified that not all elements are equally represented in the literature, despite the fact all elements are important to those working with syndromic surveillance. The development and implementation of syndromic surveillance systems were greatly researched, in comparison to the definition and evaluation which have room for improvement.

INTRODUCTION

Historically, infectious disease surveillance has consisted of laboratory-based or self-reported illness records. These traditional surveillance strategies suffer from limitations, such as under-reporting and reporting time lags, which hinders the ability and timeliness of the public health response and intervention initiation (1). Syndromic surveillance can, in part, mitigate these issues and capture a more complete and rapid state of disease activity in populations (2,3).

Syndromic surveillance is the “real-time (or near-real-time) collection, analysis, interpretation and dissemination of health-related data to enable the early identification of the impact (or absence of impact) of potential human health... threats,” (4). This type of surveillance relies on pre-existing health data (e.g. purchases at pharmacies, visits to family physicians) and non-specific disease syndromes and symptoms to diagnose cases. These novel surveillance systems are used for both detecting disease activity in populations or documenting a lack of disease activity, and are used to detect and respond to health threats prior to laboratory diagnoses as early in the disease course as possible (1,5–11).

Early work in the field of syndromic surveillance focussed on the potential threat of bioterrorism activity, where its purpose is to provide some evidence that a population has not been deliberately exposed to a hazardous biological agent (5,7,10). The application of syndromic surveillance has subsequently expanded to a diverse range of both communicable and non-communicable diseases and environmental events such as earthquakes and heat waves (7). These systems have promising applications internationally in countries seeking to expand surveillance to monitor a wide variety of diseases and health outcomes in populations.

Syndromic surveillance systems must be carefully planned and developed. Work must be focussed on the definition, development, implementation, and evaluation of such systems. While these themes have been described in various individual studies (12–14), the insights and recommendations from this body of literature have not been previously synthesized using a transparent, structured approach. A comprehensive scoping review was performed to locate all relevant literature describing the definition, development, implementation, and evaluation of syndromic surveillance systems.

METHODS

Question and scope

A full scoping review protocol was devised *a priori* following internationally recommended methods (15). The research question was, “What are the essential elements involved in the definition, development, implementation, and evaluation of a fully functional syndromic surveillance system for human infectious disease?” The inclusion criteria were: primary research articles, review articles, government and research institute reports, theses, conference

proceedings, commentaries, and book chapters describing syndromic surveillance and its framework in terms of human health and infectious disease surveillance. Although syndromic surveillance systems encompass a diverse range of communicable and non-communicable disease, and natural health events, systems aimed to identify infectious disease were the focus of this review due to the abundance of information in the literature. Inclusion criteria included documents written in English and published from 2004 until March 2016 (at the time of the searches). A cut-off date of 2004 was selected based on when some notable studies on the framework of syndromic surveillance systems were published (5,11,14,16). Exclusion criteria for the scoping review included all other types of documents (i.e. websites, Microsoft PowerPoint presentations) which did not detail the essential elements of syndromic surveillance systems.

Search strategy

A detailed search algorithm was created using combinations of keywords extracted from the titles, abstracts, and full-texts of nine predetermined relevant documents; seven peer-reviewed articles and two government reports. The search algorithm was broken into three parts: the term "syndromic surveillance" as well as seven alternative phrases, followed by infectious disease, enteric, and gastrointestinal terms, and lastly by article description terms (define, implement or review). The full search algorithm used is described in *Appendix 2.1*.

The searches were performed on March 29, 2016 in four bibliographic databases: Scopus, PubMed, CINAHL, and ProQuest Dissertations and Theses Abstract & Indexing (A&I). Additional searches were also performed to identify grey literature from national and international sources (i.e. government and research agency reports). Two Google searches were

performed on March 30, 2016, searching for syndromic surveillance documents; however, for logistical reasons, only the first 100 hits in each search were examined for potential inclusion in the scoping review (full details of these searches available in *Appendix 2.1*). In addition, 10 targeted searches were performed on various government and non-profit agency websites for formal report documents. A search verification was then performed on the reference lists from all located documents for inclusion in the scoping review; however, all identified documents were either duplicates or did not satisfy the inclusion criteria.

Relevance screening

The titles and abstracts of all identified citations were reviewed independently by two reviewers for relevance to the review question. Only documents which pertained to syndromic surveillance, infectious disease, human health, and the framework (definition, development, implementation, and evaluation) of syndromic surveillance systems were included at this stage.

Data characterization and extraction

The full texts of all documents deemed potentially relevant at the title/abstract level were independently reviewed by two reviewers using a structured data characterization and extraction form (*Appendix 2.1*). This form, consisting of 21 questions, was devised *a priori* and was used to extract as much key information as possible from the relevant documents.

Review management and analysis

Search results were imported into the citation software Mendeley (Elsevier, Amsterdam, Netherlands) for duplicate removal and document management. Citations were then uploaded

into the systematic review software Covidence (Alfred Health, Melbourne, Australia) to facilitate relevance screening and characterization. Each review step was performed and independently verified by two reviewers; verification at each step was measured using Kappa agreement. Both reviewers met to discuss the results and clarify any discrepancies or areas of confusion. Once this was complete, the screening process proceeded. Conflicts were all resolved by the primary author. All findings from the data characterization and extraction step were entered into Microsoft Excel spreadsheets for descriptive analysis.

RESULTS

A total of 1237 unique citations were assessed in the title and abstract screening, which led to 142 full-texts screened. Out of these, 55 articles were confirmed as relevant for inclusion in the review (Figure 2.1). The information extracted from the 55 articles (*Appendix 2.2*) was organized into the four themes (definition, development, implementation, and evaluation of syndromic surveillance), as well as individual elements, using descriptions in literature for guidance.

The majority (36/55) of the selected articles were published in the United States, with others published in Canada and across Europe (Table 2.1). The number of articles published per year was stable over the review period (Table 2.1). Primary research articles were the main source of essential elements of human infectious disease syndromic surveillance systems. Other common types of articles were review articles, book chapters, and government reports (Table 2.1).

The term "syndromic surveillance" was used most commonly (53/55 articles; Figure 2.2). The term biosurveillance was the second most frequently used term (14/55), and is increasing in popularity, with many other terms also used in the literature (Figure 2.2).

Most of the articles which described the essential elements of syndromic surveillance systems for human infectious disease also described their applications (43/55) (Table 2.2). The three overarching uses of syndromic surveillance systems were described in the articles, with early warning the most commonly mentioned, followed by detection, and lastly situational awareness (Table 2.2).

The articles which described the definition of syndromic surveillance either focussed on terminology or determining the purpose and outcome(s) of syndromic surveillance systems. The various terms used to refer to syndromic surveillance (Figure 2.2) were described and defined in six and three articles, respectively (Table 2.2).

Similarly, many aspects of the development of new syndromic surveillance systems were described in the literature including: stakeholder identification, staff composition, data structure, data sources, classification of syndromes, and the development of alert response protocols and statistical methods (Table 2.2). The most commonly described element was the array of data sources used in the systems to detect disease (27/55), followed by the statistical methods employed to improve the detection of signals and alerts (21/55), and then by the classification of syndromes and structure of data sources (19/55) (Table 2.2).

The implementation of syndromic surveillance systems involves a variety of cost and infrastructure inputs and solutions to overcome operational and legal barriers, and requires collaborations with other types of public health surveillance to be successful and effective. Operational challenges were described most often (22/55), followed by the need for collaboration with other types of public health surveillance (13/55), and then costs and infrastructure (11/55) (Table 2.2).

The fourth theme in the creation of syndromic surveillance systems for human infectious disease is evaluation. There are many essential elements required for a thorough evaluation of the system(s) to be performed and for improvements to be made where necessary. The essential elements regarding the evaluation of syndromic surveillance systems described in the literature are: data quality metrics (sensitivity, specificity, representativeness and completeness, and timeliness), and system performance. Timeliness was discussed most frequently in the identified articles (21/55), as was specificity (12/55); the remaining essential elements were described in approximately 2-6 of the total articles (Table 2.2).

DISCUSSION

This scoping review presents an international perspective of the current state of syndromic surveillance systems used in the monitoring of human infectious disease. It synthesized existing published literature, analyzed the essential elements of such systems, and identified the most commonly described areas as well as areas in need of further exploration. This process revealed the development and implementation of syndromic surveillance systems were researched and

discussed the most frequently in the literature, and the majority of work was performed in the U.S. From this scoping review, knowledge strengths and gaps are identified for future research.

Most of the relevant articles (n=36) originated from the U.S. indicating the historical focus of syndromic surveillance in the U.S. Only nine other countries were represented in the review, with very few articles identified per country (between 1-7 articles per country). In addition, the papers were overwhelmingly published from higher income nations, with only research from China and Peru representing lower income countries. The limited geographic distribution may be a result of several factors, including the current state of research internationally and the inclusion criteria imposed in the review process (i.e. exclusion of non-English language articles). High-income countries may be more advanced in terms of research due to the availability of surveillance data and resource availability when compared to low and/or middle-income countries (1); in addition, the focus of human infectious disease syndromic surveillance will differ depending on the country, with some countries being more advanced than others on certain elements (17).

Most relevant documents in this review (43/55) were peer-reviewed journal articles. This large percentage of journal articles is likely because a significant proportion of articles on the topic were published by research groups who had written about specific syndromic surveillance systems they had created. This is in comparison to many of the review articles, which focussed on more general aspects of the essential elements of such systems. All four of the government reports used in the scoping review were reviews, as were three of the four book chapters and the single commentary.

There was great diversity worldwide in the most-heavily researched and reported elements in the definition, development, implementation, and evaluation of syndromic surveillance systems for human infectious disease. Several elements were well-represented, while others were less represented. These data demonstrate the progress in the field and where future research is required; these data also describe which areas of research appear to be the most pertinent and perhaps crucial to the creation of syndromic surveillance systems, therefore garnering the most attention.

The most common term used to refer to syndromic surveillance was "syndromic surveillance". This term is often debated given that not all syndromic surveillance systems examine syndromic data; rather, many systems utilize non-specific health indicator data sources (i.e. those which provide chief complaint information). Chief complaints as health indicators have become popular in the field of syndromic surveillance, and are now very frequently used as inputs to indicate the onset of disease for such systems. These indicators are used by many researchers and systems internationally (18,19). The various terms that are used to refer to syndromic surveillance require clarification to decide where particular terms are most appropriate. In addition, the purpose and desired outcome(s) must be predetermined when a system is first being created. Systems can have either specific or non-specific outcomes; selecting the outcome, in combination with defining the purpose of the system, is necessary for determining which datasets are the most appropriate and help in the construction of the surveillance system.

Most information found in the literature focussed on the development of new syndromic surveillance systems. This is likely a result of two overarching reasons. There are more individual elements comprising the development of syndromic surveillance when compared to the other three themes, giving rise to more research areas. In addition, some of the individual elements are complex and can be broken down further. Two examples of such complex elements in the development of syndromic surveillance include types of data sources and statistical methods; both elements are crucial to the creation of a syndromic surveillance system and can be broken down further into subcategories. Conversely, the process for defining syndromic surveillance was the least represented in literature. This was likely due to the fact there are existing definitions (13,11) that are routinely accepted, and fewer elements comprising this theme. These drawbacks will ideally give rise to future research in the definition of syndromic surveillance as its importance becomes more recognized and utilized.

Across all the four themes, the most commonly researched elements were data sources, operational challenges, timeliness, statistical methods, and the classification of syndromes. All are highly complex elements which can be subdivided into further subcategories. Data sources include pre-clinical (e.g. nurse telehealth lines) and clinical (e.g. emergency department chief complaints) pre-diagnostic sources. The classification of syndromes is the means by which data sources are classified into disease syndromes for analyses; this can be a complex process and needs to be performed meticulously to ensure all information is captured and properly organized to trigger alarms when disease is present (13,20). The number of syndrome categories will also impact the performance and effectiveness of the system, with some syndrome classifiers being broader or more detailed (21,22). The list of operational challenges is lengthy and includes

communication between stakeholders, interested parties, and all involved public health professionals, and also the various legal barriers involved in the implementation of such systems (13). In the development of an alert response protocol, there are many different ways data can be analyzed, and some techniques are more effective at achieving certain outcomes and weighing certain data than others; therefore, there is also a list of statistical analyses that must be considered. Lastly, the timeliness of syndromic surveillance systems is one of their most well-known benefits and is a metric which should be considered in evaluations.

Conversely, the least researched elements across all four themes of creating human infectious disease syndromic surveillance systems were data representativeness and completeness, overall system performance, determining the purpose and outcome(s) of systems, and management/staff composition. Describing overall system performance was a somewhat broad category; most articles which described the elements of the evaluation of syndromic surveillance systems did so with high granularity. Therefore, this element was often overlooked. Considering the representativeness and completeness of data is crucial to a system's success, however was not consistently described in the articles. This may have been due to the fact this was very specific and was sometimes not described explicitly or separately from data sources. Determining the purpose and outcome(s) of systems is a step which should be performed prior to the creation of every syndromic surveillance system to design and target it effectively. Henning (2004) describes the purpose and use of syndromic systems in great detail; however, additional work needs to be completed in this area to adequately cover the topic area and position public health professionals looking to employ syndromic surveillance systems in the future (5). This also holds true for the resource requirements from management/staff; the balance and time requirements of

the management and staff should be well described and determined prior to the creation of such systems and requires additional research.

A few limitations arose from this study. Articles may have been excluded because of the “enteric” term in the search string, and additional synonyms for "syndromic surveillance" which were not listed in the search string. Additional articles would likely have been captured if the review cut-off date was extended beyond 2004; however, this date was selected due to it being the year several pivotal articles were published and the desire to use the most up-to-date and relevant literature in this review.

CONCLUSION

In summary, this scoping review identified that not all elements involved in creating a syndromic surveillance system for human infectious disease have been researched and described equally in the global literature. Future work is required on the definition of such systems (i.e. outlining the importance of a system), and the evaluation.

REFERENCES

1. Chretien J-P, Burkom HS, Sedyaningsih ER, Larasati RP, Lescano AG, Mundaca CC, et al. Syndromic surveillance: Adapting innovations to developing settings. *PLoS Med.* 2008;5(3):367–72.
2. Public Health Ontario. *Syndromic Surveillance Discussion Paper*. Toronto, ON; 2012.
3. Forslund DW, Joyce EL, Burr T, Picard R, Wokoun D, Umland E, et al. Setting standards for improved syndromic surveillance. *IEEE Eng Med Biol Mag.* 2004;23(1):65–70.

4. Triple-S Project. Assessment of syndromic surveillance in Europe. *Lancet*. 2011;378(9806):1833–4.
5. Henning KJ. What is syndromic surveillance? *MMWR Suppl*. 2004;53(Sept):7–11.
6. Ziemann A, Rosenkötter N, Garcia-Castrillo Riesgo L, Fischer M, Krämer A, Lippert FK, et al. Meeting the International Health Regulations (2005) surveillance core capacity requirements at the subnational level in Europe: The added value of syndromic surveillance. *BMC Public Health*. 2015;15(1):107.
7. Berger M, Shiau R, Weintraub JM. Review of syndromic surveillance: Implications for waterborne disease detection. *J Epidemiol Community Health*. 2006;60:543–50.
8. Buehler JW, Whitney EA, Smith D, Prietula MJ, Stanton SH, Isakov AP. Situational uses of syndromic surveillance. *Biosecurity Bioterrorism Biodefense Strateg Pract Sci*. 2009;7(2):165–77.
9. Buckeridge DL. Outbreak detection through automated surveillance: A review of the determinants of detection. *J Biomed Inform*. 2007;40(4):370–9.
10. Chen D, Cunningham J, Moore KM, Tian J. Spatial and temporal aberration detection methods for disease outbreaks in syndromic surveillance systems. *Ann GIS*. 2011;17(4):211–20.
11. 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*. 2004;53:1–13.
12. Pavlin JA. Syndromic surveillance for infectious diseases. In: *Infectious Disease Surveillance: Second Edition*. 2nd ed. Blackwell Publishing Ltd; 2013. p. 470–81.
13. Triple-S Project. Guidelines for designing and implementing a syndromic surveillance

- system. 2013.
14. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: A practical guide informed by the early experience. *J Am Med Informatics Assoc.* 2004;11(2):141–50.
 15. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005 Feb 1;8(1):19–32.
 16. Stoto MA. Syndromic surveillance. *Issues Sci Technol.* 2005;21(3):49–56.
 17. Frerichs RR. Epidemiologic surveillance in developing countries. *Annu Rev Public Health.* 1991;12:257–80.
 18. Conway M, Dowling JN, Chapman WW. Using chief complaints for syndromic surveillance: A review of chief complaint based classifiers in North America. *J Biomed Inform.* 2013;46(4):734–43.
 19. Loveridge P, Cooper D, Elliot AJ, Harris J, Gray J, Large S, et al. Vomiting calls to NHS Direct provide an early warning of norovirus outbreaks in hospitals. *J Hosp Infect.* 2010 Apr;74:385–93.
 20. Haas SW, Travers D, Waller A, Mahalingam D, Crouch J, Schwartz TA, et al. Emergency Medical Text Classifier: New system improves processing and classification of triage notes. *Online J Public Health Inform.* 2014;6(2):e178.
 21. Jayatilleke A, Kriseman J, Bastin LH, Ajani U, Hicks P. Syndromic surveillance in an ICD-10 world. In: *AMIA Annual Symposium Proceedings.* United States; 2014. p. 1806–14.
 22. Yan P, Chen H, Zeng D. Syndromic surveillance systems. In: *Annual Review of Information Science and Technology.* University of Arizona; 2009. p. 425–95.

TABLES

Table 2.1. Descriptive characteristics of 55 selected articles describing the essential elements of human infectious disease syndromic surveillance systems

<i>Characteristic</i>	All selected articles		Primary research articles		Review articles	
	<i>Number</i>	<i>Percent (%)</i>	<i>Number</i>	<i>Percent (%)</i>	<i>Number</i>	<i>Percent (%)</i>
Publication date						
2004-2007	19	34.5	8	30.8	11	37.9
2008-2011	19	34.5	9	34.6	10	34.5
2012-2016	17	30.9	9	34.6	8	27.6
Document type						
Journal article	43	78.2	23	88.5	20	69.0
Book chapter	4	7.3	1	3.8	3	10.3
Conference proceeding	2	3.6	1	3.8	1	3.4
Government report	4	7.3	0	0	4	13.8
Thesis	1	1.8	1	3.8	0	0
Commentary	1	1.8	0	0	1	3.4
Study location¹						
United States	36	65.5	19	67.9	17	58.6
Canada	7	12.7	2	7.1	5	17.2
France	3	5.5	1	3.6	2	6.9
Italy	2	3.6	2	7.1	0	0
Netherlands	2	3.6	0	0	2	6.9
Germany	2	3.6	2	7.1	0	0
China	2	3.6	2	7.1	0	0
Other ²	3	5.5	0	0	3	10.3

¹Column percentages do not sum to 100% due to the fact some studies were conducted in more than one location

²Additional countries included Australia (n=1), Peru (1), and the United Kingdom (1).

Table 2.2. Characteristics of syndromic surveillance as described in the 55 selected articles used in the scoping review of human infectious disease syndromic surveillance systems

<i>Characteristic</i>	<i>Number</i>	<i>Percent (%)</i>
Does the article describe the uses of syndromic surveillance? (n=55)		
Yes	43	78.2
No	12	21.8
Use(s) of syndromic surveillance described in the article¹ (if “yes” above, n=43)		
Early warning	37	86.0
Detection	24	55.8
Situational awareness	22	51.1
How many articles describe the elements of the definition of syndromic surveillance? (n=55)		
Terminology	6	10.9
Determining purpose and outcome of system	3	5.5
How many articles describe the elements of the development of syndromic surveillance? (n=55)		
Data sources	27	49.1
Statistical methods	21	38.2
Classification of syndromes	19	34.5
Data structure	14	25.5
Development of alert response protocol	6	9.1
Stakeholder identification and participation	5	9.1
Management/staff composition	4	7.3
How many articles describe the elements of the implementation of syndromic surveillance? (n=55)		
Operational challenges	22	40.0
Integration with other public health surveillance tools	13	23.6
Costs and infrastructure	11	20.0
How many articles describe the elements of the evaluation of syndromic surveillance? (n=55)		
Data quality metrics:		
Timeliness	21	38.2
Specificity	12	21.8
Sensitivity	6	10.9
Representativeness and completeness	2	3.6
System performance	3	5.5

FIGURES

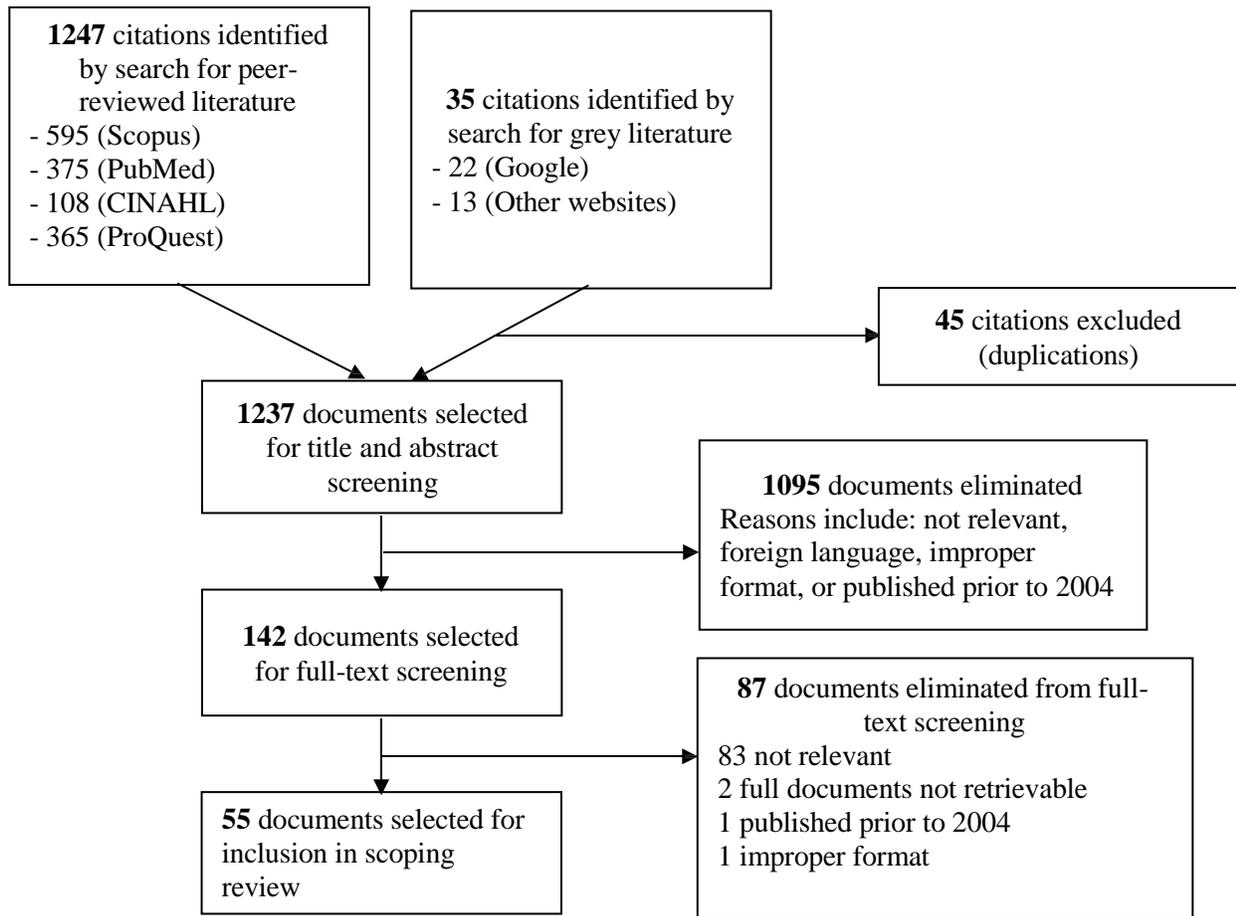


Figure 2.1. Document screening and selection process for scoping review of human infectious disease syndromic surveillance systems

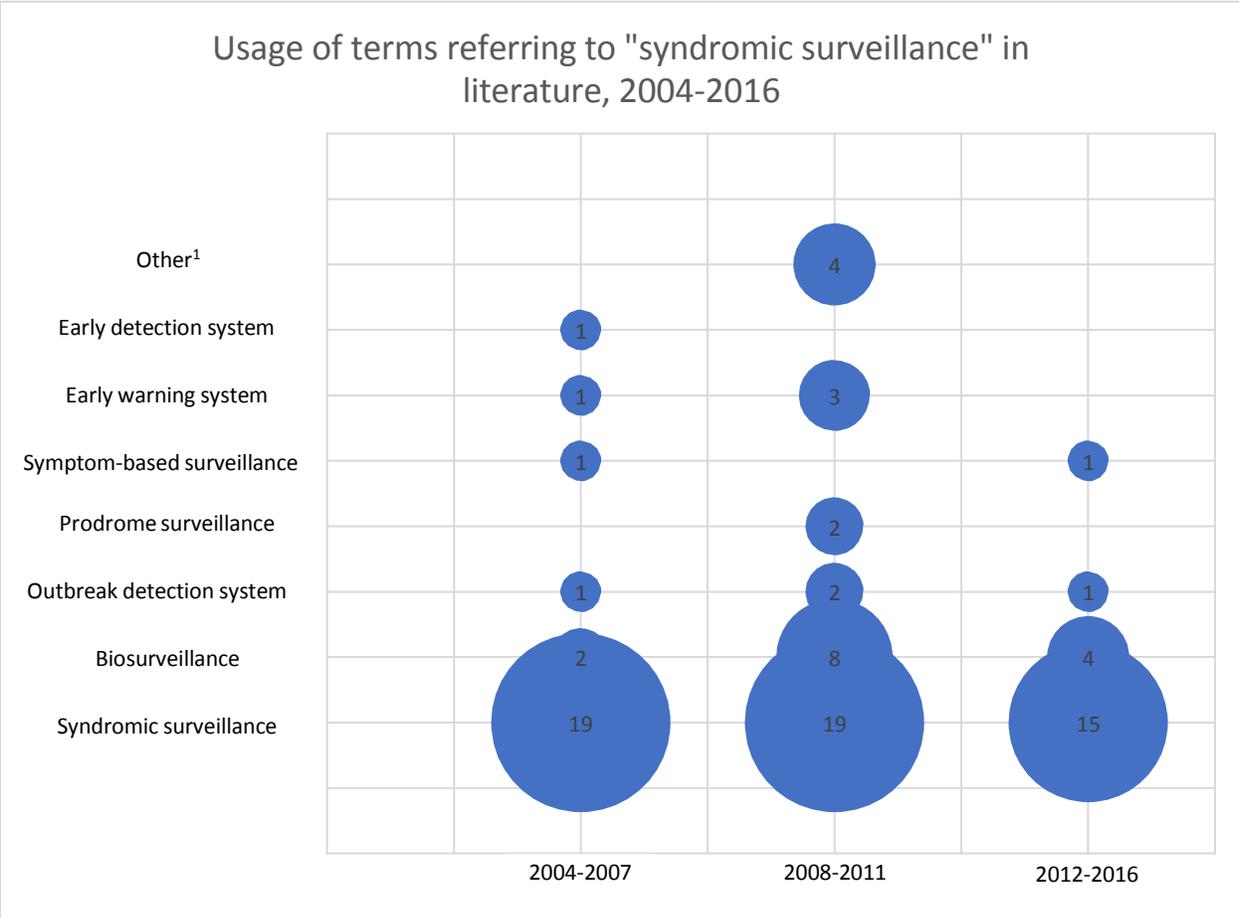


Figure 2.2. Analysis of the various terms used in the 55 documents to refer to “syndromic surveillance” and their frequency, by publication date.

¹Other terms used in articles included "health indicator surveillance," "enhanced surveillance," "information system-based sentinel surveillance," "population health surveillance," and "drop-in surveillance."

CHAPTER 3

Essential elements of human infectious disease syndromic surveillance systems: A narrative synthesis

ABSTRACT

Syndromic surveillance is a type of health surveillance that is rapidly expanding in public health practice for early detection of cases and outbreaks. Syndromic surveillance systems utilize existing data sources and can detect emerging diseases and more accurately track the progression of illness in a population when compared to traditional surveillance. This narrative synthesis set out to describe the essential elements in the definition, development, implementation, and evaluation of syndromic surveillance systems for human infectious disease.

A structured scoping review was performed prior to the writing of this synthesis, which identified 55 relevant articles. Key study characteristics and data were extracted from these 55 articles by two independent reviewers. This information was used as a foundation to conduct a narrative synthesis on the essential elements of syndromic surveillance systems.

Many essential elements of syndromic surveillance systems were identified. These elements included the definition of syndromic terminology, data structure & sources, developing alert response protocols, cost and infrastructure barriers, operational challenges, and various aspects of evaluation criteria necessary for syndromic surveillance system improvements.

Over the past decade, the role of syndromic surveillance in public health practice has grown immensely; these systems have now become an integral part of public health surveillance in

several countries. The results from this narrative synthesis demonstrate that various elements are required to effectively define, develop, implement, and evaluate these systems to ensure their optimal performance.

INTRODUCTION

Traditionally, public health surveillance has focussed on laboratory and outbreak reports. While effective at monitoring disease, substantial data gaps exist as a result of under-reporting and time lags [1,2]. Novel surveillance techniques and intervention strategies were required to address these gaps; hence, the field of syndromic surveillance emerged [3,4].

By definition, syndromic surveillance is referred to as the “real-time (or near-real-time) collection, analysis, interpretation and dissemination of health-related data to enable the early identification of the impact (or absence of impact) of potential human health... threats,” [5]. Such systems are effective at both detecting disease activity in populations and also observing when a health event(s) is not present in populations [1]. In doing this, public health and disease specialists can detect and respond to health threats earlier than would otherwise be possible with traditional surveillance methods [1,2,6,10,12–15].

Syndromic surveillance systems must be carefully developed to ensure their optimal performance. Therefore, a considerable amount of effort must go into the definition, development, implementation, and evaluation of these systems. While these themes have been described in several studies, information has been segmented [5,16,17]. A comprehensive scoping review was performed to locate relevant literature describing the essential elements of

human infectious disease syndromic surveillance systems [18]. This article reports a narrative synthesis of the articles (n=55) identified in that review. Narrative synthesis is an analytical technique used in scoping reviews and other knowledge syntheses that compiles and amalgamates findings and evidence into overall themes [19]. The approach aims to generate new insights and research questions for further studies. The objective of this synthesis was to combine the information outlining the definition, development, implementation, and evaluation of syndromic surveillance to inform future syndromic surveillance research and practice. It will also serve as an informative toolkit for public health professionals looking to develop such systems moving forward.

METHODS

Scoping review

A scoping review protocol was developed and performed prior to the development of this narrative synthesis using the research question, “What are the essential elements of a fully functional syndromic surveillance system for human infectious disease?” Eligible articles for the scoping review included review articles, primary research articles, government papers, research institute reports, conference proceedings, commentaries, theses, and book chapters. A total of 1237 unique citations were identified from the search; following relevance and full-text screenings, 55 articles were selected. Each step of the review was conducted and independently verified by two reviewers. A description of the scoping review process can be located in [18] and a complete list of the 55 articles identified can be seen in *Appendix 2.2*.

Data extraction

A data extraction form, consisting of 21 questions, was developed *a priori* (Appendix 2.1) and used to extract information from the 55 articles selected for inclusion (Appendix 2.2). The form was applied to all 55 articles independently and information was transcribed into a Microsoft Excel spreadsheet.

Narrative synthesis

Using information gathered from the data extraction, a narrative synthesis was then conducted, bringing together evidence from the relevant articles. Data were organized based on common themes, i.e. the essential elements of syndromic surveillance systems for human infectious disease. Four overarching themes of these elements were identified (definition, development, implementation, and evaluation of syndromic surveillance), as well as subcategories (Table 3.1). Information was structured into these subcategories based on categorizations presented in the literature and knowledge from familiarity with the subject matter.

RESULTS

Stage 1: Definition of syndromic surveillance

Prior to the development of a syndromic surveillance system, it is imperative that it is well-defined; its purpose must be stated, what information it will collect must be described, and the illnesses(s) targeted detailed [5]. What is meant by the term "syndromic surveillance" as well as the uses of syndromic surveillance should be thoroughly understood to ensure syndromic surveillance practices are appropriate [6,16].

Various other terminologies exist to describe such systems, however, "syndromic surveillance" is the dominant term in the field [1,5,6,10,16–33]. Similarly, syndromic surveillance systems can either be defined as having specific or non-specific outcomes of interest [16].

The purpose and proposed uses of the syndromic surveillance system must be clearly defined prior to its creation to determine how to effectively construct the surveillance system, target it, and potentially link it to other surveillance systems already in place [5,9]. Three overarching uses of syndromic surveillance are widely recognized: the early detection of a pre-established health event, the discovery of a previously unrecognized illness, and the routine monitoring of populations to confirm that there is a lack of disease activity/impact of active disease activities (situational awareness) [5,6,10,13,20,29,32,34–37].

Stage 2: Development of syndromic surveillance

2.1 Stakeholder identification & participation

Interested stakeholders are required for the initiation, operation, and evaluation of a syndromic surveillance system. Their expertise should cover a wide range of fields, and their continued collaboration and communication is essential for the success of the system. Parties include data providers, educational institutions, epidemiologists, funding bodies, physicians, nurses, infection control practitioners, geographic information system (GIS) experts, statisticians, public health professionals, and government [1,5,9,13,37].

2.2 Management/staff composition

An appropriate and knowledgeable management structure and staff team must be assembled to develop and operate a syndromic surveillance system, such as the team lead, data monitor, data analyst, administrative assistant, facility liaison, technical support, and epidemiologist [5,6,42,48].

2.3 Data coding structure

Available coded data may be available in different formats based upon the source of data and whether coded data are entered as pre-determined (closed) or open-ended entries. While the decision of which data coding structure to select may be determined solely by the available system, if alternatives are available then selecting to use closed entries (i.e. drop-down menus) may help keep analyses consistent; however, this limits the range of information which can be gathered when compared to open-ended free-text [13,40]. Although open-ended free-text (i.e. triage notes, sentence or point form notes recorded by nurses when triaging patients) allows for a more diverse array of information to be collected, there is often a lack of consistency and difficulty with analyses. Responses can vary, leaving staff to deal with editorial issues [13]. Additional decisions may be made on whether to use chief complaint entries (primary symptoms) or clinical diagnosis codes. Chief complaints are considered to be more timely as they are collected at the first available point of health care [16]. However, diagnoses are more information-rich as they are typically provided by medical professionals. The downside to diagnoses is that they can be left unassigned until weeks following medical care [13]. The International Classification of Diseases (ICD) and Health Level-7 (HL7) codes are frequently used for assigning diagnoses [1,12,13,24,26,30,32,40,46,50].

2.4 Data sources

An array of data sources are currently being employed in the field of syndromic surveillance. The majority of these sources are pre-existing prior to their inclusion in syndromic surveillance and can be exchanged in a timely manner electronically (16). Syndromic surveillance data sources can be categorized into one of two groups, pre-clinical pre-diagnostic data or clinical pre-diagnostic data (Table 3.2) [1–3,5,6,8,9,12,13,17,18,20–22,29,31,32,34,44,46–49,51,52].

2.5 Classification of syndromes

Once the sources of coded data are identified, they need to be grouped into disease syndromes for analyses, such as respiratory or gastrointestinal disease. The grouping can be performed using specific data structures and computer classifier programs, or manually using customized syndromes or indicators [26]. Automated programs can be structured as natural language processors, Bayesian classifiers, positive predictive value criteria, text string searchers, support vector machines, decision trees, or vocabulary abstractors [32]. Manually, a working group can map entries to pre-determined disease syndromes. However, not all data sources allow for this as it depends on the level of granularity; some data sources are non-specific and cannot be grouped [12,13,20,26,31,40,47,49,50,53–55].

2.6 Development of an alert response protocol

Activity thresholds must be devised prior to the implementation of a syndromic surveillance system. Thresholds play an important role in the activity of such systems, distinguishing normal disease activity from outbreak situations [37]. Alarm thresholds can also be developed to detect

geographical aberrations, demographic information, and/or concurrent alarms detected using multiple syndromic surveillance systems or across various surveillance systems [6].

A response protocol for public health professionals following the triggering of an alarm is also required [51]. When necessary, detailed outbreak investigations are carried out [34].

2.6.1 Statistical methods

Statistical methods used to analyze syndromic surveillance data must be considered to determine if public health action is required. Various statistical algorithms have been devised to accomplish such tasks with differing degrees of accuracy and success based on temporal, spatial, or spatial-temporal analyses [2,5,8,12,13,18,19,23,24,27,28,30–32,41–43,45–48].

i) Temporal analyses

Temporal data analysis techniques are those which assess disease anomalies over time. These analyses often incorporate historical data and observed/expected mean values. Popular examples of temporal detection algorithms include cumulative sums (CUSUMS), recursive least squares (RLS), and exponentially weighted moving average (EWMA) (32).

ii) Spatial analyses

Spatial data analysis techniques are those which identify disease clusters and cases by their geographical distribution. Spatial algorithms can also track the spread of outbreaks and identify vulnerable populations. Such techniques often incorporate observed/expected case counts in space, and spatial characteristics such as ZIP/post code and health unit. Popular examples of

spatial detection algorithms include the Generalized Linear Mixed Modeling (GLMM) and Small Area Regression and Testing (SMART) (32).

iii) Spatial-temporal analyses

Spatial-temporal data analysis techniques are those which incorporate both spatial and temporal analyses. These algorithms assess disease anomalies and observed/expected values in both space and time to determine whether an outbreak is occurring. Popular examples of spatial-temporal algorithms include What's Strange About Recent Events (WSARE) and Population-wide ANomaly Detection and Assessment (PANDA) (32).

Step 3: Implementation of syndromic surveillance

3.1 Costs and infrastructure

Various costs and resources are required to realize a syndromic surveillance system, which will be entirely dependent on local circumstances. Costs can be organized into several categories: initial set up, implementation, general operating, improvement, and per-event costs [5,39].

Staff are obligatory for the setting up and maintenance of systems; they are imperative for the daily monitoring of alerts and carrying out of responses [5,13,38]. Depending on the sources and situation, there may be costs involved with accessing data for syndromic surveillance [5,13].

Lastly, additional IT resources to power the underlying system are needed [2,5,13,36,38,48].

3.2 Operational challenges

Communication is key during the implementation and continued performance of a syndromic surveillance system [5]. This includes routine (daily) communication between the public health practitioners co-ordinating the system and stakeholders and other interested parties to notify of activity alerts, and ad-hoc communication during incidents or events [5,12]. Communication is also important when drafting data sharing agreements between public health professionals and data providers [5]. The method with which these results are disseminated vary; the most appropriate means should be selected based on the recipients and desired outcome. Methods include person-to-person contact, surveillance bulletins, websites, telephone, social media, and e-mail [1,5].

Obtaining the legal permissions to access and use health (or other sources) data for syndromic surveillance is a key step. Approval by local or national governance and ethics groups and any legal business contracts outlining data sharing and use need to be resolved [5,57]. The use of anonymized aggregated data can circumvent some legal and governance issues, but explicit permission must be sought from both patients and data providers [13].

3.3 Integration with other public health surveillance tools

Surveillance systems must be integrated within existing public health programs; syndromic surveillance systems complement and strengthen traditional surveillance by offering additional information on infectious disease activity and seasonal disease events [2,6,7,16,21,26,29,31,32,47].

Step 4: Evaluation of syndromic surveillance

The evaluation of a syndromic surveillance system is crucial to the long-term success of syndromic surveillance systems and improvements. There is existing literature and guidance on evaluating public health surveillance systems [9]. The system should be assessed using data quality metrics such as sensitivity, specificity, representativeness, and timeliness, as well as overall performance [5,12,13,34,36]. Evaluation determines the usefulness of a system and can highlight areas that require development and/or improvement.

DISCUSSION

There are many components necessary for the definition, development, implementation, and evaluation of syndromic surveillance systems for human infectious disease. In the articles analyzed in this narrative synthesis, various key elements were identified and summarized for future consideration by researchers and public health professionals.

Syndromic surveillance systems must be developed with a clear goal in mind, clearly stating its intents and purposes. Fricker et al. (2008) and Uscher-Pines et al. (2009) discuss the developments and new directions of syndromic surveillance since its inception; early detection, bioterrorism, and situational awareness are all highlighted as main purposes and uses of such systems [37]. In addition, the Triple-S Project (2013) stresses the need for defining a system's purpose to build the system [5,19].

Interested stakeholders are required for the initial discussions surrounding the syndromic surveillance system, as well as maintenance and continuous improvement. They are drawn from a wide variety of expertise, which adds knowledge and value from multiple perspectives to the system being created. For example, a CDC-led working group helps map clinical patient visit information to syndrome groups in the National Bioterrorism Syndromic Surveillance Demonstration Program in the United States [48]. A team of international research partners working with BioPortal™ collaborate to deliver an online information-sharing and analysis platform with multilingual capabilities [32]. The staff are necessary for the development and daily operation of the system, as well as alert response. Both the coding structure and sources of data must be pre-determined since these data act as the core of the system. Lastly, an alert response protocol should be developed prior to the implementation of the system; such a protocol involves analyzing the data using one or several statistical methods to rule in or out a public health threat.

In the implementation of syndromic surveillance, the essential elements focussed on the setup of the system and the operational requirements needed for it to be initiated and function. A key function is the requirement of financial resources. Setup and general operating costs need to be considered: staff resources play an important role throughout the implementation. Physical resources play an important role in terms of data analysis and storage and are critical components of syndromic surveillance systems. Funding is also sometimes required to form agreements between data providers and those who wish to use the data for syndromic purposes. The availability to form agreements between data providers and syndromic users may pose barriers to

implementation and are separate from monetary barriers. Privacy, governance, and ethical issues can come in to play during implementation and potentially prevent the sharing of data.

Ding et al. (2015) measured the costs of various stages of syndromic surveillance systems. Some of the most notable costs described include staff salaries, computers and other technology equipment, training, and quality control measures. They also make note of the operating costs of data collection in some facilities are higher than others, e.g. data collection from pharmacies tends to cost more to run in comparison to primary schools [38,39]. In addition, Kirkwood et al. (2007) makes note that although cost can be a barrier in setting up and operating a syndromic surveillance system, these costs typically account for a small percentage of overall surveillance costs [56].

Stakeholders, public health practitioners, and the syndromic surveillance system staff must be in frequent contact to discuss current findings and next steps. Once implemented, public health practitioners must be informed of significant aberrations in the syndromic surveillance data so that they may investigate possible public health threats and initiate action. Communication with other public health experts is important because syndromic surveillance best functions in tandem with other surveillance methods. Ziemann et al. (2015) stress the importance of collaborations between public health authorities and health care institutions (who often serve as data suppliers). Such collaborations help strengthen data quality, interpretation, and use in syndromic surveillance systems and better prepare systems for crisis management (7).

Following implementation, the evaluation of syndromic surveillance systems is necessary to ensure their long-term usefulness and success. Existing evaluation guidelines highlight the importance of reviewing: the purpose and operation of systems; the ability of the system to detect outbreaks, including timeliness and data quality; the system experience, including usefulness, flexibility, stability, and cost (9).

As an example, Ding et al. (2015) describe the importance of evaluating whether a system has met its stated purpose(s) and objective(s) and how well these have been met [39]. They also comment on the need to evaluate the cost-effectiveness of a system, an often-overlooked attribute. More data sources can be integrated into systems to improve performance and monitor more disease syndromes, and richer more timely data may be incorporated.

The limitations of this narrative review include the search being limited to English language documents and publication years 2004-2016. The former criterion may have, in turn, limited the geographic range of articles selected for inclusion, and the latter would have excluded any older publications that may have contained additional insights; however, this date range was chosen because many pivotal articles were published in 2004 and the most up-to-date literature was preferred. Moreover, some articles may not have been identified and included in the synthesis because of additional synonyms for "syndromic surveillance" not being included in the scoping review search string. Lastly, the steps outlined in this synthesis should be interpreted as a general guide and should not be interpreted for each system in its literal form. The outline presented should be adjusted and adapted as necessary.

CONCLUSION

Syndromic surveillance is a rapidly expanding field of public health due to its unique ability to detect adverse health events earlier than traditional methods of surveillance. Their composition is complex and should be carefully considered. The definition, development, implementation, and evaluation of syndromic surveillance systems for human infectious disease are comprised of many key elements which should be considered by public health professionals when creating or adopting such systems. This narrative synthesis will serve as a toolkit for professionals in the field looking to define, develop, implement, and evaluate syndromic surveillance systems moving forward.

REFERENCES

1. Buehler JW, Whitney EA, Smith D, Prietula MJ, Stanton SH, Isakov AP. Situational uses of syndromic surveillance. *Biosecurity Bioterrorism Biodefense Strateg Pract Sci*. 2009;7(2):165–77.
2. Berger M, Shiau R, Weintraub JM. Review of syndromic surveillance: Implications for waterborne disease detection. *J Epidemiol Community Health*. 2006;60:543–50.
3. Public Health Ontario. *Syndromic Surveillance Discussion Paper*. Toronto, ON; 2012.
4. Forslund DW, Joyce EL, Burr T, Picard R, Wokoun D, Umland E, et al. Setting standards for improved syndromic surveillance. *IEEE Eng Med Biol Mag*. 2004;23(1):65–70.
5. Triple-S Project. *Guidelines for designing and implementing a syndromic surveillance system*. 2013.
6. Henning KJ. What is syndromic surveillance? *MMWR Suppl*. 2004;53(Sept):7–11.
7. Ziemann A, Rosenkötter N, Garcia-Castrillo Riesgo L, Fischer M, Krämer A, Lippert FK,

- et al. Meeting the International Health Regulations (2005) surveillance core capacity requirements at the subnational level in Europe: The added value of syndromic surveillance. *BMC Public Health*. 2015;15(1):107.
8. Chen D, Cunningham J, Moore KM, Tian J. Spatial and temporal aberration detection methods for disease outbreaks in syndromic surveillance systems. *Ann GIS*. 2011;17(4):211–20.
 9. 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*. 2004;53:1–13.
 10. Chretien J-P, Burkom HS, Sedyaningsih ER, Larasati RP, Lescano AG, Mundaca CC, et al. Syndromic surveillance: Adapting innovations to developing settings. *PLoS Med*. 2008;5(3):367–72.
 11. Buckeridge DL. Outbreak detection through automated surveillance: A review of the determinants of detection. *J Biomed Inform*. 2007;40(4):370–9.
 12. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: A practical guide informed by the early experience. *J Am Med Informatics Assoc*. 2004;11(2):141–50.
 13. Pavlin JA. Syndromic surveillance for infectious diseases. In: *Infectious Disease Surveillance: Second Edition*. 2nd ed. Armed Forces Health Surveillance Center, Silver Spring, MD, United States: Blackwell Publishing Ltd; 2013. p. 470–81.
 14. Hughes SL, Elliot AJ, McEwen SA, Greer AL, Young I, Papadopoulos A. Key elements of infectious disease syndromic surveillance systems : A scoping review. 2017;9(1):2579.
 15. Mays N, Pope C, Popay J. Systematically reviewing qualitative and quantitative evidence

- to inform management and policy-making in the health field. *J Health Serv Res Policy*. 2005;10(suppl 1):6–20.
16. Katz R, May L, Baker J, Test E. Redefining syndromic surveillance. *J Epidemiol Glob Health*. 2011 Dec;1(1):21–31.
 17. Aller RD. Pathology's contributions to disease surveillance: Sending our data to public health officials and encouraging our clinical colleagues to do so. *Arch Pathol Lab Med*. 2009;133(6):926–32.
 18. 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(May):18–22.
 19. Fricker RDJ, Hegler BL, Dunfee DA. Comparing syndromic surveillance selection methods: EARS versus a CUSUM-based methodology. *Stat Med*. 2008;27(17):3407–29.
 20. Generous N, Margevicius KJ, Taylor-McCabe KJ, Brown M, Daniel WB, Castro L, et al. Selecting essential information for biosurveillance - A multi-criteria decision analysis. *PLoS One*. 2014;9(1).
 21. Hartley DM, Nelson NP, Arthur RR, Barboza P, Collier N, Lightfoot N, et al. An overview of internet biosurveillance. *Clin Microbiol Infect*. 2013;19(11):1006–13.
 22. Kaydos-Daniels SC, Smith LR, Farris TR. Biosurveillance in outbreak investigations. *Biosecurity and Bioterrorism*. 2013;11(1):20–8.
 23. Najmi A-H, Burkom H. Recursive least squares background prediction of univariate syndromic surveillance data. *BMC Med Inform Decis Mak*. 2009;9(1):4.
 24. Shmueli G, Burkom H. Statistical Challenges Facing Early Outbreak Detection in Biosurveillance. *Technometrics*. 2010;52(1):39–51.

25. Sintchenko V, Gallego B. Laboratory-guided detection of disease outbreaks: Three generations of surveillance systems. *Arch Pathol Lab Med* [Internet]. 2009;133(6):916–25. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-67049095736&partnerID=40&md5=2fa83a2479008fe135796b4268394ca4>
26. Sniegowski CA. Automated syndromic classification of chief complaint records. *Johns Hopkins APL Tech Dig*. 2004;25(1):68–75.
27. Szarka III JL, Gan L, Woodall WH. Comparison of the early aberration reporting system (EARS) W2 methods to an adaptive threshold method. *Stat Med*. 2011;30:489–504.
28. Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the prospective detection of infectious disease outbreaks: A review. *J R Stat Soc*. 2012;175(1):49–82.
29. 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. *Eurosurveillance*. 2011;16(9):1–8.
30. VanStone N, Wu C-C, Van-Dijk A, Belanger P, Moore KM. Syndromic surveillance report: Validation of the EDSS system’s categorization of syndromes. 2015.
31. Yan P, Zeng D, Chen H. A review of public health syndromic surveillance systems. In: S. Mehrotra et al, editor. *Lecture Notes in Computer Science*. University of Arizona; 2006. p. 249–60.
32. Yan P, Chen H, Zeng D. Syndromic surveillance systems. In: *Annual Review of Information Science and Technology*. University of Arizona; 2009. p. 425–95.
33. Mostashari F, Hartman J. Syndromic surveillance: a local perspective. *J Urban Health*. 2003;80(2 Suppl 1):i1-7.

34. Hurt-Mullen KJ, Coberly J. Syndromic surveillance on the epidemiologist's desktop: making sense of much data. *MMWR Morb Mortal Wkly Rep.* 2005;54 Suppl:141–6.
35. Katz R, May L, Baker J, Test E. Redefining syndromic surveillance. *J Epidemiol Glob Health.* 2011;1(1):21–31.
36. Stoto MA. Syndromic surveillance. *Issues Sci Technol.* 2005;21(3):49–56.
37. Uscher-Pines L, Farrell C, Cattani J, Moskal MD, Babin SM, Charlotte A, et al. Framework for the Development of Response Protocols for Public Health Syndromic Surveillance Systems. 2008;53:2008.
38. Yih WK, Caldwell B, Harmon R, Kleinman K, Lazarus R, Nelson A, et al. National Bioterrorism Syndromic Surveillance Demonstration Program. *Morb Mortal Wkly Rep.* 2004;53(Supplement):43–9.
39. Lawson BM, Fitzhugh EC, Hall SP, Franklin C, Hutwagner LC, Seeman GM, et al. Multifaceted syndromic surveillance in a public health department using the Early Aberration Reporting System. *J Public Heal Manag Pract.* 2005 Jul;11(4):274–281 8p.
40. Haas SW, Travers D, Waller A, Mahalingam D, Crouch J, Schwartz TA, et al. Emergency Medical Text Classifier: New system improves processing and classification of triage notes. *Online J Public Health Inform.* 2014;6(2):e178.
41. Conway M, Dowling JN, Chapman WW. Using chief complaints for syndromic surveillance: A review of chief complaint based classifiers in North America. *J Biomed Inform.* 2013;46(4):734–43.
42. Tsui K-L, Chiu W, Gierlich P, Goldsman D, Liu X, Maschek T. A review of healthcare, public health, and syndromic surveillance. *Qual Eng.* 2008;20(4):435–50.
43. Kirian ML, Weintraub JM. Is syndromic surveillance of OTC drug sales effective in

- detecting outbreaks of waterborne GI disease? *Am Water Work Assoc.* 2011;103(8):63–73.
44. Moore K. Real-time syndrome surveillance in Ontario, Canada: the potential use of emergency departments and Telehealth. *Eur J Emerg Med.* 2004;11(1):3–11.
 45. Triple-S Project. About data sources for syndromic surveillance. 2013.
 46. Velardi P, Stilo G, Tozzi AE, Gesualdo F. Twitter mining for fine-grained syndromic surveillance. *Artif Intell Med.* 2014;61:153–63.
 47. Zhang Y, Dang Y, Chen H, Thurmond M, Larson C. Automatic online news monitoring and classification for syndromic surveillance. *Decis Support Syst.* 2009;47:508–17.
 48. Chapman WW, Christensen LM, Wagner MM, Haug PJ, Ivanov O, Dowling JN, et al. Classifying free-text triage chief complaints into syndromic categories with natural language processing. *Artif Intell Med.* 2004;33(1):31–40.
 49. Gerbier S, Yarovaya O, Gicquel Q, Millet A-L, Smaldore V, Pagliaroli V, et al. Evaluation of natural language processing from emergency department computerized medical records for intra-hospital syndromic surveillance. *BMC Med Inform Decis Mak.* 2011;11(1).
 50. Magruder SF, Henry J, Snyder M. Linked analysis for definition of nurse advice line syndrome groups, and comparison to encounters. *MMWR Morb Mortal Wkly Rep.* 2005;54 Suppl:93–7.
 51. Smith GE, Elliot AJ, Ibbotson S, Morbey R, Edeghere O, Hawker J, et al. Novel public health risk assessment process developed to support syndromic surveillance for the 2012 Olympic and Paralympic Games. *J Public Health (Bangkok)* [Internet]. 2016;39(3):e111-117. Available from: <https://academic.oup.com/jpubhealth/article->

lookup/doi/10.1093/pubmed/fdw054

52. Henriques G, Stacey D. An ontology-based framework for syndromic surveillance method selection. *4th Int Conf Knowl Eng Ontol Dev.* 2012;396–400.
53. Lescano AG, Larasati RP, Sedyaningsih ER, Bounlu K, Araujo-Castillo R V, Munayco-Escate C V, et al. Statistical analyses in disease surveillance systems. *BMC Proc.* 2008;2(Suppl 3):S7.
54. Stoto MA, Fricker RD, Jain A, Diamond A, Davies-Cole JO, Glymph C, et al. Evaluating statistical methods for syndromic surveillance. In: *Statistical Methods in Counterterrorism: Game Theory, Modeling, Syndromic Surveillance, and Biometric Authentication.* 2006. p. 141–72.
55. Ding Y, Sauerborn R, Xu B, Shaofa N, Yan W, Diwan VK, et al. A cost-effectiveness analysis of three components of a syndromic surveillance system for the early warning of epidemics in rural China health policies, systems and management in low and middle-income countries. *BMC Public Health.* 2015;15(1).
56. Ding Y, Fei Y, Xu B, Yang J, Yan W, Diwan VK, et al. Measuring costs of data collection at village clinics by village doctors for a syndromic surveillance system-a cross sectional survey from China. *BMC Health Serv Res.* 2015;15(1):287.
57. Hu J. *Privacy-Preserving Data Integration in Public Health Surveillance.* ProQuest Dissertations and Theses. [Ann Arbor]: University of Ottawa (Canada); 2011.
58. Kirkwood A, Guenther E, Fleischauer AT, Gunn J, Hutwagner L, Barry MA. Direct cost associated with the development and implementation of a local syndromic surveillance system. *J Public Heal Manag Pract.* 2007;13(2):194–9.

TABLES

Table 3.1. List of all essential elements of syndromic surveillance systems for human infectious disease, organized by category and associated subcategories.

Category	Associated subcategories
Definition	Terminology
	Determining purpose and outcome of system
Development	Data sources
	Statistical methods
	Classification of syndromes
	Data coding structure
	Development of alert response protocol
	Stakeholder identification and participation
	Management/staff composition
Implementation	Operational challenges
	Collaboration with other public health surveillance
	Costs and infrastructure
Evaluation	Data quality metrics:
	Timeliness
	Specificity
	Sensitivity
	Representativeness and completeness
	Overall system performance

Table 3.2. List of data sources utilized in syndromic surveillance systems.

a) Syndromic surveillance pre-clinical pre-diagnostic data sources

Pre-clinical pre-diagnostic data
Absenteeism (school, work)
Calls to nurse-operated telephone help lines
Poison control call centres
Over-the-counter pharmacy purchases
Prescription sales
Sewage water pathogen testing
Social media posts
Internet search queries
Voluntary surveys
Insurance claims

b) Syndromic surveillance clinical pre-diagnostic data sources

Clinical pre-diagnostic data
Emergency department chief complaints
Emergency department discharge diagnoses
Calls to emergency telephone help lines
Visits to general practitioners

CHAPTER 4

Estimating the burden of norovirus gastroenteritis in Ontario, Canada -- 2009-2014

ABSTRACT

Norovirus is well established as the most common cause of acute gastroenteritis in Canada and most developed countries worldwide. The illness has high morbidity and also high societal costs. While work has been performed on the burden of norovirus and norovirus-like illnesses in several countries, little research is available in the province of Ontario, Canada. Such studies are necessary to help direct public health promotion campaigns and reduce the economic burden of disease in this region.

Three datasets were acquired from the government of Ontario: two traditional surveillance datasets (outbreak and laboratory), as well as syndromic data (Telehealth), all spanning Jan 2009 – Dec 2014. All outbreaks, laboratory submissions, and calls to Telehealth were first assessed for total viral gastroenteritis (VGE). Norovirus and norovirus-like illness totals were then calculated as a proportion of total VGE to estimate the agent-specific burden of illness. Affected institution types, genders, and age groups were also analyzed.

Between 2009-2014, 41.5% of VGE outbreaks, 63.4% of VGE laboratory submissions, and 36.6% of all acute gastroenteritis-related (not restricted to viral causes) calls to Telehealth were attributed to norovirus and norovirus-like illness in Ontario. The most commonly affected institution type was long-term care homes, the most commonly affected age groups included younger (<5) and older (≥ 65) individuals, and females were only slightly more affected than males.

Norovirus and norovirus-like illnesses were the leading cause of VGE in Ontario between 2009-2014. They comprised the greatest percentage of VGE when compared to all other VGE-associated viruses. Additional work is needed to determine all component costs and necessary public health actions to reduce the burden of disease.

INTRODUCTION

Norovirus causes a significant level of illness worldwide. It is characterized by the sudden onset of nausea, vomiting, diarrhea, abdominal cramps, and malaise and is transmitted via the fecal-oral route and aerosolized vomitus (1,2). Vomiting is moderately more prevalent than diarrhea in adults; whereas children more commonly present with diarrhea as the predominant symptom (3,4). The incubation period is short (approximately 10-48h) and symptoms typically clear in 1-3 days; however, this is often longer in higher-risk individuals, such as the very young and elderly (1,2,5). Sequelae and serious side-effects, such as irritable bowel syndrome, necrotizing enterocolitis, or death are rare but possible (1,6).

Norovirus has been identified as the most common cause of infectious gastroenteritis in Ontario, Canada and worldwide (1-3). It comprises roughly 50% of acute gastroenteritis (AGE) (all aetiologies), and >90% of viral gastroenteritis (VGE) worldwide (4). A global systematic review performed in 2014 demonstrated that the overall prevalence of norovirus in patients with acute gastroenteritis was 18% (5). Its high morbidity rate is due to its low infectious dose (roughly 18-1000 viral particles), multiple transmission routes, extended viral shedding period, short immunity, and persistence in the environment (6,7). The burden of disease is high, with an estimated 3.4 million annual cases in Canada and 21 million cases in the United States (8,9).

Costs are estimated to be 21 million CAD per year in Canada and 500 million USD per year in the United States from hospital expenditures (1,10).

Surveillance of norovirus is necessary to help inform public health, develop appropriate messages and intervention measures, and, as a result, help decrease the burden of illness.

However, a lack of formal reporting mechanisms for norovirus (and AGE in general) creates data gaps; research has shown that only approximately 15% of individuals suffering from AGE seek medical care, and of those, diagnostic samples are only requested from 13% (11). Therefore, in the absence of complete data, we have attempted to describe the burden of VGE and specifically norovirus in the province of Ontario, Canada from the following sources: 1) outbreak data from Public Health Ontario; 2) laboratory data from Public Health Ontario; and 3) syndromic surveillance (telehealth) data from the Ontario Ministry of Health and Long-Term Care.

METHODS

Datasets

The integrated Public Health Information System (iPHIS) dataset represents confirmed outbreaks of VGE in institutions in Ontario reported to local public health units from 2009 to 2014. Such institutions include long-term care homes, retirement homes, daycares, and correctional facilities.

The Public Health Ontario Laboratories (PHOL) dataset represents all samples submitted to Public Health Ontario (PHO) from 2009 to 2014 with suspected VGE for confirmatory testing; more specifically, this dataset contains all samples sent to PHO with suspected norovirus or rotavirus infection.

The Telehealth Ontario (THO) dataset represents all calls made to the provincial nurse telehealth service with gastrointestinal symptoms as the chief complaint from 2009 to 2014. These gastrointestinal calls represent a collection of AGE symptoms, encompassing a broader scope than just VGE calls, captured by the nurses at THO. Callers may be ill with these gastrointestinal symptoms for a myriad of reasons including norovirus; therefore, calls with selected chief complaints (vomiting and diarrhea) are referred to as the “relevant chief complaints” and were used as a proxy for norovirus in this study. These relevant chief complaints were chosen due to their alignment with the main symptoms of norovirus illness. The breakdown of all AGE chief complaints collected by THO and those considered “relevant chief complaints” in this study are described in Table 4.1.

For more detailed descriptions of all three datasets, please see *Appendix 4.1*.

Data analyses

All data acquired were anonymized (no personal identifiers). All three datasets in this study were analyzed for total VGE outbreaks and the proportion attributed to norovirus (or in the case of THO data, gastroenteritis illness due to the inability to confirm presence of norovirus). These percentages were used to estimate the burden of norovirus in Ontario. The total number of each type of institution analyzed, and Census data of the population of Ontario, were used as denominator data to standardize particular analyses.

Descriptive analyses were performed on the datasets to examine institution type, virus type, age, and time using SAS v.9.4 (Cary, NC, USA) and Microsoft Office (Excel) 2010 (Redmond, WA, USA). All graphs were constructed using Microsoft Excel 2010.

RESULTS

i) iPHIS

There was a total of 3,100 VGE outbreaks in Ontario during the years 2009-2014, inclusive; 1,291 were determined to be norovirus, either by case definition (12) and/or laboratory confirmation (1,279). Therefore 41.6% (1,291/3,100) of all VGE outbreaks were attributed to norovirus, and the remaining 58.4% (1,809/3,100) were attributed to adenoviruses, astroviruses, enteroviruses/echoviruses, rotavirus, other caliciviruses, and gastroenteritis unspecified/other. During the period of 2009 and 2014, there were 1,398 VGE outbreaks in long-term care homes, 958 in child care facilities, 700 in retirement homes, eight in correctional facilities, and 36 in other settings. This order of affected institution type remained relatively consistent when only norovirus outbreaks were analyzed, in which case retirement homes replaced child care facilities as the second most frequently affected institution type. Between 2009 and 2014, there were 800 norovirus outbreaks in long-term care homes, 340 in retirement homes, 138 in child care facilities, 4 in correctional facilities, and 9 other. This demonstrates most (57.2%, 800/1,398) of the VGE outbreaks in long-term care homes can be attributed to norovirus. The results for institutional breakdown for both VGE and norovirus-only outbreaks are shown in Figure 4.1. The year assigned to outbreaks is based on the onset dates of the index case of each outbreak.

The number of norovirus outbreaks per year was as follows over the six years; 211 in 2009, 265 in 2010, 178 in 2011, 247 in 2012, 215 in 2013, and 175 in 2014. This is comparable to the stability of VGE outbreaks across the same period: 551 in 2009, 629 in 2010, 482 in 2011, 539 in 2012, 492 in 2013, and 407 in 2014. The seasonal distribution of norovirus outbreaks by month and year is shown in Figure 4.2, with a rise in activity in the winter of 2009-2010. The average duration of VGE outbreaks was 12.6 days (range 1-78), and for norovirus outbreaks was 14.1 days (range 1-52 days).

ii) PHOL

There was a total of 29,459 submitted samples for rotavirus and norovirus-like VGE to PHO between 2009-2014, inclusive. The majority of these samples were negative, with 75.2% (22,147/29,459) samples showing no virus detected. When removing these negative samples and only analyzing the positive submissions (7,312), these results indicate that 63.4% (4,633/7,312) of the VGE-positive samples could be attributed to norovirus, with a total of 4,633 confirmed to be norovirus either by PCR, EM, or ICT (Figure 4.3). The remaining 36.6% (2,679/7,312) were composed of various other VGE aetiologies: adenovirus, astrovirus, other caliciviruses, picornaviruses, rotavirus, and/or sapovirus. The second most commonly detected virus was rotavirus, with 2,170 positive samples (29.7%, 2,170/7,312) (Figure 4.3).

Out of all VGE-positive submissions (7,312), females were affected more than males, with 2,928 submissions compared to 2,366, respectively; however, a large number of unknown patient sex samples were present (2,018). When analyzing only norovirus-positive samples (4,633), this pattern was repeated with 1,822 female submissions and 1,006 male submissions (1,805

unknown). The 65+ age group saw the highest number of VGE-positive submissions, with 2,546 in total. Second to this was the 0-4 age group with 2,426. Of this, 2,380 (93.5%) and 371 (15.3%) were attributed to norovirus, respectively (Figure 4.4).

A total of 62.3% (4,559/7,312) of the VGE-positive submitted samples were determined to be outbreak-related. In terms of the outbreak locations, there were many missing observations in the dataset (3,850/7,312 positive VGE submissions). However, once all missing entries were removed and the remaining 3,462 were analyzed, long-term care homes were the most commonly affected location type (1,824, 52.7%), followed by hospitals (495, 14.3%), retirement homes (441, 12.7%), day cares (343, 9.9%), restaurants (78, 2.3%), gatherings (52, 1.5%), schools (39, 1.1%), food supplies (37, 1.1%), camps (18, 0.5%), psychiatric hospitals (13, 0.4%), correctional facilities (12, 0.3%), clinics (6, 0.2%), and other (104, 3.0%). The seasonality of norovirus positive laboratory submissions broken down by month and year closely mimics the seasonality of norovirus outbreaks (Figure 4.2), also showing a rise in activity in the winter of 2009-2010.

iii) THO

A total of 320,834 telehealth calls were recorded for AGE illness in the period 2009-2014. There were 117,275 calls with vomiting (86,000), vomiting without diarrhea (22,218), or vomiting with diarrhea (9,057) as the chief complaint. These three symptoms are hereafter collectively referred to as the “relevant chief complaints” based on the fact vomiting is the predominant symptom of norovirus. This demonstrates 36.6% (117,275/320,834) of calls were due to gastroenteritis-

related illness. The percentage of AGE calls attributed to the relevant chief complaints fluctuated between 31-41% during 2009-2014 (Figure 4.5).

Callers most often phoned in to THO on behalf of a female more often than males, with 200,774 AGE calls; calls regarding males comprised 115,231 calls, and 397 unknown. (From the total number of calls, 4,432 were excluded due to data entry errors, i.e. missing observations or entries other than “female,” “male,” or “unknown.”) When analyzing the relevant chief complaints, this pattern was repeated with females comprising 66,044 calls in comparison to 49,528 calls from males and 162 unknown. A subset of 1,541 calls were excluded due to data entry errors.

Out of the calls made to THO, the 15-44 age range was the focus of the call (the “patient”) for the highest number of AGE calls to Telehealth Ontario with 131,271 calls between 2009-2014; this was not the case when analyzing the relevant chief complaints where the 0-4 age range displayed the highest call volumes (60,058 calls). The 65+ group consistently had the lowest number of calls. The youngest age groups (0-4, 5-14) displayed higher percentages of AGE calls attributed to the relevant chief complaints in comparison to the older age groups (15-44, 45-64, 65+) (Figure 4.4).

DISCUSSION

Norovirus was the most common cause of VGE cases and outbreaks in Ontario during the years 2009-2014. Overall, it was determined that 41.6% of all VGE outbreaks in iPHIS were attributed to norovirus, 63.4% of VGE PHOL laboratory submissions were attributed to norovirus. In addition, 36.6% of calls to THO were due to vomiting with and without diarrhea as a chief complaint and, therefore, likely attributed to norovirus. This work confirms previous research

that has identified norovirus as the most common cause of VGE and intestinal infections in the community (3,11–13).

The iPHIS data demonstrated that long-term care homes were the most commonly implicated institution type for both VGE and norovirus (Fig. 4.1). This is to be expected due to the higher incidence of VGE in older adults; the virus disproportionately causes more severe illness in vulnerable populations, such as young children, the elderly, and those with reduced/absent immune systems (3). The number of VGE outbreaks per year in child care facilities decreased after 2011 (Figure 4.1); using child care centres as a proxy for a young age group, this is likely due to the introduction of the Rotarix vaccine in the Ontario childhood vaccination schedule in August 2011 (14). Rotavirus is a common illness of young children (<5 yrs), its primary symptom being diarrhea (with vomiting being common). Therefore, the presence of rotavirus infection in these data likely impacts the number of VGE outbreaks in that age group. Other studies, both in Ontario and in countries worldwide, have verified this drop in rotavirus activity following the introduction of the vaccine (14,15).

Both the iPHIS and PHOL data showed a rise in norovirus activity in the winter of 2009/2010 when compared to the other winters analyzed. A rise in norovirus activity typically occurs with the introduction of a new strain into a population, mostly due to the quick mutation rate of the virus and lack of herd immunity in the population (13,16). The introduction of new strains can cause shifts in seasonality and/or increases in the number of outbreaks (17). It is likely that this was a result of the emergence of two novel strains: the GII.4 New Orleans strain which caused a global epidemic of norovirus, as well as a GII.12 strain (18,19). The GII.4 New Orleans strain

caused so many outbreaks and was so widespread that it was still detected in high numbers up until 2013; this may also explain the higher peak seen in Figure 4.2 for the 2011/2012 season (20).

There were 2,170 positive samples for rotavirus from the total amount of submitted samples where virus was detected, second only to norovirus (4,633 positive samples), demonstrating roughly a 2:1 ratio of norovirus:rotavirus. It is well known that norovirus tends to disproportionately affect very young (<5 yrs) and very old (65+ yrs) individuals when compared to middle-aged healthy individuals; as well, outbreaks are very common in high-density areas, such as daycares and preschools, as well as retirement homes and long-term care homes (4,21). Rotavirus has a similar outcome in that it disproportionately affects young children (<5 yrs), also resulting in outbreaks in daycares, preschools, etc. (3,14). The elderly (65+ yrs) and those living in long-term care homes may be privy to more acute medical care and a higher likelihood of samples being collected and submitted to public health authorities, which might explain the findings. It should be noted that the difference in norovirus- and rotavirus-positive samples was present even before the introduction of Rotarix into the Ontario childhood vaccination schedule in August 2011.

When the results of iPHIS and PHOL are matched, interesting patterns are evident. In theory, the PHOL results should be a multiple of the iPHIS results; the iPHIS dataset represents the number of VGE (and, as a subset, norovirus) outbreaks in Ontario, whereas the PHOL dataset represents the number of individual VGE (and, as a subset, norovirus) cases in Ontario. However, this was not the case. For example, the iPHIS dataset indicates there were 1,291 outbreaks of norovirus

from 2009-2014, yet PHOL indicates only 4,633 positive submitted samples of norovirus during the same timeframe which indicates roughly 3.6 cases per outbreak. This is in comparison to the average number of norovirus cases per outbreak from 2009-2014 as determined in iPHIS to be 30 (Hughes *et al*, unpublished data). These numbers may be higher in iPHIS because the dataset does not indicate the amount of people seeking medical care to report their symptoms and have samples taken. Rather, outbreaks are often reported by employees of the institution(s) where the outbreaks occur and declared by the medical officer of health. In addition, oftentimes outbreaks are declared as such after only a small percentage of ill individuals are tested for laboratory confirmation; the remaining ill individuals are left untested and determined solely on case definition. In contrast, the PHOL data may be artificially deflated because the dataset represents the people who have pursued medical care, therefore highlighting the lack of reporting of norovirus.

The difference in highest call volumes among age range groups (ages 15-44 for AGE, and 0-4 for the relevant chief complaints) seen in the THO dataset is likely due to the strong influence of rotavirus. This illness, whose primary symptoms are diarrhea and vomiting, disproportionately affects infants and young children; this would lead to a higher than average call volume in the 0-4 age group. The rotavirus vaccine was introduced in August 2011 and significantly decreased rotavirus activity in young children in Ontario (14). It has also been established that many callers phoning in to THO are parents concerned about their children; vomiting (with or without diarrhea) in 0-4 children would lead to a high call volume of parents concerned for their children's well-being (Adam Jones, personal communication). In addition to this, the 15-44 age

group likely had the highest number of callers for AGE because of the large number of illnesses encompassed within this category.

While the iPHIS and PHOL datasets contained a great deal of information and helped to describe the state of VGE and norovirus in Ontario, there was one clear disadvantage: a lack of community data (i.e. people suffering from illness at-home). The iPHIS dataset represented only institutional outbreaks, and the PHOL dataset was largely comprised of samples obtained from institutions. This is because it is only mandatory for institutional cases of norovirus and other VGE to be reported to Ontario public health authorities. In addition, many VGE cases (and norovirus specifically) suffer from underreporting by much of the population (22,23). The inclusion of THO data helps to bridge this gap in that it primarily collects community data. Although biases do exist in the THO dataset (i.e. the data are predominantly made up of middle-aged adults and parents phoning in on behalf of children (Adam Jones, personal communication)), it covers the part of the population that is not well-represented by the iPHIS or PHOL data. Syndromic data are known for their ability to reduce underreporting and represent a higher percentage of the population (24). Therefore, including THO data alongside iPHIS and PHOL covers a great deal of VGE in Ontario and helps to reduce reporting bias. These advantages emphasize the importance of developing and including unique and novel surveillance methods.

Syndromic data are becoming increasingly more common and frequently utilized in public health practice due to their array of advantages. They are timely, can detect new or emerging threats, supplement data produced by traditional surveillance systems, and are non-specific (can detect a

wide range of disease) (25). Telehealth data are particularly beneficial as they represent one of the timeliest syndromic data options available; this is because telephone medical helplines are one of the first points of medical care for those that are ill, after only home self-care (25). These data are also known for their availability in real-time, and ease of access electronically.

However, telehealth data are not as specific as other data sources and cannot necessarily be used to detect specific or severe disease outbreaks (25). In this study, telehealth data does not need to be specific because non-specific gastrointestinal calls provide the early warning of illness required for the surveillance system designed, and will be the most effective at observing norovirus and similar gastrointestinal illness in Ontario when combined with laboratory and outbreak data.

There are a few notable limitations to these data and analyses. In the iPHIS dataset, there was no age range to use for analyses. Rather, the institution type was used as an age proxy. In addition, there were many “gastroenteritis unspecified” and “gastroenteritis other” entries in iPHIS which likely contained additional norovirus cases. There was a lack of standardized reporting for norovirus and VGE in Ontario. Only institutional cases of norovirus were required to be reported to Ontario public health, and even that requirement suffered from underreporting and time lags. As a result, analyses of norovirus and VGE activity and burden are challenging due to the fact that many gaps in data currently exist, as well as notable biases in age group reporting across the province. Lastly, the acute gastroenteritis calls to THO may have included bacterial or other reasons for calling, but this could not be explored.

In this study, we have attempted to present the burden of VGE and, specifically, norovirus in the province of Ontario, Canada between 2009-2014. It is clear that norovirus is a highly prevalent illness and the most dominant cause of VGE in Ontario. Our findings are in line with those of similar studies that have been performed internationally, demonstrating norovirus as the leading cause of VGE (1–3). While a vaccine has been introduced in Ontario and countries worldwide to mitigate rotavirus infection, there is no such medical intervention currently in place for norovirus. Introducing preventative interventions, such as a vaccine, are ideal; however, other public health actions, such as novel surveillance techniques, are also necessary and highly effective strategies to monitor cases and reduce the impact of outbreaks. Improving surveillance on norovirus and similar gastrointestinal illness in Ontario would help to detect cases and outbreaks as early as possible and help public health professionals implement proper infection control measures to reduce illness scope. A combination of traditional and novel surveillance techniques is required to best represent the population of Ontario and reduce bias in surveillance. These actions would, in turn, decrease overall norovirus activity and the burden of disease.

ACKNOWLEDGEMENTS

We are grateful for the assistance provided from Public Health England, Public Health Ontario, and Sykes Assistance Services Corporation. In addition, many thanks are necessary to William Sears at the University of Guelph for all of his assistance with the analyses in this study.

REFERENCES

1. Morton VK, Thomas MK, McEwen SA. Estimated hospitalizations attributed to norovirus and rotavirus infection in Canada, 2006 – 2010. *Epidemiol Infect.* 2015;143:3528–37.

2. Yen C, Hall AJ. Challenges to estimating norovirus disease burden. *J Pediatric Infect Dis Soc.* 2013;2(1):61–2.
3. Patel MM, Hall AJ, Vinjé J, Parashar UD. Noroviruses: A comprehensive review. *J Clin Virol.* 2009;44(1):1–8.
4. Karst SM. Pathogenesis of noroviruses, emerging RNA Viruses. *Viruses.* 2010 Mar 23;2(3):748–81.
5. Ahmed SM, Hall AJ, Robinson AE, Verhoef L, Premkumar P, Parashar UD, et al. Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. *Lancet Infect Dis.* 2014;14(8):725–30.
6. Glass RI, Parashar UD, Estes MK. Norovirus gastroenteritis. *N Engl J Med.* 2009;361:1776–85.
7. Teunis PFM, Moe CL, Liu P, Miller SE, Lindesmith L, Baric RS, et al. Norwalk virus: How infectious is it? *J Med Virol.* 2008;80:1468–76.
8. Scallan E, Hoekstra RM, Angulo FJ, Tauxe R V., Widdowson M-A, Roy SL, et al. Foodborne illness acquired in the United States— major pathogens. *Emerg Infect Dis.* 2011;17(1):7–15.
9. Thomas MK, Murray R, Flockhart L, Pintar K, Pollari F, Fazil A, et al. Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne Pathog Dis.* 2013 Jul;10(7):639–48.
10. Bartsch SM, Lopman BA, Hall AJ, Parashar UD, Lee BY. The potential economic value of a human norovirus vaccine for the United States. *Vaccine.* 2012;30(49):7097–104.
11. Hall AJ, Rosenthal M, Gregoricus N, Greene SA, Ferguson J, Henao OL, et al. Incidence of acute gastroenteritis and role of norovirus, Georgia, USA, 2004–2005. *Emerg Infect*

- Dis. 2011;17(8):1381–8.
12. Loveridge P, Cooper D, Elliot AJ, Harris J, Gray J, Large S, et al. Vomiting calls to NHS Direct provide an early warning of norovirus outbreaks in hospitals. *J Hosp Infect.* 2010;74:385–93.
 13. Gastañaduy PA, Hall AJ, Curns AT, Parashar UD, Lopman BA. Burden of norovirus gastroenteritis in the ambulatory setting - United States, 2001-2009. *J Infect Dis.* 2013;207(7):1058–65.
 14. Wilson SE, Rosella LC, Wang J, Le Saux N, Crowcroft NS, Harris T, et al. Population-level impact of Ontario’s infant rotavirus immunization program: Evidence of direct and indirect effects. *PLoS One.* 2016;11(5):e015340.
 15. Thomas SL, Walker JL, Fenty J, Atkins KE, Elliot AJ, Hughes HE, et al. Impact of the national rotavirus vaccination programme on acute gastroenteritis in England and associated costs averted. *Vaccine.* 2016;35(4):680–6.
 16. Karst SM, Baric RS. What is the reservoir of emergent human norovirus strains? *J Virol.* 2015;89(11):5756–9.
 17. Allen DJ, Adams NL, Aladin F, Harris JP, Brown DWG. Emergence of the GII-4 norovirus Sydney2012 strain in England, winter 2012-2013. *PLoS One.* 2014;9(2).
 18. Vega E, Vinjé J. Novel GII.12 norovirus strain, United States, 2009-2010. *Emerg Infect Dis.* 2011;17(8):1516–8.
 19. Yen C, Wikswo ME, Lopman BA, Vinje J, Parashar UD, Hall AJ. Impact of an emergent norovirus variant in 2009 on norovirus outbreak activity in the United States. *Clin Infect Dis.* 2011;53(6):568–71.
 20. Medici MC, Tummolo F, De Grazia S, Calderaro A, De Conto F, Terio V, et al.

- Epidemiological dynamics of norovirus GII.4 variant New Orleans 2009. *J Gen Virol.* 2015;96(9):2919–27.
21. Said MA, Perl TM, Sears CL. Gastrointestinal flu: Norovirus in health care and long-term care facilities. *Clin Infect Dis.* 2008 Nov;47(9):1202–8.
 22. Tam CC, Rodrigues LC, Viviani L, Dodds JP, Evans MR, Hunter PR, et al. Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. *Gut.* 2012;61(1):69–77.
 23. Bernard H, Werber D, Höhle M. Estimating the under-reporting of norovirus illness in Germany utilizing enhanced awareness of diarrhoea during a large outbreak of Shiga toxin-producing *E. coli* O104:H4 in 2011 - a time series analysis. *BMC Infect Dis.* 2014;14:116.
 24. da Silva AD, Evangelista M do SN. Syndromic surveillance: Etiologic study of acute febrile illness in dengue suspicious cases with negative serology. Brazil, Federal district, 2008. *Rev Inst Med Trop Sao Paulo.* 2010;52(5):237–42.
 25. Triple-S Project. Guidelines for designing and implementing a syndromic surveillance system. 2013.

TABLE

Table 4.1. List of acute gastroenteritis chief complaints to Telehealth Ontario and those selected as the “relevant chief complaints” for norovirus and gastroenteritis-related illness

THO ¹ AGE ² chief complaints	Relevant chief complaints
Vomiting	Vomiting
Diarrhea	Vomiting with diarrhea
Vomiting with diarrhea	Vomiting without diarrhea
Vomiting without diarrhea	
Abdominal pain	
Vomiting with blood	

¹Telehealth Ontario
²Acute gastroenteritis

FIGURES

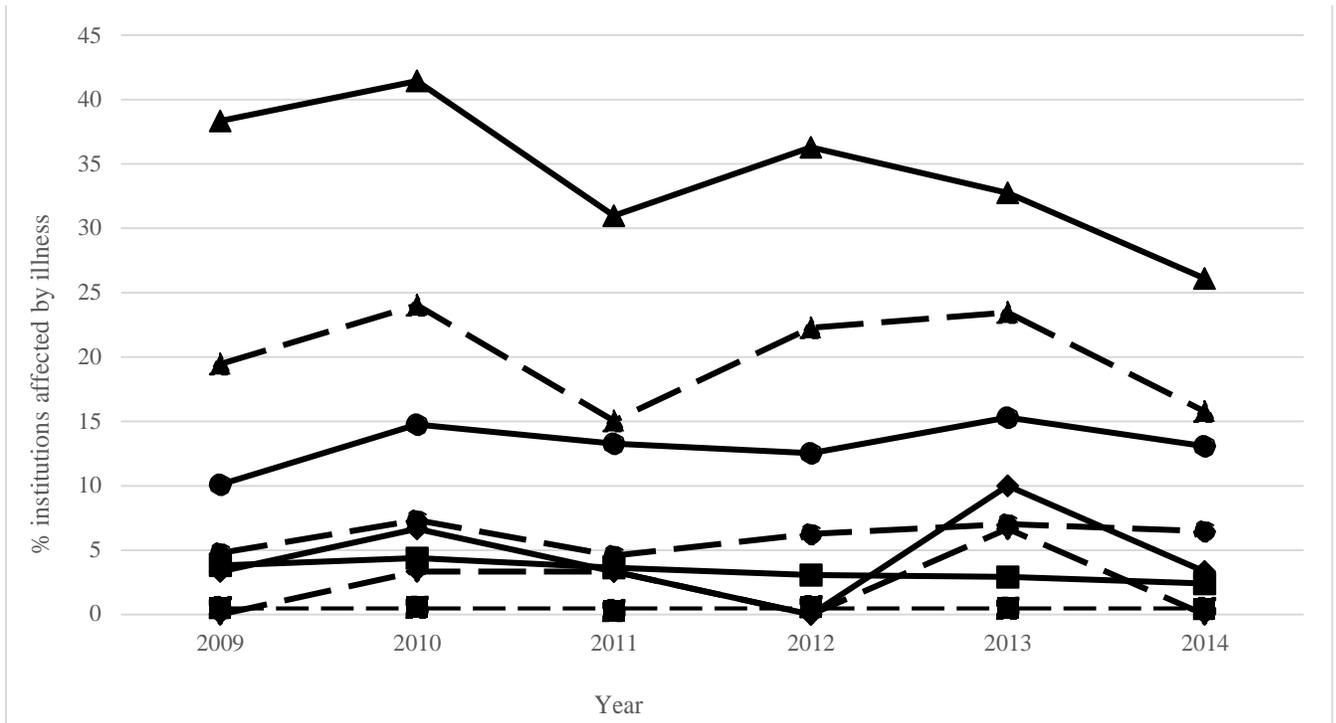


Figure 4.1. Percent of institutions affected by viral gastroenteritis (solid lines) and norovirus (dashed lines) outbreaks in Ontario, 2009-2014. Triangle: long-term care homes, circle: retirement homes, square: child care centre, diamond: correctional facility.

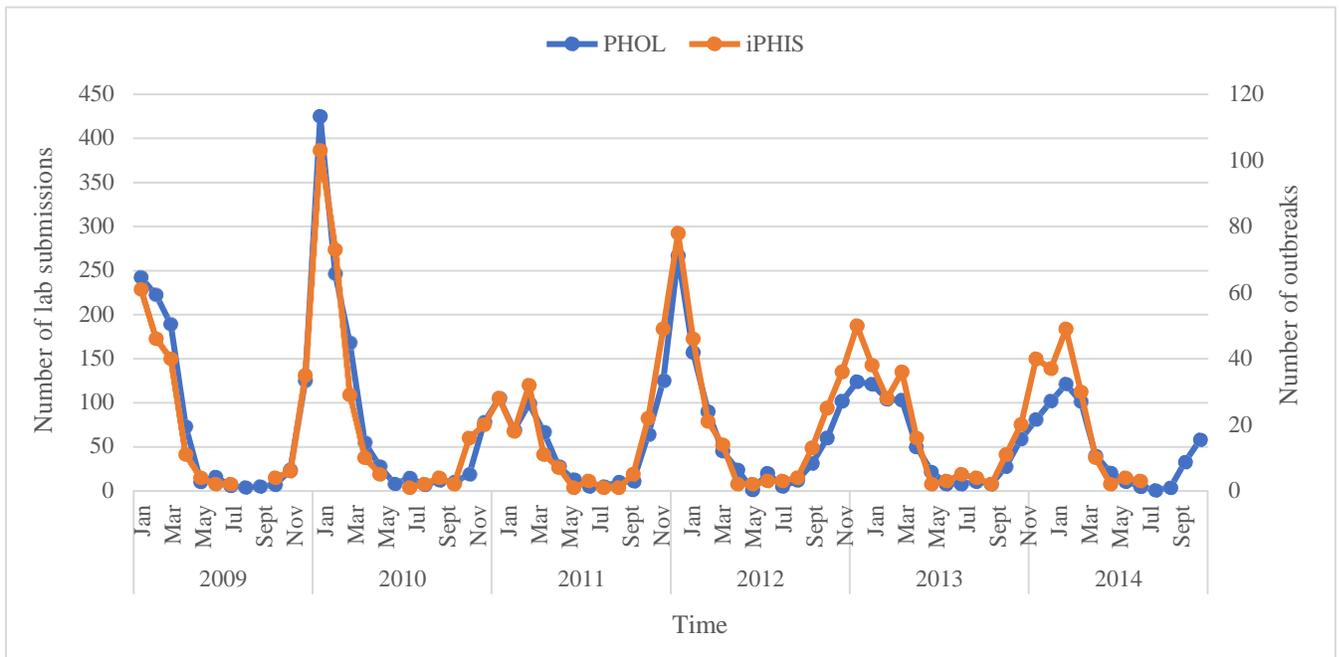


Figure 4.2. Seasonality of norovirus laboratory submissions and outbreaks in Ontario by month and year, 2009-2014. PHOL: Public Health Ontario Laboratories (laboratory submissions); iPHIS: integrated Public Health Information System (outbreaks).

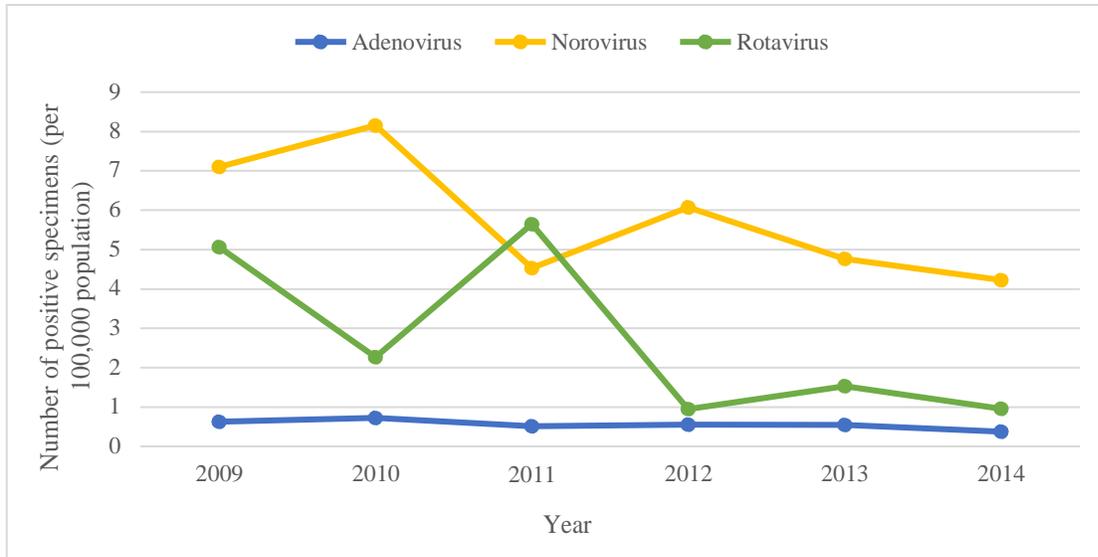


Figure 4.3. Number of positive specimens submitted to PHO by virus type and year, 2009-2014 (per 100,000 population). Additional virus types (astrovirus, sapovirus, other Picornaviridae, and other Caliciviridae) removed due to rates <0.10 positive specimens per 100,000 population.

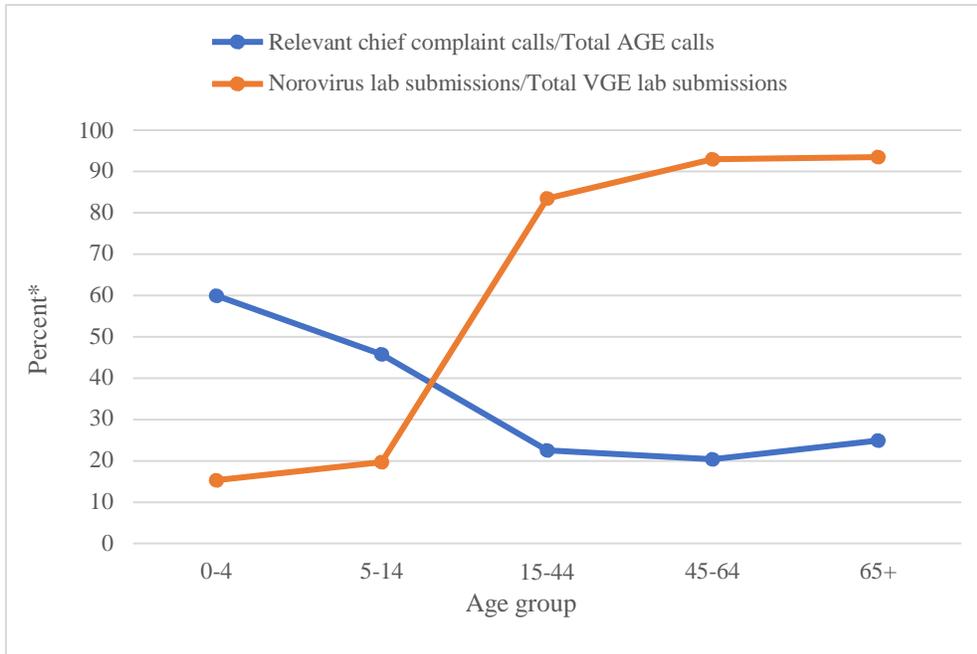


Figure 4.4. Percent of viral gastroenteritis (VGE) activity attributed to norovirus. Orange line: norovirus-positive specimens as a percent of VGE-positive specimens submitted to Public Health Ontario (PHO) by age group, 2009-2014. Blue line: relevant chief complaint (vomiting, diarrhea, vomiting+diarrhea) calls to Telehealth Ontario (THO) as a percent of all acute gastroenteritis (AGE)-related calls by age group, 2009-2014. *Because calls to THO are less specific than those to PHO, conclusions are therefore specific only to vomiting and related gastroenteritis calls as a percentage of AGE calls.

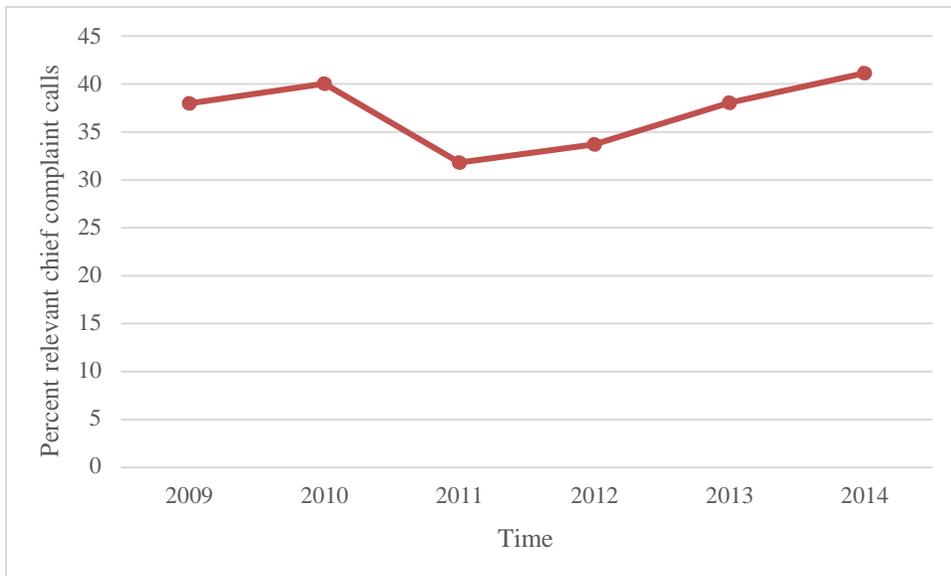


Figure 4.5. Percentage of acute gastroenteritis calls to Telehealth Ontario attributed to relevant chief complaints (vomiting, diarrhea, and vomiting+diarrhea calls) by year, 2009-2014.

CHAPTER 5

Monitoring telehealth vomiting calls as a potential public health early warning system for seasonal norovirus activity in Ontario, Canada

ABSTRACT

Norovirus is a predominant cause of acute gastroenteritis. Norovirus gastroenteritis occurs with regular seasonality in winter months, yet underreporting and delays in laboratory surveillance data have the potential to limit public health mitigation capabilities. This study presents a proposal for an early warning system for norovirus activity in Ontario, Canada using telehealth syndromic surveillance call data.

Positive norovirus laboratory submissions and telehealth call data for vomiting and diarrhea chief complaints (used as a proxy for norovirus) were analyzed for Ontario, Canada over the period 2009-2015. Negative binomial regression analyses were fitted to determine if statistically significant relationships existed between the two series of data including the presence of time lags. Early warning alarm thresholds based on telehealth data were devised using Shewhart (control) chart methods.

Negative binomial regression analyses using a range of chief complaints against laboratory data revealed vomiting calls to have the strongest correlation, with a statistically significant two-week lag between the vomiting calls to telehealth and norovirus-positive laboratory test data ($p < 0.05$; pseudo- $R^2 = 0.1861$). Early warning alarm thresholds demonstrated an onset range of vomiting call volumes of 419-474 calls per week and weeks of onset of 42-45 (Oct-Nov), signaling the start of the winter norovirus season.

Telehealth vomiting calls demonstrated potential to provide early warning of norovirus activity. Early warning alarm thresholds have the ability to alert public health authorities to increasing seasonal norovirus activity, thus enabling interventions to be adopted.

INTRODUCTION

Norovirus is a predominant cause of infectious gastroenteritis in countries worldwide (1–5). It accounts for approximately 50% of acute gastroenteritis (AGE) and greater than 90% of viral gastroenteritis (6,7). The incubation period ranges between 10-48 hours and illness duration is generally 1-3 days with self-limiting symptoms; however, this duration is often longer (e.g. 4-6 days) in vulnerable populations such as hospital patients or young children (2,8). Symptomatic infection of norovirus presents as acute vomiting, diarrhea, abdominal cramps, and nausea, with severe vomiting and diarrhea (non-bloody) being most common (2,5,9).

Norovirus activity occurs with predictable seasonality. Although the specific months vary, activity peaks during cold weather winter months in temperate climates, similar to other common respiratory viruses (e.g. influenza and rhinovirus) (2,10). Due to its predictable winter seasonality, norovirus is often referred to as “winter vomiting disease” (11,12). However, outbreaks can occur throughout the year, especially in years when a new norovirus variant has emerged (13).

Traditional surveillance of norovirus activity plays an essential role in the identification of seasonal trends, disease burden in the population, and the extent of geographical spread. This traditional surveillance is reliant upon datasets such as laboratory submissions and outbreak case

reporting to carry out these tasks. However, due to the self-limiting nature and short duration of norovirus illness, public health surveillance is limited by under-reporting and subsequent time lags (14–17). Hall et al. (2011) demonstrated that only 15% of individuals affected with AGE illness seek medical intervention, and of those, diagnostic testing samples were only requested from 13% (14). Other studies have described the absence of routine surveillance and frequent under-reporting, which affect the overall ability to monitor disease activity (15,18). The time lags associated with traditional laboratory surveillance methods have long been recognized as a limitation for disease surveillance due to their delayed effects on public health interventions. A study performed by Ashford et al. (2003) demonstrated across a number of public health scenarios that traditional reporting techniques had a delay of 0-26 days between the index case and when the problem was identified through surveillance (17). Such lags include the time taken for general practitioners to submit samples to the laboratory, and for the specific laboratory tests to be performed and reported (17). These delays reduce the timeliness of disease monitoring and affect the ability to implement effective public health interventions.

This paper presents a proposal for an early warning syndromic surveillance system using telehealth syndromic surveillance call data to detect winter norovirus activity in Ontario, Canada. The use of real-time surveillance for monitoring norovirus activity could greatly improve the ability of public health authorities to monitor the illness and implement interventions to reduce viral spread.

METHODS

Syndromic surveillance telehealth call data

Telehealth Ontario call data were obtained from the Ontario Ministry of Health and Long-Term Care and Sykes Assistance Services Corporation for the period 17/6/2011 – 31/8/2015. The data were received as daily call counts and aggregated into weekly counts for sufficient sample sizes in analyses. All AGE calls were requested which included the chief complaints: vomiting, diarrhea, vomiting and diarrhea, vomiting without diarrhea (subsequently coded as “vomiting”), abdominal pain, and vomiting with blood. For the purposes of this study, the “vomiting,” “diarrhea,” and “vomiting and diarrhea” chief complaints were included in analyses (subsequently referred to as the “relevant chief complaints”).

Norovirus laboratory reports

Norovirus laboratory data were obtained from Public Health Ontario for the period 17/6/2011 – 30/9/2014 for samples submitted for norovirus confirmatory testing in Ontario. The data were received as daily norovirus test results and aggregated into weekly counts for sufficient sample sizes in analyses. For the purposes of this study, the collection dates (the date of sample collection from infected patients), and only samples which tested positive for the presence of norovirus were utilized in analyses.

Regression analyses

Both the telehealth and laboratory datasets were analyzed over the period 17/6/2011 – 30/9/2014 (week 24, 2011 – week 40 2014). A “public holiday” variable was added to both datasets to indicate the occurrence of an Ontario-wide statutory holiday; this was included because previous

studies have shown the effects public holidays have on call volumes and the necessity to treat holidays differently than normal days (19). A “week commencing” variable was also added to the datasets to indicate the start of the week analyzed; this was included as a trend variable to assess long-term patterns in the data.

The laboratory dataset was used at time 0 and shifted one, two, and three weeks forward to construct lag variables (t_0 , t_1 , t_2 , and t_3 , respectively) to observe whether a statistically significant time lag existed between the datasets. Prior to performing this study, an assumption was made that the telehealth data might precede the laboratory data by a number of weeks.

Sets of four negative binomial regression models were constructed with the original laboratory data (t_0), and the t_1 , t_2 , and t_3 lag variables as predictor variables. These predictors were regressed against weekly counts of calls to telehealth using the following parameter combinations: all-ages and all relevant chief complaints (“index parameter model”), all ages and vomiting-only chief complaints (“vomiting,” and “vomiting and diarrhea” calls only), 0-4 ages and all relevant chief complaints, ≥ 65 ages and all relevant chief complaints, 0-4 ages and vomiting-only chief complaints, and ≥ 65 ages and vomiting-only chief complaints. These various combinations were tested to observe whether the index parameter model was the most sensitive to early warning, or if a combination of indicators were most sensitive. The specific age groups selected were tested as they are known to be disproportionately affected by norovirus infection, and the vomiting-only chief complaints were tested because vomiting is the primary symptom of norovirus.

Included in each of these models were the public holiday and week commencing variables. The (McFadden's) pseudo- R^2 values of these models were compared to each other to determine the best-fitting model; however, it should be noted that pseudo- R^2 values should not be interpreted the same as R^2 for normal data.

All negative binomial regression analyses were performed in STATA v.13.0 (StataCorp LLC, College Station, Texas, USA).

Early warning system threshold calculations

The telehealth data were plotted from 19/6/2011 – 29/8/2015 (week 25, 2011 – week 34, 2015). A downward trend in call volumes was observed during this time period; therefore, a linear trendline was inserted to calculate the overall slope. This slope (-0.46) was subsequently used to standardize the overall baseline of the data, setting the very first week as the referent.

The periods of increased laboratory reports of norovirus were observed to be regular, within weeks 44-24 (Oct/Nov-June) and were therefore defined as the active season; weeks 25-43 (June-Oct) were defined as out-of-season. Using these weeks as approximate measures for the winter norovirus season, the average weekly call volume for the out-of-season periods was calculated (referred to as the baseline). Using this baseline, the standard deviation (SD) was then calculated and modeled incrementally (up to 4SDs) above the overall average for the out-of-season call volumes for the period of 19/6/2011 to 29/8/2015. This was done using Shewhart (control) chart methods (20), a method which displays the mean of the statistic on a graph and

models standard deviations of the statistic above and/or below the mean to indicate a threshold of data variation.

All threshold calculations were executed using Microsoft Office (Excel) 2016.

RESULTS

Regression analyses

The negative binomial regression models demonstrated that the all-ages, vomiting-only chief complaint t_0 , t_1 , t_2 , and t_3 models were the best-fitting of all the parameter combinations tested. This parameter combination produced the highest pseudo- R^2 values when compared to all other parameter combinations (Table 5.1). In this combination, the t_0 and t_1 variables were not significant ($p > 0.05$); however, the t_2 and t_3 lag variables were significant [$p = 0.004$ (95% CI 0.00097-0.0051)], 0.003 (95% CI 0.00093-0.0047), respectively) (Table 5.2). Although both the t_2 and t_3 lag variables were significant, the fit of the t_2 model (pseudo- $R^2 = 0.1861$) was higher than that of the t_3 model (pseudo- $R^2 = 0.1852$) (Table 5.1). A visual comparison/overlay of the telehealth and laboratory datasets over the timeframe used in the regression analyses (17/6/2011 – 30/9/2014) is displayed in Figure 5.1, and the full model results are described in Table 5.2.

The public holiday variable was not significant in any of the t_0 , t_1 , t_2 , and t_3 all-ages, vomiting-only chief complaint models ($p \geq 0.05$). However, due to its practical significance and importance in syndromic models demonstrated by other research groups worldwide, it was kept in the final models. The week commencing variable was observed to be significant in all of the t_0 , t_1 , t_2 , and

t₃ all-ages, vomiting-only chief complaint models ($p < 0.05$), and was therefore included in the final models.

Early warning threshold calculations

Trend-adjusted call volumes were plotted and visualized (Figure 5.2). SDs were calculated on the baseline (average number of weekly calls recorded during the out-of-season time period) of vomiting-only calls. Each SD level of vomiting-only calls was set as an early warning threshold across each of the four winter norovirus seasons analyzed. Call volumes were observed noting the week during which each SD threshold was breached, and the peak week of activity during each season (Table 5.3). Where weekly vomiting call volumes breached the SD level of vomiting-only calls to telehealth for three consecutive weeks, this was noted as a winter norovirus season early warning alarm.

Applying the recommended control chart standard of using the mean + 3SDs to devise an early warning threshold (20), five false alarms would have been triggered during the 19/6/2011 – 29/8/2015 period assessed. Alarms were declared as false if they surpassed the 3SD level of vomiting-only calls, but not for three consecutive weeks. There were two occasions in 2012 (week 25, 433 vomiting-only calls; week 40, 426 vomiting-only calls), and three occasions in 2013 (week 25, 448 vomiting-only calls; week 28, 454 vomiting-only calls; week 39, 430 vomiting-only calls) that surpassed the 3SD level of vomiting-only calls. This is in comparison to the 8 false alarms that would have been triggered using the 2SD level of vomiting-only calls, which has also been recommended by other researchers (21). Therefore, the recommended

control chart standard of the mean + 3SD is ideal for this system. The presence of false alarms was assessed during the out-of-season weeks.

DISCUSSION

Main findings

This study provides an initial validation that there is a statistical association between vomiting calls made to a telehealth system in Ontario and norovirus-positive laboratory test results.

Regression analysis provided evidence of a two-week lag of telehealth vomiting calls ahead of norovirus-positive laboratory reports, thus revealing potential to provide early warning of seasonal norovirus activity by monitoring telehealth vomiting calls. A series of early warning thresholds were developed based upon previously published methods and tested for the triggering of alerts in the study period.

Telehealth calls with vomiting as the chief complaint in all ages demonstrated the timeliest prediction of the winter norovirus season, thus were the focus of this analysis. This timely indication over all other telehealth parameters assessed is likely due to vomiting being the most predominant symptom of the illness in adults (5,9). These results were comparable to those found by Loveridge et al. (2010), who demonstrated that vomiting-only calls in the <5 age group gave the most sensitive prediction of the winter norovirus season (22). A significant two-week prediction of telehealth calls ahead of norovirus laboratory reports was also observed, emphasizing the value of using community-based surveillance systems that monitor initial presentation of symptoms at a population level, rather than subsequent patient-level laboratory reports. This also highlights the importance of integrating syndromic surveillance systems with

traditional systems for enhanced disease surveillance. Loveridge et al. (2015) found an early warning of 6-11 weeks in the <5 ages; however, were able to calculate AGE calls as a proportion of total calls to telehealth and adjust the system to the most sensitive indicators (22).

A set of winter norovirus early warning alarm thresholds were devised for the four winter periods analyzed, indicating the week during which an early warning alert would be appropriate. A recommendation of using a threshold based on 3SDs was most appropriate at capturing the winter norovirus season without setting off a high number of false alerts in the out-of-season periods, and without significant time delay. This 3SD level agrees with several studies in the fields of syndromic surveillance and/or norovirus, which have followed this metric to accurately identify variation within sample statistics (20,23–25). However, using the 2SDs in addition to the 3SD threshold could also prove to be useful as a possible additional early alert signal to warn public health authorities that an alarm may be triggered the following week. This would allow earlier public health action to be considered even before an official alert is triggered.

While this telehealth early warning system demonstrated the strongest association with a two-week prediction of the norovirus laboratory reports, additional time lags must be considered. The average time delay between sample collection and laboratory reporting was 2.6 days (Hughes, S.L. et al., unpublished data). However, it is important to note that this was based upon analysis of retrospective data; prospectively, these data would not necessarily be tested or reported in a timely manner, with results often batched and/or reported at a later date, extending the delay to reporting. Therefore, prospective reporting of laboratory data always needs to incorporate these additional reporting delays, and further highlights the importance of having real-time data available for use in surveillance systems to detect disease activity as early as possible.

The observed decrease in the telehealth weekly vomiting calls was evident over the entire period analyzed in this study. Such decreases may be due to the overall decrease in telehealth utilization in Ontario and increased usage of the internet for at-home health care advice. A high percentage of ill patients in countries worldwide are now utilizing the internet as their first point of medical care (26,27); an American study in 2013 concluded that 80% of internet users had searched the web for health care advice, in comparison to 62% in 2001 (28). These statistics, in combination with the decreases observed in telehealth call numbers as far back as 2009, may explain the general downward trend observed in the data in this study. To verify whether this decrease was limited just to the gastroenteritis reasons for call, or affected all telehealth calls, additional calculations comparing both the telehealth gastroenteritis call volumes and total call volumes (all reasons for calls) during the same timeframe would be required. However, these data were not available for this study.

Strengths and limitations

This study successfully utilized two existing datasets to provide an understanding of seasonal norovirus activity in the community. The advantage of using telehealth data was the improved representativeness of community activity. Laboratory data are frequently under-representative of disease burden and are biased, with sampling predominantly occurring in patients from higher risk or vulnerable groups and those from institutionalized outbreaks, particularly from hospital or long-term care settings (29,30). Telehealth Ontario provides a service to all Ontarians, and better captures those patients who do not present to other healthcare services. In addition, this study was able to highlight the early warning telehealth data were able to provide over the laboratory

data, a key piece of information in early warning and reducing the spread of illness in populations of high risk.

There were several limitations associated with the datasets used in this study. It was only possible to undertake a comparative analysis across three norovirus seasons due to a lack of data from 2015 in the laboratory dataset. The lack of sufficient daily data in both datasets prevented analysis at a more granular daily level, as laboratory submissions and telehealth call counts were aggregated to weekly counts due to low daily counts. Additionally, the overall decreasing trend of weekly vomiting calls to telehealth over the study period made the early warning threshold calculations challenging. Adding a slope adjustment to calculate an adjusted baseline was required; however, if adopted as a prospective system, further decreasing call volumes would need to be continually monitored and thresholds adjusted accordingly. Finally, the telehealth data monitored the presentation of patients' reported symptoms and were not a confirmation of norovirus illness. The laboratory data were used to determine the sensitivity of vomiting calls to norovirus activity; however, it is still important to note that the telehealth data can only be used as a proxy for norovirus activity.

Implications for public health practice

It is recommended that the early warning system proposed in this study be further validated with subsequent implementation and routine use in the province of Ontario, Canada. Adopting syndromic surveillance alongside existing laboratory and outbreak surveillance methods would improve public health monitoring of disease activity and help identify threats and outbreaks as early as possible. This work has highlighted the utility of telehealth call data to detect the onset

of the winter norovirus season as early as possible to allow health care professionals to implement infection control guidelines, including hand hygiene, patient isolation, and ward closures in high-risk settings (i.e. hospitals, long-term care homes) (31).

Future directions

If adopted into routine public health practice, the norovirus surveillance and early warning system developed in this study should be updated prospectively with daily data, should sufficient daily counts be available in Ontario, to provide the timeliest surveillance information. In addition, as telehealth data become more commonly utilized in public health practice, it will become more feasible to develop similar, well-adapted telehealth syndromic surveillance systems for the detection of a wider range of public health issues, including other communicable diseases, (e.g. influenza) and environmental impacts (e.g. heatwaves). Additional sources of syndromic data could be used to complement telehealth data, such as pharmacy purchases, internet search queries, school absenteeism, and general practitioner diagnoses, and integrated into existing laboratory and outbreak surveillance methods (32,33). In doing so, data gaps and delays would be reduced and more effective public health actions performed to protect Ontarians.

ACKNOWLEDGEMENTS

We recognize the Ontario Ministry of Health and Long-Term Care, Sykes Assistance Services Corporation, and Public Health Ontario for their assistance with retrieving the datasets used in this study.

REFERENCES

1. Morton VK, Thomas MK, McEwen SA. Estimated hospitalizations attributed to norovirus and rotavirus infection in Canada, 2006 – 2010. *Epidemiol Infect.* 2015;143:3528–37.
2. Patel MM, Hall AJ, Vinjé J, Parashar UD. Noroviruses: A comprehensive review. *J Clin Virol.* 2009;44(1):1–8.
3. Yen C, Hall AJ. Challenges to estimating norovirus disease burden. *J Pediatric Infect Dis Soc.* 2013;2(1):61–2.
4. Widdowson M-A, Sulka A, Bulens SN, Beard RS, Chaves SS, Hammond R, et al. Norovirus and foodborne disease, United States, 1991–2000. *Emerg Infect Dis.* 2005;11(1):95–102.
5. Kaplan JE, Gary GW, Baron RC, Singh N, Schonberger LB, Feldman R, et al. Epidemiology of Norwalk gastroenteritis and the role of Norwalk virus in outbreaks of acute nonbacterial gastroenteritis. *Ann Intern Med.* 1982;96:756–61.
6. Karst SM. Pathogenesis of noroviruses, emerging RNA Viruses. *Viruses.* 2010 Mar 23;2(3):748–81.
7. Lopman BA, Reacher MH, Duijnhoven Y Van, Hanon F-X, Brown D, Koopmans M. Viral gastroenteritis outbreaks in Europe, 1995 – 2000. *Emerg Infect Dis.* 2003;9(1):90–6.
8. Lopman BA, Reacher MH, Vipond IB, Sarangi J, Brown DWG. Clinical manifestation of Norovirus gastroenteritis in health care settings. *Clin Infect Dis.* 2004;39:318–24.
9. Rockx B, de Wit M, Vennema H, Vinje J, Bruin E De, van Duynhoven Y, et al. Natural history of human Calicivirus infection : A prospective cohort study. *Clin Infect Dis.* 2002;35:246–53.
10. Mounts AW, Ando T, Koopmans M, Bresee JS, Noel J, Glass RI. Cold weather

- seasonality of gastroenteritis associated with Norwalk-like viruses. *J Infect Dis.* 2000;181(Suppl 2):S284-287.
11. Kapikian AZ, Wyatt RG, Dolin R, Thornhill TS, Kalica AR, Chanock RM. Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. *J Virol.* 1972;10(5):1075–81.
 12. Zahorsky J. Hyperemesis Hiemis or the winter vomiting disease. *Arch Pediatr Adolesc Med.* 1929;46:391–5.
 13. Lopman BA, Reacher M, Gallimore C, Adak GK, Gray JJ, Brown DWG. A summertime peak of “winter vomiting disease”: Surveillance of noroviruses in England and Wales, 1995 to 2002. *BMC Public Health.* 2003;3(13).
 14. Hall AJ, Rosenthal M, Gregoricus N, Greene SA, Ferguson J, Henao OL, et al. Incidence of acute gastroenteritis and role of norovirus, Georgia, USA, 2004-2005. *Emerg Infect Dis.* 2011;17(8):1381–8.
 15. Moore MD, Goulter RM, Jaykus L-A. Human Norovirus as a Foodborne Pathogen: Challenges and Developments. *Annu Rev Food Sci Technol.* 2015;6(1):411–33.
 16. Tam CC, Rodrigues LC, Viviani L, Dodds JP, Evans MR, Hunter PR, et al. Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. *Gut.* 2012;61(1):69–77.
 17. Ashford DA, Kaiser RM, Bales ME, Shutt K, Patrawalla A, McShan A, et al. Planning against biological terrorism: Lessons from outbreak investigations. *Emerg Infect Dis.* 2003;9(5):515–9.
 18. Scallan E, Hoekstra RM, Angulo FJ, Tauxe R V., Widdowson M-A, Roy SL, et al. Foodborne illness acquired in the United States— major pathogens. *Emerg Infect Dis.*

- 2011;17(1):7–15.
19. Buckingham-Jeffery E, Morbey R, House T, Elliot AJ, Harcourt S, Smith GE. Correcting for day of the week and public holiday effects: improving a national daily syndromic surveillance service for detecting public health threats. *BMC Public Health*. 2017;17(1):477.
 20. Best M, Neuhauser D. Walter A Shewhart, 1924, and the Hawthorne factory. *Qual Saf Heal Care*. 2006;15(2):142–3.
 21. Wang R, Jiang Y, Michael E, Zhao G. How to select a proper early warning threshold to detect infectious disease outbreaks based on the China infectious disease automated alert and response system (CIDARS). *BMC Public Health*. 2017;17:1–10.
 22. Loveridge P, Cooper D, Elliot AJ, Harris J, Gray J, Large S, et al. Vomiting calls to NHS Direct provide an early warning of norovirus outbreaks in hospitals. *J Hosp Infect*. 2010 Apr;74:385–93.
 23. Larsson J, Landstad BJ, Wiklund H, Vinberg S. Control charts as an early-warning system for workplace health outcomes. *Work*. 2011;39(4):409–25.
 24. Dijk A van, Dawson E, Moore KM, Belanger P. Risk assessment during the Pan American and Parapan American Games, Toronto, 2015. *Public Health Rep*. 2017;132(Supplement 1):1065–105.
 25. Siqueira JAM, Nunes AAGC, Ferreira MS, Linhares A da C, Ramos EMLS, Gabbay YB. Application of time series control charts to model and monitor the seasonality of norovirus. *Rev Pan-Amazônica Saúde*. 2015;6(2):61–8.
 26. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ*. 2015;351:h3480.

27. Tan SS-L, Goonawardene N. Internet Health Information Seeking and the Patient-Physician Relationship: A Systematic Review. *J Med Internet Res.* 2017;19(1):e9.
28. NBC News. More people search for health online [Internet]. 2013 [cited 2017 Dec 15]. Available from: <http://www.nbcnews.com/id/3077086/t/more-people-search-health-online/#.WjW-PhNzJPY>
29. Van-Dijk A, McGuinness D, Rolland E, Moore KM. Can Telehealth Ontario respiratory call volume be used as a proxy for emergency department respiratory visit surveillance by public health? *Can J Emerg Med.* 2008;10(1):18–24.
30. Trainor E, Lopman B, Iturriza-Gomara M, Dove W, Ngwira B, Nakagomi O, et al. Detection and molecular characterisation of noroviruses in hospitalised children in Malawi, 1997 – 2007. *J Med Virol.* 2013;85:1299–306.
31. MacCannell T, Umscheid CA, Agarwal RK, Lee I, Kuntz G, Stevenson KB, et al. Guidelines for the Prevention and Control of Norovirus Gastroenteritis Outbreaks in Healthcare Settings. Centers for Disease Control and Prevention. 2011.
32. Triple-S Project. Guidelines for designing and implementing a syndromic surveillance system. 2013.
33. Pavlin JA. Syndromic surveillance for infectious diseases. In: *Infectious Disease Surveillance: Second Edition.* 2nd ed. Armed Forces Health Surveillance Center, Silver Spring, MD, United States: Blackwell Publishing Ltd; 2013. p. 470–81.

TABLES

Table 5.1. Pseudo-R² values for all negative binomial regression models

Model parameters		Model type (stage of lag)*			
Age (years)	Telehealth chief complaints	t ₀	t ₁	t ₂	t ₃
All	Includes vomiting ¹	0.1733	0.1828	<u>0.1861</u>	0.1852
0-4	Includes vomiting	0.1760	0.1834	<u>0.1849</u>	0.1841
0-4	Includes vomiting and/or diarrhea ²	0.1663	0.1764	0.1783	<u>0.1791</u>
All	Includes vomiting and/or diarrhea	0.1561	0.1702	<u>0.1758</u>	0.1748
≥65	Includes vomiting and/or diarrhea	0.0878	0.0845	<u>0.0993</u>	0.0831
≥65	Includes vomiting	0.0779	0.0767	<u>0.0899</u>	0.0802

The highest value in each row is underlined.

¹Vomiting-only calls encompass “vomiting” and “vomiting and diarrhea” telehealth chief complaints

²Vomiting and diarrhea calls encompass “vomiting,” “diarrhea,” and “vomiting and diarrhea” telehealth chief complaints

*Laboratory data used in negative binomial models was regressed at time 0 (t₀), and shifted one (t₁), two (t₂), and three (t₃) weeks in the future, separately against telehealth calls

Table 5.2. Model results for norovirus laboratory data predictor variable in all negative binomial regression models

Model parameters		Model type (stage of lag)*											
Age (years)	Telehealth chief complaints	t ₀			t ₁			t ₂			t ₃		
		Coefficient	p-value	95% CI ³	Coefficient	p-value	95% CI	Coefficient	p-value	95% CI	Coefficient	p-value	95% CI
All	Includes vomiting ¹	-0.0010	0.442	(-0.0036)- 0.0016	0.000055	0.964	(-0.0023)- <u>0.0024</u>	0.0030	<u>0.004</u>	0.00097- 0.0051	0.0028	<u>0.003</u>	0.00093- 0.0047
0-4	Includes vomiting	0.00070	0.630	(-0.0021)- 0.0035	0.00036	0.804	(-0.0025)- 0.0032	0.0027	<u>0.040</u>	0.00013- 0.0052	0.0025	<u>0.039</u>	0.00013- 0.0049
0-4	Includes vomiting and/or diarrhea ²	0.00080	0.544	(-0.0018)- 0.0034	0.0013	0.306	(-0.0012)- 0.0038	0.0027	<u>0.017</u>	0.00047- 0.0049	0.0031	<u>0.002</u>	0.0011- 0.0051
All	Includes vomiting and/or diarrhea	0.00038	0.738	(-0.0019)- 0.0026	0.0018	0.064	(-0.00010)- 0.0037	0.0034	<u>0.000</u>	0.0019- 0.0050	0.0031	<u>0.000</u>	0.0017- 0.0046
≥65	Includes vomiting and/or diarrhea	0.0082	<u>0.000</u>	0.0048- 0.012	0.0082	<u>0.000</u>	0.0047- 0.012	0.0096	<u>0.000</u>	0.0067- 0.012	0.0081	<u>0.000</u>	0.0052- 0.011

Significant p-values in all rows are underlined

¹Vomiting-only calls encompass “vomiting” and “vomiting and diarrhea” telehealth chief complaints

²Vomiting and diarrhea calls encompass “vomiting,” “diarrhea,” and “vomiting and diarrhea” telehealth chief complaints

³CI = Confidence interval

*Laboratory data used in negative binomial models was regressed at time 0 (t₀), and shifted one (t₁), two (t₂), and three (t₃) weeks in the future, separately against telehealth calls

Table 5.3. Telehealth vomiting call thresholds and start week for winter norovirus activity; values by standard deviation increment used for thresholds.

	Mean + 2 SD		Mean + 3 SD		Mean + 4 SD	
Year	Week	# calls	Week	# calls	Week	# calls
2011	44	462	44	462	44	462
2012	42	433	42	433	43	465
2013	43	419	44	474	44	474
2014	45	444	45	444	45	444

FIGURES

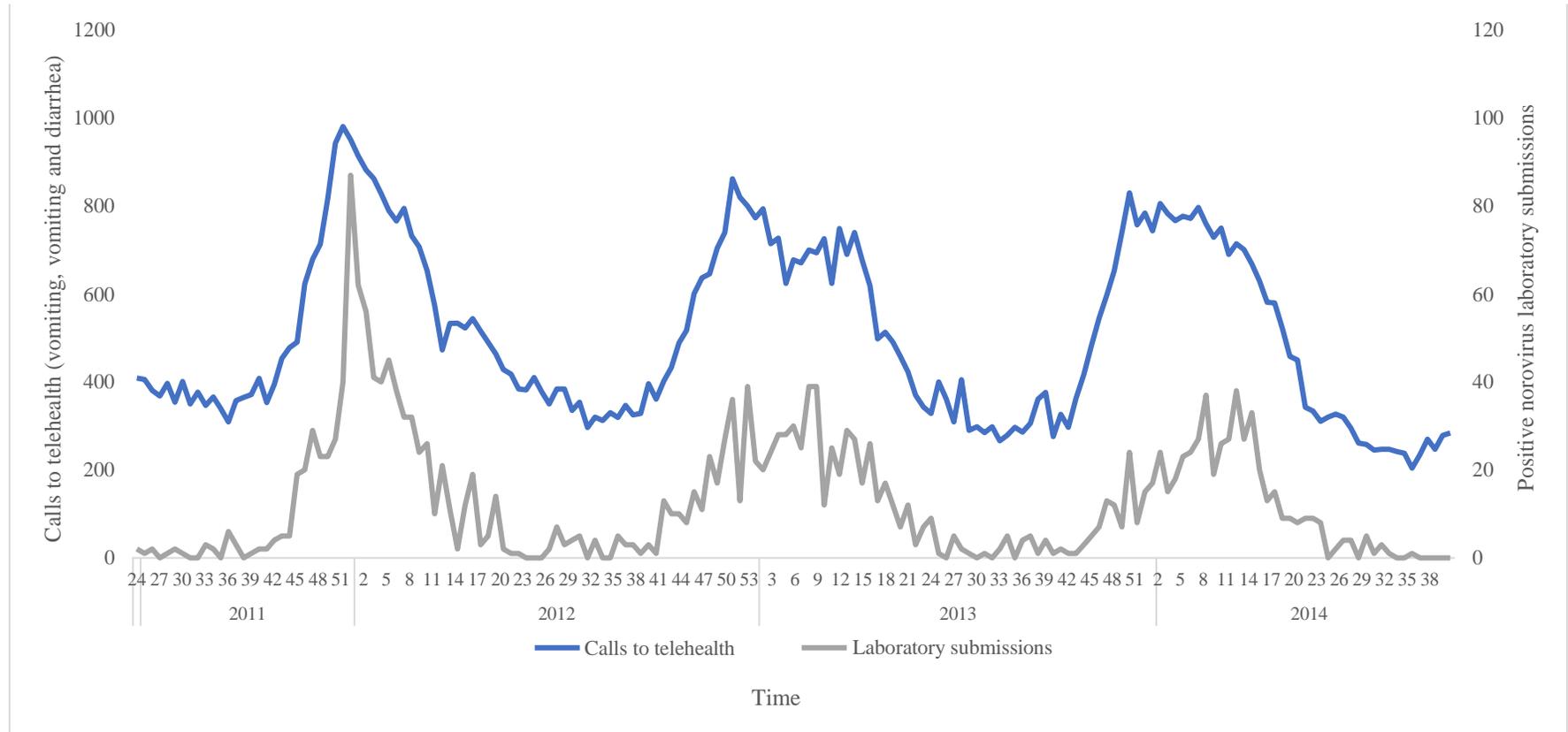


Figure 5.1. Weekly vomiting telehealth calls (“vomiting” and “vomiting and diarrhea” chief complaints) and laboratory positive norovirus submissions, 17/6/2011 – 30/9/2014.

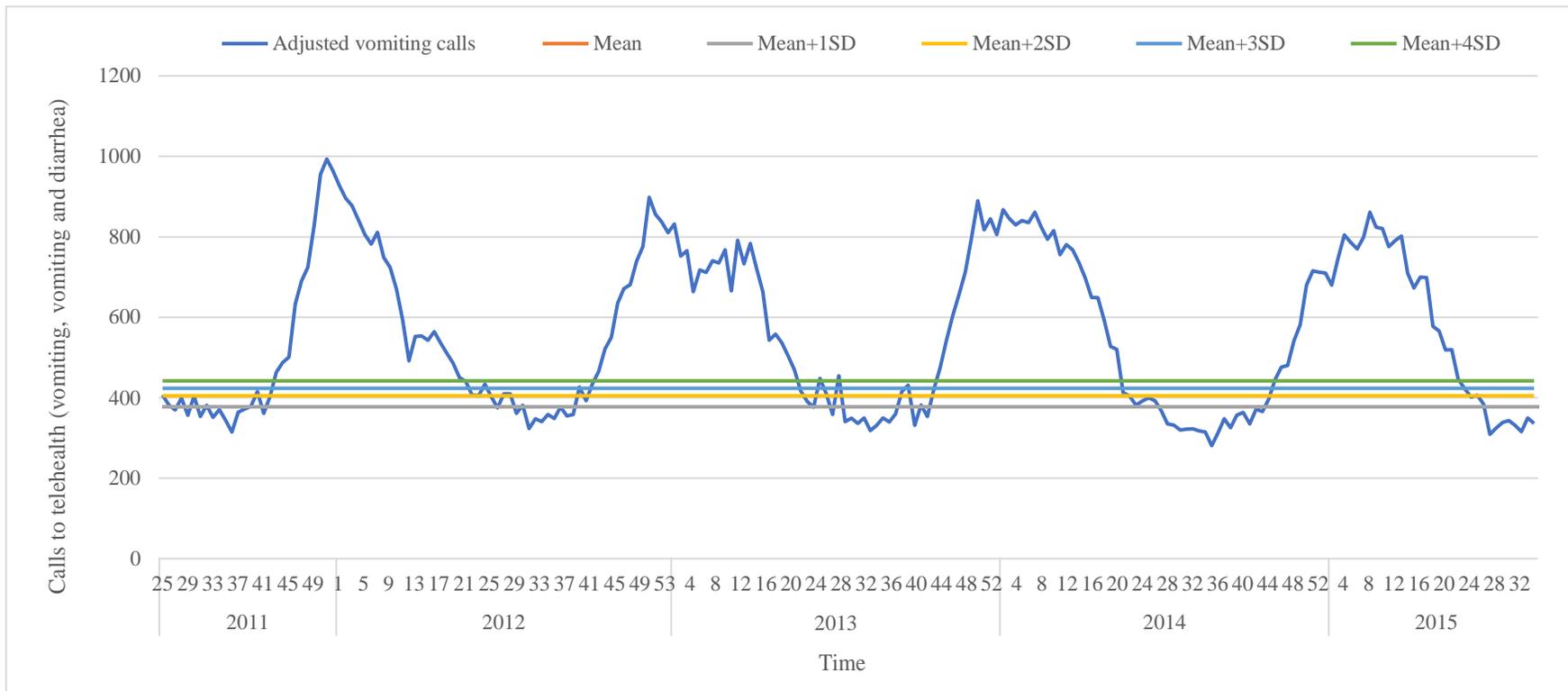


Figure 5.2. Weekly telehealth vomiting calls adjusted to remove long-term trend, with the ‘baseline’ mean number of calls and standard deviations marked, 19/6/2011 – 29/8/2015.

CHAPTER 6

Conclusion

Norovirus is the most common cause of gastroenteritis in countries worldwide. It is estimated to be the etiologic agent of more than 50% of all-cause gastroenteritis outbreaks, and more than 90% of non-bacterial outbreaks (1–6). Initially described as the agent of “winter vomiting disease”, noroviruses have now been identified and described structurally and genetically; however, the original name remains due to its pronounced winter seasonality. The self-limiting nature and short duration of the illness discourage many sufferers from seeking healthcare advice when ill, resulting in greatly under-reported case numbers (7,8). In addition, in the event of a patient consult, there is a small probability of the clinician requesting a clinical sample for laboratory confirmatory testing. Current laboratory and outbreak surveillance methods fail to capture the full scope of illness due to these under-reporting issues, and are biased towards vulnerable populations seeking physician assistance (i.e. $<5, \geq 65$ years) (7–12).

Syndromic surveillance is a newer form of disease surveillance which aims to capture a wider scope of illness and detect cases earlier than laboratory and outbreak methods (13–15). It utilizes pre-diagnostic and pre-existing data sources to monitor the earliest points of medical care individuals may seek when ill. Syndromic surveillance was initially developed to monitor potential bioterrorism events, but has now been adapted to a variety of communicable diseases as well as environmental health events (15–17). Due to its improved representation of community disease activity, it has been used as an effective addition to strengthen existing laboratory and outbreak surveillance methods and to detect under-reported illnesses.

MAJOR FINDINGS

The studies described in this thesis aimed to describe syndromic surveillance systems, norovirus activity in the province of Ontario, Canada, and develop a novel syndromic surveillance system in Ontario for the winter norovirus season. Initially, a scoping review was conducted to better understand all necessary components of a syndromic surveillance system for human infectious disease, including the definition, development, implementation, and evaluation, and to help identify research gaps. The review identified various gaps and described all elements in detail for those wishing to create a syndromic surveillance system from scratch. Work performed in this thesis also sought to analyze norovirus activity in Ontario during the years 2009 to 2014. All of the above information was applied to construct a syndromic surveillance system for the winter norovirus season in Ontario. Creating this system will help to better understand the full community burden of illness in Ontario and detect the winter norovirus season earlier than currently available surveillance methods. Descriptive statistics were used to assess data from 2009 to 2014 to properly design and adjust the syndromic surveillance system for the winter vomiting season in Ontario; the in- and out-of-season winter vomiting season periods were defined from these data. Negative binomial regression modeling and Shewhart (control) chart methods were applied to telehealth data alongside norovirus laboratory data from 2009 to 2014. The goal of this was to identify whether the telehealth data predicted the laboratory data, and, if so, by how many weeks. In these analyses, specific alarm thresholds (i.e. week numbers and call volumes) of winter norovirus season onset were devised.

The creation of the novel syndromic surveillance system in Chapter 5 demonstrated that vomiting calls to telehealth can be used as an appropriate proxy for norovirus activity, and can effectively identify the onset the winter norovirus season. The negative binomial regression analyses showed that the telehealth data provided two weeks' early warning ahead of the laboratory data. This pivotal result shows how much timelier syndromic data can be in comparison with non-syndromic (traditional, e.g. laboratory) data, provides a better understanding of the winter norovirus season in Ontario, and helps fill gaps in community data. Using only laboratory (traditional) data, data gaps exist in norovirus illness reporting and biases toward institutional norovirus activity become apparent. The telehealth data used in this study are more representative of the population and provide a more complete picture of norovirus illness in the province. Most importantly, determining statistically that vomiting calls to telehealth provide early warning of the winter norovirus season in Ontario creates a basis for a novel surveillance system to be utilized prospectively. The result of Chapter 5 in this thesis demonstrates that the early warning system developed in this thesis could be effectively used to detect the onset of the winter norovirus season and reduce the impact of illness for years to come.

In applying this novel syndromic surveillance system, the impact of the winter norovirus season can be mitigated; areas of high-risk (i.e. hospitals, long-term care homes, day cares) can be notified immediately after an early alarm threshold is surpassed and infection control measures can be implemented. These measures function to decrease the impact of disease on the most vulnerable populations in society. These results emphasize the ability of syndromic surveillance to strengthen and improve existing laboratory and outbreak reporting methods, and its ability to

reduce the impact of disease. The THO data are already collected on a daily basis by nurses at THO, and can effectively be utilized for surveillance purposes.

There were some notable findings identified from the study performed in Chapter 2: the fact that the majority of syndromic surveillance systems for human infectious disease research exists in the U.S., and that the definition and evaluation of syndromic surveillance are not well-described. The field of syndromic surveillance is a relatively new area of research; however, interest is rapidly expanding. The distribution of research performed internationally was unequal, with gaps identified in places outside the U.S., Canada, and select European countries. This may be a result of disease priorities in lower income versus higher income countries. Differences in research areas were also noted, with the development and implementation of syndromic surveillance systems researched much more than the definition and evaluation. This may be a result of identified research priorities in the field, i.e. the development and implementation seen as more pertinent. In addition, predictable winter activity of norovirus was observed in all three datasets analyzed (telehealth, laboratory, and outbreak), and long-term care homes were the most affected institution type in the province.

STRENGTHS

Several notable strengths were present throughout this study. These strengths facilitated the success of the research, and allowed for all research objectives to be met. A sufficient number of peer-reviewed and grey literature studies were synthesized to successfully perform a scoping review on the creation of syndromic surveillance systems for human infectious disease. These studies were available and allowed for research gaps to be identified, and for key elements to be

described. With the scoping review, online software (Covidence) helped facilitate the process and vastly improved the efficiency of the scoping review. Non-anonymized laboratory and outbreak data were available for request in Ontario, with the assistance of Public Health Ontario. Telehealth data had been routinely collected and entered electronically in Ontario since 2006, allowing such data to be readily available for use in a syndromic surveillance system. Utilizing these data helped to improve the representativeness of norovirus surveillance in Ontario. Most notably, the studies in this thesis integrated various datasets to inform public health authorities on the status of norovirus in Ontario and create a novel syndromic surveillance system.

The results from the studies in this thesis show which countries worldwide are the most active in research regarding the creation of syndromic surveillance systems for human infectious disease, and where improvements can be made. Additionally, the descriptive analyses of norovirus in Ontario from 2009 to 2014 will help to reduce the economic burden of disease with targeted health promotion campaigns. Lastly, the two weeks early warning demonstrates the significant future potential for public health in Ontario to take earlier action on the winter norovirus season.

LIMITATIONS

Limitations existed in this study, although they did not affect the ability to determine outcomes. There was a clear bias in country of study publication in the scoping review, with the majority of studies being published in the U.S. The laboratory dataset was unavailable prior to 2009 as the data had not yet been entered electronically. In addition, certain variables within the laboratory dataset changed over the 2009-2014 period, complicating analyses. The outbreak dataset did not contain demographic data, and included a great deal of undiagnosed “gastroenteritis

unspecified/other” entries, which could have encompassed many additional entries of norovirus. The syndromic dataset (telehealth) had the greatest limitations, including varying formats of variables, missing data, open-ended entries, low daily call counts (which were subsequently aggregated to weekly counts), as well as vomiting calls used as a proxy for norovirus and could not be officially attributed to the illness.

Although limitations existed, they did not prevent meaningful results from being elucidated from this study. This thesis describes the great potential for syndromic surveillance in Ontario and how existing health service data can be applied for disease surveillance purposes. The field of syndromic surveillance is still new in the province, with much room to grow. Other countries, such as the U.S. and U.K., have developed and implemented a variety of systems addressing several disease syndromes and environmental health events using multiple data sources.

Telehealth data are collected in the province on a daily basis and should be applied for public health purposes. By developing additional surveillance systems and integrating information with existing laboratory and outbreak methods, our knowledge of disease activity in the population of Ontario will improve, as will disease mitigation activities.

FUTURE RESEARCH

Using the results of this study, there are a few areas of future research to be considered. These areas would improve norovirus, and additional disease and adverse health event, surveillance in Ontario. These future research aims would also help to target health promotion campaigns for norovirus and reduce the burden of disease (including economic).

- Analyzing norovirus laboratory submissions, outbreaks, and calls to THO in Ontario by geographic region, i.e. public health unit. More granular descriptive statistics will help to inform public health authorities on norovirus patterns and “hotspots” throughout the province. More granular (spatial) early warning thresholds could therefore be devised.
- Using daily THO data in the creation of early warning thresholds. This would allow for more precise (temporal) early alarm thresholds.
- Implementing the results from this study with existing Ontario public health laboratory and outbreak surveillance activities to strengthen norovirus surveillance capacities in Ontario. This would improve community coverage, above current capabilities.
- Expanding this surveillance system to include a variety of syndromes, encompassing both communicable diseases and environmental health events. This would mimic the successful syndromic surveillance systems developed worldwide. In doing so, a wide array of illnesses and adverse events would be detected early, and Ontario public health would have great situational awareness across various sectors.
- Expanding this surveillance system to include additional syndromic datasets, such as pharmacy purchases and/or internet queries. There are numerous additional syndromic datasets which are being utilized by public health authorities worldwide. Each one of the datasets has its own benefits and contributes to disease detection. This would expand the number of potential illnesses identified by the system, and increase situational awareness.
- Conducting economic analyses on the cost of norovirus in long-term care homes in Ontario. Few economic analyses have been conducted on norovirus in Ontario, and none have estimated the burden in long-term care homes, where the illness disproportionately affects residents.

- Cost-benefit analyses of implementing syndromic surveillance in Ontario with existing laboratory and outbreak surveillance systems to justify the long-term benefits of syndromic surveillance; and
- Confirmatory testing on callers to THO similar to studies conducted in other countries, including the U.K., where callers to telehealth have given consent and mailed stool sample containers to confirm the aetiology of their gastrointestinal distress symptoms. Performing these analyses would allow public health authorities to confirm the vomiting calls to THO as norovirus, rather than merely as a proxy for the disease.

The studies performed in this thesis successfully evaluated the current state of research in the creation of syndromic surveillance systems. Gaps in research areas from countries worldwide were identified, and important points in the creation of syndromic surveillance systems were comprised into one document for use by public health practitioners. The studies in this thesis also successfully described norovirus activity in Ontario, Canada from 2009 to 2014, which was applied to create a novel telehealth syndromic surveillance system for the winter norovirus season in the province. Although norovirus illness has plagued humans for many years to date, the results of this thesis give solutions to improve the monitoring of the disease and decrease its impact.

REFERENCES

1. Patel MM, Hall AJ, Vinjé J, Parashar UD. Noroviruses: A comprehensive review. *J Clin Virol.* 2009;44(1):1–8.
2. van Duynhoven YTHP, de Jager CM, Kortbeek LM, Vennema H, Koopmans MPG, van

- Leusden F, et al. A one-year intensified study of outbreaks of gastroenteritis in The Netherlands. *Epidemiol Infect.* 2005;133(1):9–21.
3. Buesa J, Collado B, López-Andújar P, Abu-Mallouh R, Diaz JR, Diaz AG, et al. Molecular epidemiology of Caliciviruses causing outbreaks and sporadic cases of acute gastroenteritis in Spain. *J Clin Microbiol.* 2002;40(8):2854–9.
 4. Ike AC, Brockmann SO, Hartelt K, Marschang RE, Contzen M, Oehme RM. Molecular epidemiology of norovirus in outbreaks of gastroenteritis in Southwest Germany from 2001 to 2004. *J Clin Microbiol.* 2006;44(4):1262–7.
 5. Reuter G, Krisztalovics K, Vennema H, Koopmans M, Szucs G. Evidence of the etiological predominance of norovirus in gastroenteritis outbreaks - Emerging new-variant and recombinant noroviruses in Hungary. *J Med Virol.* 2005;76(4):598–607.
 6. Lopman BA, Reacher MH, Duijnhoven Y Van, Hanon F-X, Brown D, Koopmans M. Viral gastroenteritis outbreaks in Europe, 1995 – 2000. *Emerg Infect Dis.* 2003;9(1):90–6.
 7. Hall AJ, Rosenthal M, Gregoricus N, Greene SA, Ferguson J, Henao OL, et al. Incidence of acute gastroenteritis and role of norovirus, Georgia, USA, 2004–2005. *Emerg Infect Dis.* 2011;17(8):1381–8.
 8. Bernard H, Werber D, Höhle M. Estimating the under-reporting of norovirus illness in Germany utilizing enhanced awareness of diarrhoea during a large outbreak of Shiga toxin-producing *E. coli* O104:H4 in 2011 - a time series analysis. *BMC Infect Dis.* 2014;14:116.
 9. Thomas MK, Murray R, Flockhart L, Pintar K, Pollari F, Fazil A, et al. Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne Pathog Dis.* 2013 Jul;10(7):639–48.

10. Yen C, Hall AJ. Challenges to estimating norovirus disease burden. *J Pediatric Infect Dis Soc.* 2013;2(1):61–2.
11. Martinez A, Dominguez A, Torner N, Ruiz L, Camps N, Barrabeig I, et al. Epidemiology of foodborne Norovirus outbreaks in Catalonia, Spain. *BMC Infect Dis.* 2008;8:47.
12. MacCannell T, Umscheid CA, Agarwal RK, Lee I, Kuntz G, Stevenson KB, et al. Guidelines for the Prevention and Control of Norovirus Gastroenteritis Outbreaks in Healthcare Settings. Centers for Disease Control and Prevention. 2011.
13. Leonardi GS, Hajat S, Kovats RS, Smith GE, Cooper DL, Gerard E. Syndromic surveillance use to detect the early effects of heat-waves: An analysis of NHS direct data in England. *Soz Praventiv Med.* 2006;51(4):194–201.
14. 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):pii=19806.
15. Triple-S Project. Guidelines for designing and implementing a syndromic surveillance system. 2013.
16. Yan P, Chen H, Zeng D. Syndromic surveillance systems. In: *Annual Review of Information Science and Technology.* University of Arizona; 2009. p. 425–95.
17. Henning KJ. What is syndromic surveillance? *MMWR Suppl.* 2004;53:7–11.

APPENDICES

Appendix 2.1. Detailed protocol for scoping review

A. Full search strategy used in scoping review searches

Peer-reviewed literature

1. Scopus

March 29, 2016

Search string: (TITLE-ABS-KEY (("syndromic surveillance" OR "biosurveillance" OR "situational awareness" OR "prodrome surveillance" OR "outbreak detection system" OR "symptom-based surveillance" OR "early warning system" OR "early detection system")) AND TITLE-ABS-KEY ((infectious OR outbreak OR disease OR enteric OR gastroin*)) AND TITLE-ABS-KEY ((defin* OR implement* OR review))) AND PUBYEAR > 2003

Hits: 595

2. Pubmed

March 29, 2016

Search string: ("syndromic surveillance" OR "biosurveillance" OR "situational awareness" OR "prodrome surveillance" OR "outbreak detection system" OR "symptom-based surveillance" OR "early warning system" OR "early detection system") AND (infectious OR outbreak OR disease OR enteric OR gastroin*) AND (defin* OR implement* OR review)

Hits: 375

3. CINAHL

March 29, 2016

Selected “Boolean/phrase search” and “apply related words”

Search string: ("syndromic surveillance" OR "biosurveillance" OR "situational awareness" OR "prodrome surveillance" OR "outbreak detection system" OR "symptom-based surveillance" OR "early warning system" OR "early detection system") AND (infectious OR outbreak OR disease OR enteric OR gastroin*) AND (defin* OR implement* OR review)

Hits: 108

4. ProQuest (Dissertations & Theses A&I)

March 29, 2016

Subjects: Public health & epidemiology

Search string: ("syndromic surveillance" OR biosurveillance OR "situational awareness" OR "prodrome surveillance" OR "outbreak detection system" OR "symptom-based surveillance" OR "early warning system" OR "early detection system") AND (infectious OR outbreak OR disease) AND (defin* OR implement* OR review) AND (enteric OR gastro*)

Hits: 365 (public health) & 199 (epidemiology)s

Grey literature

1. Google search (1) (www.google.ca)

March 30, 2016

Search string: “syndromic surveillance”

Hits: 212, 000

Limited to the first 100 (most relevant) hits

Novel citations created: 10

2. Google search (2) (www.google.ca)

March 30, 2016

Search string: “syndromic surveillance framework”

Hits: 356, 000

Limited to the first 100 (most relevant) hits

Novel citations created: 12

3. International Society for Disease Surveillance (www.syndromic.org)

March 30 2016

Search strings: “syndromic surveillance”, “syndromic surveillance framework”, “syndromic surveillance system”

Novel citations created: 2

4. Public Health Ontario (www.publichealthontario.ca)

March 30 2016

Search strings: “syndromic surveillance”

Novel citations created: 1

5. Eurosurveillance (www.eurosurveillance.org)

March 30, 2016

Search strings: “syndromic surveillance”

Novel citations created: 0

6. Syndromic Surveillance Systems in Europe (www.syndromicsurveillance.eu)

March 31 2016

Search strings: Scanned through “publications” page

Novel citations created: 8

7. Public Health Agency of Canada (www.phac-aspc.gc.ca)

April 1 2016

Search strings: “Syndromic surveillance”

Novel citations created: 0

8. Kingston, Frontenac, Lennox and Addington Public Health (www.kflapublichealth.ca)

April 1 2016

Search strings: “Syndromic surveillance”

Novel citations created: 0

9. Kingston, Frontenac, Lennox and Addington Public Health Informatics (www.kflaphi.ca)

April 1 2016

Search strings: “Syndromic surveillance”

Novel citations created: 1

10. Centers for Disease Control and Prevention (www.cdc.gov)

April 4 2016

Search strings: “Syndromic surveillance”

Novel citations created: 1

11. European Centers for Disease Prevention and Control (ecdc.europa.eu)

April 4 2016

Search strings: “Syndromic surveillance”

Novel citations created: 0

12. Public Health England (www.gov.uk/government/organisations/public-health-england)

April 4 2016

Search strings: “Syndromic surveillance” on main page and “publications” page

Novel citations created: 0

B. Relevance screening form

1. Does the document pertain to the concept of syndromic surveillance?

Yes

No

2. Is the document a primary research article, review paper, book chapter, government report, thesis, conference paper, or commentary?

Yes

No

3. Does the article pertain to human infectious disease?

Yes

No

4. Is the document written in English?

Yes

No

5. Was the document published after 2003?

Yes

No

6. Does the document describe the individual elements/framework of a syndromic surveillance system?

Yes

No

7. What elements does the document describe?

Statistical methods

Costs (HR, IT, responding to alarms, operating, R&D, acquiring data -- staff, resources)

Classification of syndromes

Naming "syndromic surveillance"

Data standards (i.e. timeliness)

Privacy & confidentiality issues

Operational challenges

Data quality (representativeness, completeness)

Alarm thresholds

Data sources

Stakeholder identification & participation

Management composition

Collaboration(s) with other public health surveillance methods

Data structure (i.e. chief complaint)

Development of alert response protocol

Evaluation criteria

8. If “data sources” was checked off in Q. 7, please answer the following; if it was left blank, please move on to Q. 9. What data source(s) of syndromic surveillance were described in the document?

School absenteeism

Calls to a telehealth line chief complaints

OTC pharmacy purchases

Prescription sales

General practitioner visits (in hours & out-of-hours)

Emergency department chief complaints

Employee absenteeism

Sewage pathogen testing

Repeated interval web surveys

Calls to 911/emergency medical services chief complaints

Other non-emergency health records (travel clinics, hospital discharges, health plan data, mortality data)

Search engine queries

Social media posts

Other: _____

9. What term(s) does the document use to refer to syndromic surveillance?

Syndromic surveillance

Biosurveillance

Situational awareness

Outbreak detection system

Prodrome surveillance

Symptom-based surveillance

Early warning system

Early detection system

10. Does the document explicitly define the term provided in Q. 9?

Yes

No

11. What is the overall focus of the document?

Defining syndromic surveillance

Uses of syndromic surveillance

Pros/cons of syndromic surveillance

Elements of syndromic surveillance systems

Creating a syndromic surveillance system

Implementing syndromic surveillance systems

Evaluating syndromic surveillance systems

Data sources used in syndromic surveillance systems

Other: _____

12. Does the document describe the uses of syndromic surveillance?

Yes

No

13. If the answer to Q. 12 was “yes”, please answer the following; if the answer was “no”, please move on to Q. 14. What use(s) of syndromic surveillance was mentioned in the document?

Early warning (early onset of expected health event)

Detection (discovery of unknown health threat)

Situational awareness (track progression of or lack of activity of known health threat)

14. Does the document describe the benefits of syndromic surveillance?

Yes

No

15. If the answer to Q. 14 was “yes”, please answer the following; if the answer was “no”, please move on to Q. 16. What benefits of syndromic surveillance were mentioned in the document?

Cost savings (use of existing data sources and personnel)

Timeliness of analyses and reporting (data analysed in real-time or near-real-time)

Fosters collaboration (PH agencies, academics, health care providers, industry, etc.)

Strengthens/complements traditional surveillance methods

Can confirm alerts generated by other public health systems/agencies

Follows the course of an outbreak and provides an estimated case count (particularly useful for under-reported illnesses)

Non-specific (can be used to identify a broad range of infectious diseases)

Signals and alerts allow for early public health action

Versatile -- can be applied to various other health events and natural disasters (i.e. heat waves, earthquakes)

Useful for both short- and long-term surveillance

Cases identified can be used for microbiological testing

Non-biased

Other: _____

16. Does the document describe the limitations of syndromic surveillance?

Yes

No

17. If the answer to Q. 16 was “yes”, please answer the following; if the answer was “no”, please move on to Q. 18. What limitations of syndromic surveillance were mentioned in the document?

Does not replace traditional surveillance methods (i.e. laboratory reports)

Does not replace traditional health care (i.e. physician visits)

Cannot be relied on as a sole outbreak detection tool

False signals/alerts may occur

Less effective at detecting single cases or smaller-scale outbreaks

Non-specific (cannot be used to identify or verify a particular infectious disease)

Issues with accessing and acquiring data (privacy issues)

Poor data quality

Time lags (i.e. physician diagnoses, laboratory reporting, data acquisition)

Data not always available

Somewhat costly (resources not universally available, i.e. developing countries)

Other: _____

18. What kind of paper is the document?

Primary research article

Review article

Formal document (government report)

Commentary

Book chapter

Conference proceeding

Thesis

19. If “primary research article” was checked off in Q. 18, please answer the following; if it was left blank, please move on to Q. 20. Which study design was used in the document?

Observational (cross-sectional, cohort, case-control)

Experimental (randomized control, non-randomized control, controlled before-and-after, non-controlled before-and-after)

Qualitative

Other: _____

20. What is the publication year of the document?

21. Where was the document written?

North America

South America

Europe

Africa

Asia

Australasia

Unknown

Appendix 2.2. List of all 55 articles selected for inclusion in scoping review

1. Aller RD. Pathology's contributions to disease surveillance: Sending our data to public health officials and encouraging our clinical colleagues to do so. *Arch Pathol Lab Med*. 2009;133(6):926–32.
2. Berger M, Shiau R, Weintraub JM. Review of syndromic surveillance: Implications for waterborne disease detection. *J Epidemiol Community Health*. 2006;60(6):543–50.
3. Buckeridge DL. Outbreak detection through automated surveillance: A review of the determinants of detection. *J Biomed Inform*. 2007;40(4):370–9.
4. Buehler JW. Review of the 2003 National Syndromic Surveillance Conference--lessons learned and questions to be answered. *MMWR*. 2004;53 Suppl(May):18–22.
5. Buehler JW, Hopkins R, Overhage JM, Sosin D, Tong V. Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks. *MMWR*. 2004;53:1–13.
6. Buehler JW, Whitney EA, Smith D, Prietula MJ, Stanton SH, Isakov AP. Situational uses of syndromic surveillance. *Biosecurity and Bioterrorism*. 2009;7(2):165–77.
7. Chapman WW, Christensen LM, Wagner MM, Haug PJ, Ivanov O, Dowling JN, et al. Classifying free-text triage chief complaints into syndromic categories with natural language processing. *Artif Intell Med*. 2004;33(1):31–40.
8. Chen D, Cunningham J, Moore KM, Tian J. Spatial and temporal aberration detection methods for disease outbreaks in syndromic surveillance systems. *Ann GIS*. 2011;17(4):211–20.
9. Chretien J-P, Burkom HS, Sedyaningsih ER, Larasati RP, Lescano AG, Mundaca CC, et al. Syndromic surveillance: Adapting innovations to developing settings. *PLoS Med*. 2008;5(3):367–72.

10. Conway M, Dowling JN, Chapman WW. Using chief complaints for syndromic surveillance: A review of chief complaint based classifiers in North America. *J Biomed Inform.* 2013;46(4):734–43.
11. Ding Y, Fei Y, Xu B, Yang J, Yan W, Diwan VK, et al. Measuring costs of data collection at village clinics by village doctors for a syndromic surveillance system—a cross sectional survey from China. *BMC Health Serv Res.* 2015;15(1):287.
12. Ding Y, Sauerborn R, Xu B, Shaofa N, Yan W, Diwan VK, et al. A cost-effectiveness analysis of three components of a syndromic surveillance system for the early warning of epidemics in rural China Health policies, systems and management in low and middle-income countries. *BMC Public Health.* 2015;15(1).
13. Drociuk D, Gibson J, Hodge Jr J. Health information privacy and syndromic surveillance systems. *MMWR.* 2004;53 Suppl(Sept):24–28.
14. Forslund DW, Joyce EL, Burr T, Picard R, Wokoun D, Umland E, et al. Setting standards for improved syndromic surveillance. *IEEE Eng Med Biol Mag.* 2004;23(1):65–70.
15. Fricker RDJ, Hegler BL, Dunfee DA. Comparing syndromic surveillance selection methods: EARS versus a CUSUM-based methodology. *Stat Med.* 2008;27(17):3407–29.
16. Generous N, Margevicius KJ, Taylor-McCabe KJ, Brown M, Daniel WB, Castro L, et al. Selecting essential information for biosurveillance - A multi-criteria decision analysis. *PLoS One.* 2014;9(1).
17. Gerbier S, Yarovaya O, Gicquel Q, Millet A-L, Smaldore V, Pagliaroli V, et al. Evaluation of natural language processing from emergency department computerized medical records for intra-hospital syndromic surveillance. *BMC Med Inform Decis Mak.* 2011;11(1).
18. Gesualdo F, Stilo G, Agricola E, Gonfiantini MV, Pandolfi E, Velardi P, et al. Influenza-

like illness surveillance on Twitter through automated learning of naive language. *PLoS ONE*. 2013;8(12):e82489.

19. Haas SW, Travers D, Waller A, Mahalingam D, Crouch J, Schwartz TA, et al. Emergency Medical Text Classifier: New system improves processing and classification of triage notes. *Online J Public Health Inform*. 2014;6(2):e178.
20. Hartley DM, Nelson NP, Arthur RR, Barboza P, Collier N, Lightfoot N, et al. An overview of internet biosurveillance. *Clin Microbiol Infect*. 2013;19(11):1006–13.
21. Henning K.J. What is syndromic surveillance?. *New York City Dep Heal Ment Hyg*. 2004;53:1–9.
22. Henriques G, Stacey D. An ontology-based framework for syndromic surveillance method selection. *4th Int Conf Knowl Eng Ontol Dev KEOD 2012*. 2012;396–400.
23. Hu, J. Privacy-preserving data integration in public health surveillance. *ProQuest Dissertations and Theses*. [Ann Arbor]: University of Ottawa (Canada); 2011.
24. Hurt-Mullen KJ, Coberly J. Syndromic surveillance on the epidemiologist's desktop: making sense of much data. *MMWR*. 2005;54 Suppl:141–6.
25. Katz R, May L, Baker J, Test E. Redefining syndromic surveillance. *J Epidemiol Glob Health*. 2011;1(1):21–31.
26. Kaydos-Daniels SC, Smith LR, Farris TR. Biosurveillance in outbreak investigations. *Biosecurity and Bioterrorism*. 2013;11(1):20–8.
27. Kirian ML, Weintraub JM. Is syndromic surveillance of OTC drug sales effective in detecting outbreaks of waterborne GI disease? *Am Water Work Assoc*. 2011;103(8):63–73.
28. Kirkwood A, Guenther E, Fleischauer AT, Gunn J, Hutwagner L, Barry MA. Direct cost associated with the development and implementation of a local syndromic surveillance system. *J*

Public Heal Manag Pract. 2007;13(2):194–9.

29. Lawson BM, Fitzhugh EC, Hall SP, Franklin C, Hutwagner LC, Seeman GM, et al. Multifaceted syndromic surveillance in a public health department using the Early Aberration Reporting System. *J Public Heal Manag Pract.* 2005 Jul;11(4):274–281 8p.
30. Lescano AG, Larasati RP, Sedyaningsih ER, Bounlu K, Araujo-Castillo R V, Munayco-Escate C V, et al. Statistical analyses in disease surveillance systems. *BMC Proc.* 2008;2(Suppl 3):S7.
31. Magruder SF, Henry J, Snyder M. Linked analysis for definition of nurse advice line syndrome groups, and comparison to encounters. *MMWR.* 2005;54 Suppl:93–7.
32. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience. 2004;11(2):141–50.
33. Moore KM. Real-time syndrome surveillance in Ontario, Canada: the potential use of emergency departments and Telehealth. *Eur J Emerg Med.* 2004;11(1):3–11.
34. Najmi A-H, Burkom H. Recursive least squares background prediction of univariate syndromic surveillance data. *BMC Med Inform Decis Mak.* 2009;9(1):4.
35. Pavlin JA. Syndromic surveillance for infectious diseases. In: *Infectious Disease Surveillance: Second Edition.* 2nd ed. 2013. p. 470–81.
36. Public Health Ontario. *Syndromic Surveillance Discussion Paper.* Provincial Infectious Diseases Advisory Committee, Toronto, ON; 2012.
37. Shmueli G, Burkom H. Statistical Challenges Facing Early Outbreak Detection in Biosurveillance. *Technometrics.* 2010;52(1):39–51.
38. Sintchenko V, Gallego B. Laboratory-guided detection of disease outbreaks: Three

- generations of surveillance systems. *Arch Pathol Lab Med.* 2009;133(6):916–25.
39. Sniegowski CA. Automated syndromic classification of chief complaint records. 2004;25(1):68–75.
 40. Stoto MA. Syndromic surveillance. *Issues Sci Technol.* 2005;21(3):49–56
 41. Stoto MA, Fricker RD, Jain A, Diamond A, Davies-Cole JO, Glymph C, et al. Evaluating statistical methods for syndromic surveillance. In: *Statistical Methods in Counterterrorism: Game Theory, Modeling, Syndromic Surveillance, and Biometric Authentication.* 2006. p. 141–72.
 42. Szarka III JL, Gan L, Woodall WH. Comparison of the early aberration reporting system (EARS) W2 methods to an adaptive threshold method. *Stat Med.* 2011;30:489–504.
 43. Triple-S Project. About data sources for syndromic surveillance. 2013.
 44. Triple-S Project. Guidelines for designing and implementing a syndromic surveillance system. 2013.
 45. Tsui K-L, Chiu W, Gierlich P, Goldsman D, Liu X, Maschek T. A review of healthcare, public health, and syndromic surveillance. *Qual Eng. H.* 2008;20(4):435–50.
 46. Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the prospective detection of infectious disease outbreaks: A review. *J R Stat Soc.* 2012;175(1):49–82.
 47. Uscher-Pines L, Farrell C, Cattani J, Moskal MD, Babin SM, Charlotte A, et al. Framework for the development of response protocols for public health syndromic surveillance systems. 2008;53:2008.
 48. 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. *Eurosurveillance*. 2011;16(9):1–8.

49. VanStone N, Wu C-C, Van-Dijk A, Belanger P, Moore KM. Syndromic surveillance report: Validation of the EDSS system's categorization of syndromes. 2015.

50. Velardi P, Stilo G, Tozzi AE, Gesualdo F. Twitter mining for fine-grained syndromic surveillance. *Artif Intell Med*. 2014;61:153–63.

51. Yan P, Chen H, Zeng D. Syndromic surveillance systems. *Annu Rev Inf Sci Technol*. 2009;42(1):425–95.

52. Yan P, Zeng D, Chen H. A review of public health syndromic surveillance systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2006. p. 249–60.

53. Yih WK, Caldwell B, Harmon R, Kleinman K, Lazarus R, Nelson A, et al. National Bioterrorism Syndromic Surveillance Demonstration Program. *Morb Mortal Wkly Rep*. 2004;53(Supplement):43–9.

54. Zhang Y, Dang Y, Chen H, Thurmond M, Larson C. Automatic online news monitoring and classification for syndromic surveillance. *Decis Support Syst*. Elsevier B.V.; 2009;47:508–17.

55. Ziemann A, Rosenkötter N, Riesgo LG-C, Fischer M, Krämer A, Lippert FK, et al. Meeting the International Health Regulations (2005) surveillance core capacity requirements at the subnational level in Europe: The added value of syndromic surveillance. *BMC Public Health*. 2015;15(1):107.

Appendix 4.1 Details of three datasets used in norovirus descriptive statistics analyses

i) integrated Public Health Information System (iPHIS)

The iPHIS dataset represents the confirmed outbreaks of viral gastroenteritis in Ontario reported to public health. All outbreaks are institutional, i.e. occurred at long-term care homes, retirement homes, child care facilities, correctional facilities, etc. A confirmed gastroenteritis outbreak is defined as, “three or more cases with signs and symptoms compatible with infectious gastroenteritis in a specific unit or floor within a four-day period OR three or more units/floors having a case of infectious gastroenteritis within 48 hours” (1). Additional information found in iPHIS includes the method of exposure (if determined), aetiologic agent (if determined), public health unit, disease confirmed status, date the outbreak was declared by the medical officer of health, date the outbreak was declared over by the medical officer of health, the initial onset date (time of index case), and final onset date (time of last case). For analyses, the onset date (time of index case) dates were used to analyze the VGE and norovirus outbreaks from 2009-2014, and the range between the onset date (time of index case) and onset date (time of last case) was used to calculate average VGE and norovirus outbreak duration.

While iPHIS data are continuously updated, public health authorities are required to enter all outbreaks from the past year by each August; therefore, incurring time delays. It should be noted laboratory confirmation is not required for an outbreak to be entered into the dataset, and when laboratory confirmation is present not all cases associated with an outbreak are laboratory tested.

ii) Public Health Ontario Laboratories (PHOL)

The PHOL dataset represents all the samples submitted to Public Health Ontario (PHO) for testing in Ontario. The data represent samples which were submitted to PHO from patients ill with suspected norovirus and/or rotavirus (i.e. symptoms of vomiting and/or diarrhea). Samples are submitted by medical professionals in the form of stool samples and are either tested by polymerase chain reaction, electron microscopy, or immunochromatographic test. The dataset provides age, gender, and public health unit information, as well as the dates the samples were collected and subsequently entered into the dataset. For analyses, the sample collection dates were used (i.e. the results represent the date the samples were collected from the ill person by the healthcare practitioner for testing). It should be noted not all VGE test requisitions from institutions and physicians are tested by PHO in Ontario; samples may also be sent to other private labs in Ontario, or in duplicates to multiple labs. As a result, not all laboratory-confirmed samples of VGE from Ontario are captured in this dataset.

iii) Telehealth Ontario (THO)

The THO dataset represents all phone calls made by residents of Ontario with gastrointestinal symptoms as chief complaints to the 24-hour, 7-days-a-week, confidential telephone hotline service THO. The service is operated by Sykes Assistance Services Corporation, who are contracted by the Ontario Ministry of Health and Long-Term Care. Any resident from the province may call the hotline for any reason, where a responding nurse provides basic medical advice and directs the caller to an appropriate next course of action (i.e. visit an emergency room immediately, see a family physician very soon or within the next day, or self-care). The service helps alleviate the pressure on emergency departments and doctor's offices, while simultaneously providing free medical advice to millions of Ontarians. The THO dataset provides information

on the date and time of call, the caller's chief complaint (primary reason for calling), the nurse's suggested next steps, age, gender, and location (city) of call.

The calls made to THO are referred to as AGE in this study rather than VGE as in the iPHIS and PHOL datasets due to the fact they are less specific and encompass all causes of AGE, not strictly viral aetiologies. In addition, the calls made to THO for norovirus-like illness has not been confirmed as norovirus. Therefore, the “relevant chief complaints” are used as a proxy for norovirus and norovirus-like illness and the burden of AGE calls cannot be attributed strictly to norovirus, but rather norovirus and norovirus-like illness.

REFERENCES

1. Ontario Ministry of Health and Long-Term Care. Ontario Public Health Standards: Appendix B: Provincial Case Definitions for Reportable Diseases. August 2015.