Integrating Approaches to Geographic Variation in Methodologies for Public Health Surveillance

by

Douglas Christopher Dover

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

School of Public Health
University of Alberta

ABSTRACT

Epidemiology is increasingly recognizing the complexity of the underlying mechanisms determining health states. Public health surveillance needs to incorporate this knowledge into their regular reporting and analysis cycles. Aggregate data related to a multitude of health related states and risk factors is produced and publicly shared by public health surveillance. These large stores of aggregate data have the potential to be combined and analyzed to capture much of the underlying complexity. The aim of this thesis is to advance the methods used in public health surveillance for combining and analyzing these disparate sources of aggregate data. Three papers address this aim by focusing on (1) developing a sound methodology using funnel plots for the analysis of aggregate health data, especially addressing the issues of policy relevant analysis and overdispersion, (2) developing a spatial scan statistic capable of identify multiple irregularly shaped clusters in aggregate space-time data, and (3) applying the funnel plot and spatial scan techniques to childhood immunization surveillance in Alberta. These papers conclude that (1) the funnel plot methodology is a robust way of creating policy relevant analysis with understandable visualizations in the presence of overdispersion, (2) the novel MultScan spatial scan performs well at cluster detection, and (3) sophisticated surveillance of childhood immunization can be undertaken accounting for a wide variety of determinants using available aggregate data.

PREFACE

Chapter 2 of this thesis has been published as D. Dover, D. Schopflocher "Using funnel plots in public health surveillance", *Population Health Metrics,* 9(1) p.58 doi:10.1186/1478-7954-9-58. I was responsible for the concept formation, design of the study, analysis of the data, interpretation of the results, and manuscript composition. D. Schopflocher was the supervisory author and was involved in the concept formation, interpretation of the results, and contributed to manuscript edits.

Chapter 3 of this thesis entitled "Multiple, Irregular, Spatial-Temporal Cluster Identification in Public Health Surveillance" is an original work by Douglas C. Dover. It received research ethics approval from the University of Alberta Health Research Ethics Board, under Project Name "Integrating approaches to Geographic Variation in Methodologies for Public Health Surveillance", Study ID Pro00058825, March 14, 2016.

Chapter 4 of this thesis is intended to be published as D. Dover, S. MacDonald "Use of funnel plots and spatial scans to guide immunization surveillance and intervention". I was responsible for the concept formation, design of the study, analysis of the data, interpretation of the results, and manuscript composition. S. MacDonald was the supervisory author and contributed to the concept formation, interpretation of the results and manuscript edits. It received research ethics approval from the University of Alberta Health Research Ethics Board, under Project Name "Integrating approaches to Geographic Variation in Methodologies for Public Health Surveillance", Study ID Pro00058825, March 14, 2016.

ACKNOWLEDGEMENTS

I would like to acknowledge the many individuals who have supported me through this journey. My committee members Dr. Don Voaklander, Dr. Candace Nykiforuk, Dr. Shannon MacDonald, and Dr. Ambikaipakan (Sentil) Senthilselvan who have supported me throughout. Their kind words of encouragement meant a lot to me. My employers Shaun Malo and Larry Svenson at Alberta Health have been unwavering in their support and patience. Many friends and colleagues have been encouraging, supportive, helpful, and understanding through the many ups and downs. All of their support (and there are too many of you for me to remember all at once) helped so much to sustain me. I really appreciate it. Dr. Don Schopflocher has been a colleague and mentor for many years now. Without him, this thesis would not have been written. Thanks so much for the many lively discussions and the devoted guidance Don.

I would like to thank my family for their love and support through my PhD. It has not been a short trip and has impacted those around me. Christopher and Patrick, thank you so much for being so understanding, patient, and supportive. Allison, thank you so much for always being there, for the empathy, and for the friendly competition. All of the love, support and encouragement means the world to me.

There are also enumerable other colleagues, friends, and family members who have been supportive throughout this journey. The kind words and supportive comments were heard – and please know that they helped.

To everyone, a very sincere Thank You.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATC	Anatomical Therapeutic Chemical
CCHS	Canadian Community Health Survey
CDC	Center for Disease Control
DTaP-IPV-Hib	Diphtheria, tetanus, acellular pertussis, polio, haemophilus influenzae type b
ICD-10	International Classification of Disease, 10th edition
IHDA	Interactive Health Data Application
Imm/ARI	Immunization and Adverse Reactions following Immunization
KSS	Kulldorff's spatial scan
LLR	Log-likelihood ratio
LR	Likelihood ratio
LTSS	Linear time spatial scan
Men C	Meningococcal conjugate
MIZ	Metropolitan Influence Zone
MMR	Measles, mumps, and rubella
MMRV	Measles, mumps, rubella, and varicella
MultScan	Multiple cluster, multiple direction, spatial scan
NHS	National Household Survey
NOB	Notice of Birth
OR	Odds ratio
PCV13	Pneumococcal conjugate
PPV	Positive predictive value
RHA	Regional Health Authority
RR	Relative risk
Std Dev	Standard deviation
WHO	World Health Organization

CHAPTER 1: INTRODUCTION

1.1 MOTIVATION

The aim of this thesis is to improve the methods used in public health surveillance, in particular, methods for analyzing aggregate data. My personal experience in public health surveillance over the last twenty years has identified that as an area where the discipline can improve substantially. There are now vast quantities of data available covering a wide range of social determinants of equity, social determinants of health, risk factors, health service use, and health outcomes. Often this data is available only in the form of aggregate data for a set of geographies, such as immunization coverage for local health services planning areas. My observation has been that these vast stores of data are rarely linked together (by geography) so that they can be included in the ongoing monitoring and systematic analysis activities that underpin public health surveillance. It is my hope that the methods presented in this thesis begin to change that.

1.2 OUTLINE

This thesis is comprised of three papers in the theme of analysis of aggregate data in public health surveillance. This chapter proceeds to provide background and literature reviews for each of the three papers. Limitations and current gaps are then discussed. Chapter 2 presents the first paper providing a methodology for the use of funnel plots for effective visualization of surveillance methods, published in Population Health Metrics. Chapter 3 presents the second paper (submitted) developing and evaluating a spatial scan statistic (referred to as MultScan). The third paper in Chapter 4 (to be submitted) provides a synthesis of these methods with an application of them to childhood immunization surveillance in Alberta. Chapter 5 closes the thesis with a summary and integration of the results, discussion of potential limitations, and future directions.

1.3 Background and Aims

Public health is “the science and art of preventing disease, prolonging life and promoting health through the organized efforts and informed choices of society, organizations, public and private, communities and individuals” (Last, 1998). These organized efforts are known as public health interventions. A key component to public health interventions is their reliance upon evidence to inform the choice of intervention and its target population. Evidence comes from a number of sources, including scientific literature and surveillance activities. Public health surveillance, which evolved within the science of epidemiology, is the "ongoing, systematic collection, analysis, and interpretation of health-related data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination to those who need to know"(Center for Disease Control (CDC), 1986).

Public health is a large-scale endeavour. It is rarely if ever possible or feasible to employ a single health intervention for an entire population. Instead, smaller subpopulations are usually targeted in order to make the practice and delivery of public health possible. One consequence is that responsibility for delivery of public health programmes is divided between administrative areas. Evaluating public health programme delivery involves comparing outcomes within and between administrative areas. The most common method for creating administrative units for the delivery of public health is based on geography.

Data collected through public health surveillance processes on health events or the individuals to whom these events occur are generally geocoded (indexed by location) in order to allow the examination of geographic variation in disease and health outcomes. Beginning early in the history of epidemiology (Graunt’s analysis of the Bills of Mortality in the 16th century), counts of deaths due to communicable diseases have been aggregated by geographic area and disseminated to decision makers (Choi and Pak, 2001). This type of data was put to use first in discovering disease causes and subsequently in designing and evaluating public health interventions (Choi, 2012). Surveillance systems today inform administrators in the planning, implementation and delivery of public health programmes in their areas.

Individual administrative areas, however well-constructed, may not capture phenomena that encompass multiple areas. Recently, analyses of geocoded health data have also included methods for the detection of clusters of disease that span multiple administrative areas in close geographic proximity. This type of analysis allows for a coordinated public health response and avoiding duplication of service delivery.

Immunization is one area of public health practice that could benefit from the application of systematic methods analyzing geographic variation. The wide body of knowledge relating to factors influencing childhood immunization is known to regional public health practitioners and administrators. Neither this knowledge nor the systematic analysis of geographic variability and clustering are systematically and routinely incorporated in childhood immunization surveillance systems.

The goal of this research is to further develop methods for analyzing and interpreting geographically aggregated data in public health surveillance. I focus on three specific aims:

- Evolve methods to separate pre-defined administrative areas that have high rates of disease (or other health outcomes) from those which have low rates, especially when areas differ in population and characteristics related to the causes of disease. This aim is addressed with the funnel plot methodology for visualizing and accounting for multivariate risk factors in Chapter 2.
- Determine whether these methods can be extended to determine whether the neighbours of an area with an unusual rate are also unusual. This aim is addressed in Chapter 3 with the development of a spatial scan method.
- Demonstrate the utility of these methods for a specific public health issue, in particular, childhood immunization. This aim is addressed in Chapter 4, assessing childhood immunization after accounting for a suite of known determinants of immunization.

Next, the background and literature around each aim is reviewed.

1.4 Literature Review

Visual representation of public health information, especially the inclusion and understanding of the effects of known risk factors, is crucial to informed decision making for public health interventions. This portion of the literature review will provide an overview of key visualization methods used for geographically aggregated data and their associated background literature.

1.5 Funnel Plots

The visual display of this data is intended to, in a methodologically sound manner, rapidly convey the information required for public health decisions. In this section, I begin by providing a context around the display of aggregate geographic data. The section concludes with a review of the literature on the use of funnel plots in public health surveillance.

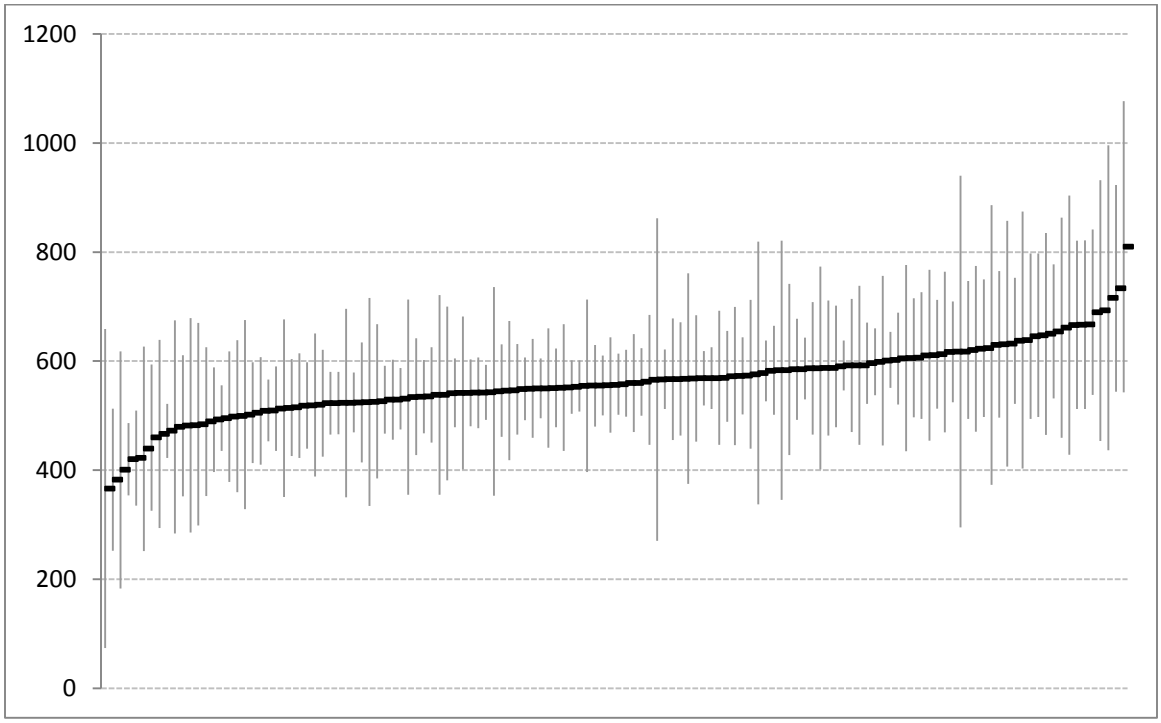
1.5.1 Displaying Small Area Data

The results of analyses conducted on data collected within a surveillance system can be used to inform public health policy and planning, to monitor the health status of a population, and to stimulate research. A functional surveillance system will provide information about the number of health events of specified types that occur within specified populations on an ongoing basis and can therefore be used to derive disease and health event rates over time in different areas or subpopulations.

Routine surveillance activities include monitoring rates of disease occurrence and other health events in small geographic areas in order to identify anomalies that might have a geographic basis and to enable the reporting of such anomalies to administrators in these areas. Substantial variability in population sizes in small areas introduces methodological challenges in the comparisons of rates as the precision of estimates of these rates depends on the size of the population over which they are measured.

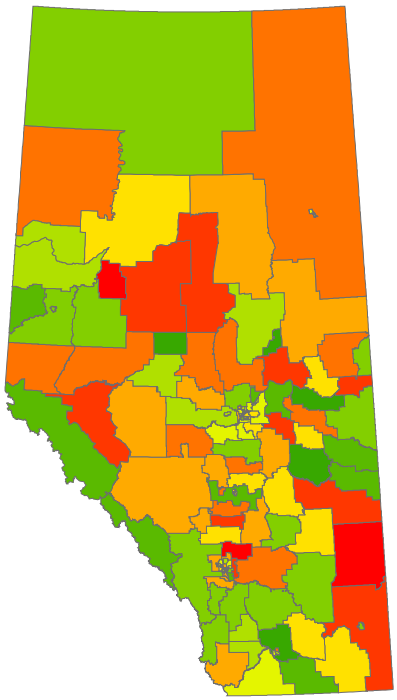
Several graphical procedures have been proposed for displaying small area rates to support the location of anomalous patterns. League plots (Woodall, 2006) and choropleth maps (Rogerson and Yamada, 2008) are two common approaches shown in Figure 1-1 and Figure 1-2, respectively. League plots are an a-spatial display of observed rates (with confidence intervals) ordered by those rates. These plots are difficult to interpret (Marshall, Mohamnmed and Rouse, 2004) because they encourage interpretation as a rank ordering, and these rank orderings are known to have extremely poor statistical properties [see for example, (Marshall and Spiegelhalter, 1998; Shen and Louis, 2000)]. Choropleth maps of rates apply differential colour schemes to chosen categorizations (often quintiles) of observed rates and colour each area on a map according to the category of its observed rate. These are also easy to misinterpret because the map reflects geographic area rather than population density and because the same data may result in maps with very different appearances, since the choice of category is arbitrary (Mazzucco *et al.*, 2017). Cartogram versions (Sui and Holt, 2008) of choropleth maps attempt to redraw areas in proportion to populations but are often difficult to reconcile to geographies and still suffer from the arbitrary category problem.

Figure 1-1. League Plot



The league plot orders geographies by rate and displays 95% confidence intervals around each rate. League plots naturally draw the viewer's attention to very high or very low rate values – regardless of the variability (confidence interval) of those rates.

Figure 1-2. Choropleth Map



The choropleth map assigns each geography a colour based on the value of the rate. It naturally draws the eye to high and low rate values – regardless of the associated variability.

1.5.2 Funnel Plots in Public Health Surveillance – Literature Review

Funnel plots are an alternative to both league plots and choropleth maps. Funnel plots are a form of scatter plot in which observed area rates are plotted against area populations. Control limits are then overlaid on the scatter plot [see Figure 1-3(a)]. The control limits represent the expected variation in rates assuming that the only source of variation is stochastic. The control limits are computed in a fashion very similar to confidence limits and exhibit the distinctive funnel shape as a result of the smaller expected variability in larger populations.

Funnel plots were first introduced in meta-analyses where they were originally used to determine whether a lack of a particular type of published finding demonstrated the presence of a publication bias (Light and Pillemer, 1984; Sterne, Egger and Smith, 2001).

This would be indicated by the absence of points in a particular region of the funnel (especially an absence of studies with a small sample size and a negative result).

The funnel plot can also be considered a form of control chart (Woodall, 2006). Control charts monitor whether a manufacturing or business process is in control. If analysis indicates that the process is currently stable, with only stochastic variation, then data from the process will vary within known limits and can be used to predict the future performance of the process. If the chart indicates that the data from the process being monitored are too variable, analysis of the chart can help determine the sources of variation, which might then be eliminated to bring the process back into control. In a funnel plot, if rate variation is only stochastic, then an appropriate proportion of the points representing area rates will tend to fall within the funnel, and importing control chart terminology, the (rate generation) process is considered to be "in control" [see Figure 1-3(a)]. Reverting to statistical terminology, the model fit is adequate (where the model is of a single stable rate). When many rates fall outside the funnel [see Figure 1-3(b)], the plot can be described as "overdispersed" and it can be said that the process is not in control or the model does not fit the data well. Points lying outside of the funnel are triggers to further investigation.

Figure 1-3(a). Funnel Plot – In Control

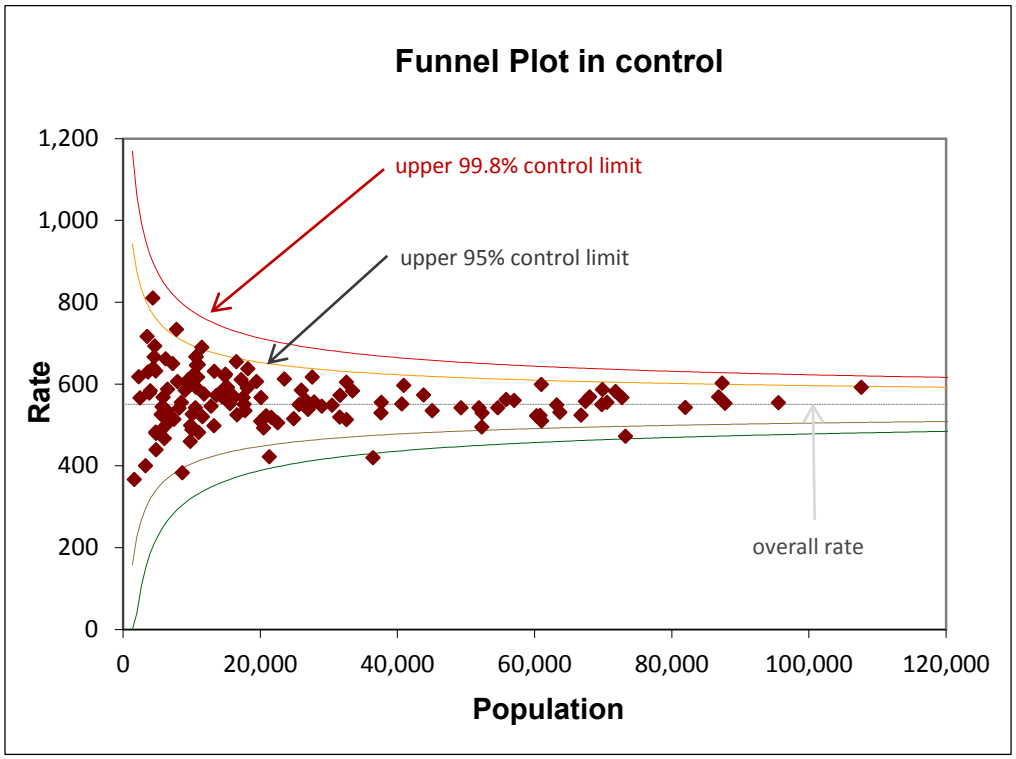
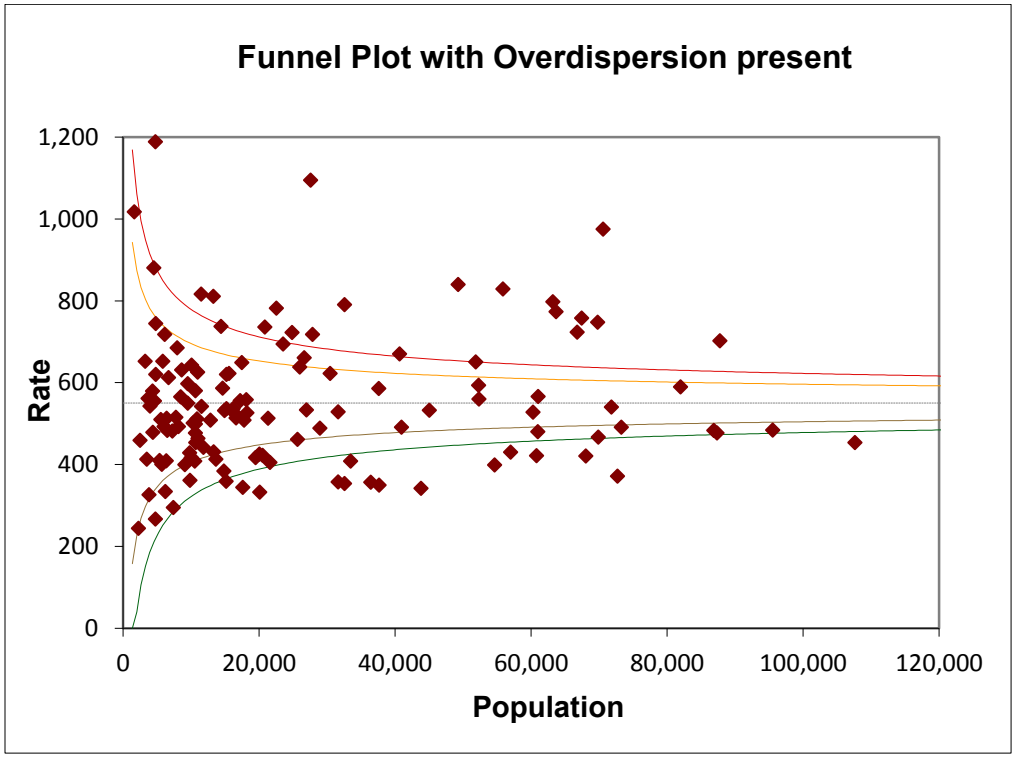


Figure 1-3(b). Funnel Plot – Overdispersion



Funnel plots have been adapted to health system performance in various jurisdictions where it is assumed that administrators within the health system or institution can exercise control over a health event-related process (Spiegelhalter 2005a). Funnel plots have been used to monitor a wide variety of institutional outcomes, such as trauma mortality by hospital (Kirkham and Bouamra, 2008), lung cancer mortality by Primary Care Trust (Nur *et al.*, 2015), colorectal postoperative mortality by National Health Service (NHS) Trust (Byrne *et al.*, 2013), and individual cardiologists' surgical mortality rates (Kunadian *et al.*, 2009). Many of the issues in institutional performance monitoring are shared by health surveillance in support of public health. Both activities deal with small domains, highly variable rates, large differences in population sizes, multiple testing issues, ongoing monitoring activities, and dissemination of results to interested parties invested with the authority or responsibility to affect change. Funnel plots use in epidemiology or population-based surveillance have thus far been limited, with examples seen for cancer incidence in municipalities (Mazzucco *et al.*, 2017) and breast and cervical cancer screening (Massat *et al.*, 2015).

1.6 Spatial Scans

This section begins with a brief history and review of the uses of spatial scan statistics in public health and epidemiology. The developments in the spatial scanning of predefined areas are now reviewed.

1.6.1 The Use of Geography in Public Health Surveillance

Public health surveillance has a number of goals. These high-level goals [adapted from (Declich and Carter, 1994)]) are reviewed with a focus on how geographic data play an important role in meeting them.

The first goal of public health surveillance is the describing of patterns of health states or of health service utilization. To do accomplish this, analyses carried out in surveillance are based upon the principles of descriptive epidemiology. In particular, many public

health surveillance analyses describe health outcomes broken down by person, place, and time. Place is normally restricted to geographic location in the practice of public health surveillance (Yiannakoulias, 2011), not necessarily accounting for the richness of what a place is and what it represents (Tunstall, Shaw and Dorling, 2004). This choice is an intentional trade off - while some richness of information is lost, a cost-effective systematic disclosure of data is maintained over time. Indeed, the careful use of administrative health data aggregated to geographic boundaries was one of many contributions made during the inception of public health surveillance by William Farr (Langmuir, 1976): registration districts were used in the analysis mortality and geographic variation identified key public health actions to be undertaken. This model of examining variation in health outcomes by geographic areas continues in public health practice today.

The second goal of public health surveillance is to detect anomalies. Similar to the descriptive analysis, anomalies are usually monitored for along person, place and time dimensions, with place being operationalized as a geographic boundary. While many methods are available for anomaly detection, the most common contemporary approach for the monitoring of spatial anomalies is the spatial scan statistic (Yiannakoulias, 2011). This approach has been successfully (if rarely) used to point public health officials in the direction of identifying a cause of elevated disease rates. Aspects of the geographic location are then identified as contributing to the observed disease rates. A prototypical example of this was the identification of exposure to the mineral erionite causing mesothelioma in Turkey (Neutra, 1990).

The third goal of public health surveillance is to suggest hypotheses for further research. Many epidemiologic hypotheses have been proposed by the examination of public health surveillance data at the geographic level. Contemporary examples of risk factors for a wide variety of diseases ranging from asthma to cervical cancer to liver cancer have been suggested by geographic analysis of health surveillance data (Pearce, 2000). Even as the discipline of public health surveillance was being established, geographically defined surveillance data was a prominent component of the analyses carried out. During the mid-1800's, William Farr's reporting on cholera included geographies defined by elevation

(the miasma theory) and by water supplier (John Snow's water theory) (Langmuir, 1976). Geographically defined data in public health surveillance have provided, and continue to provide, substantial contributions to the scientific understanding of population health.

These many and varied uses of geography in public health surveillance point to the importance of developing robust methods for analyzing and visualizing geographically aggregated public health surveillance data.

1.6.2 Spatial Scans in the Public Health Surveillance Context

Public health surveillance and epidemiology both benefit from the identification of disease and health state geographic clusters. Public health surveillance uses this information to intervene and epidemiology uses this information to suggest possible causal hypotheses. The use of spatial analysis in public health surveillance dates back to John Graunt's neighbourhood analysis of the London Bills of Mortality in the 1600s (Choi and Pak, 2001). Analyses did not change substantially until William Farr in the 1800s began analyzing influenza surveillance data to identify geographic locations with an excess of observed to expected cases (Langmuir, 1976). Similar surveillance and dissemination activities leading to control measures continue to this day (Walter, 2000). For example, weekly influenza surveillance activities at the Alberta Ministry of Health compare influenza across different geographic health service zones in the province and compare the current to expected counts [see for example (Alberta Health, 2017)]. This information is then disseminated to Medical Officers of Health and other public health officials across the province.

While many of the fundamental principles of surveillance have not changed, the tools and methods have evolved over time. Cluster detection is now an integral component of public health surveillance activities. Spatial scan statistics are a commonly used method to detect clusters (Rogerson and Yamada, 2008). A spatial scan statistic is a test statistic that (a) searches the data (in this case, geographies) to identify a cluster, and (b) assesses the statistical significance of the identified cluster. If the geographic area with the potential cluster were known *a priori*, a statistical comparison would be straightforward

between that area and another. Spatial scan methods are more complex because the geography defining the cluster is unknown. Scan procedures would ideally search all possible geographies finding the area with the largest objective function (commonly the likelihood ratio) and hence most likely to be the cluster, and then evaluate the statistical significance of that most likely cluster.

While spatial scans can search point data or aggregated data, this research focuses only on spatial scans that search data aggregated over predefined geographic regions as these are the most applicable in a public health surveillance context. Before detailing spatial scan statics themselves, it is important to first consider the implications of using a "pre-defined geography". The defining of boundaries has profound effects on what can be analyzed, what the results will be, and the value of said results. Form an analysis perspective, the scale of the regions and the choice of groupings of areas at different levels affects the findings. This is known as the modifiable areal unit problem (Waller and Gotway, 2004). It is possible to find radically different results — high positive correlations, no correlation, large negative correlations — through different aggregations of smaller areas. Additionally, the scale itself affects results — as there are fewer (larger) areas at a higher level, effect measures are attenuated. These results point to how fundamental the choice of geographies is.

Public health surveillance is fundamentally an applied discipline. The choice of geographies will need to fit to the requirements of a surveillance system. In particular, surveillance systems need to maintain comparability, be cost effective, and be sustainable over time. These needs push towards defining geographic regions once (or only a few times). This is in tension with the need to tailor geographic boundaries to the scale of the health outcome of interest to obtain analytically meaningful results and to the need to continuously create complex definitions of place relevant to the outcome of interest (Yiannakoulias, 2011). Public health surveillance takes a compromise path. Beginning in the 1800's, William Farr utilized the pre-existing 2,000 registration districts in the analysis of the Bills of Mortality and also promoted their use in the census to collect high quality denominators (population values) (Langmuir, 1976). This represents using an organic, community created definition of an area. This has the advantages of capturing

some aspects of place while maintaining its applied value. Considerations are similar today, but the processes can differ. As an example, the local geographic areas used in Chapter 4 and sub-regionals used in Chapter 2 were created by the Ministry of Health and the Health Regions. These geographies were created trying to balance their utility for health service planning, their utility for surveillance and reporting, and their having somewhat homogeneous populations. While not community created, the concept of place continues to be recognized. Boundaries such as these provide the predefined geographies that form the basis of spatial scan statistics.

For predefined geographic units, spatial scans take advantage of distance or adjacency information when searching for a cluster. In Kulldorff's original spatial scan (Kulldorff and Nagarwalla, 1995), every possible circle, starting at every area centroid, was considered. Once the most likely cluster is identified, the hypothesis that the area has a rate higher than the remainder of the areas is tested. The likelihood ratio statistic for the most likely cluster is computed, under the assumption of a known geographic choice. Then, to account for the multiple testing involved in searching for the most likely cluster, Monte Carlo simulation determines the distribution of this test statistic. If the test statistic is significant, the identified cluster is considered real.

By using only distance between centroids of areas, Kulldorff's spatial scan does not in fact guarantee the clusters identified contain adjacent areas. If the areas are extremely irregular in their shapes, it is possible that the next closest area, measured by distance to the centroid, is not actually adjacent to the first. Other spatial scan techniques do take advantage of the graph structure to ensure that clusters are in fact connected. The graph structure relates to topologic structure of the geographic areas. Each area is considered a vertex, and adjacent areas are connected by an edge. In this way the graph theoretic representation of the geographies abstracts away a level of detail, the exact size and shape of each area, to represent the map using only the vertices and edges. It also clearly defines the problem for cluster detection: to search all possible connected areas (sets of vertices joined by edges). Assunção *et al.*, 2006, explicitly use graph theoretic results to optimize the search over connected areas. Weights are associated with each edge using the difference between area rates in the adjacent vertices. A minimum spanning tree

(Gower and Ross, 1969) is then constructed using these weights. A minimum spanning tree removes edges from the graph such that there is still a way to reach every connected area, but each connection is not maintained. If the weights are unique, there is a unique spanning tree created that minimizes the weights (costs of moving between areas). This spanning tree is substantially smaller (measured by the number of edges) than the full graph, making it computationally feasible to search.

Kulldorff's spatial scan transformed the field of cluster identification because this approach was the first to meet the three criteria (Waller and Gotway, 2004) of (a) identifying the existence of a cluster, (b) identifying the spatial location of the cluster, and (c) soundly evaluating the significance of the cluster.

1.6.3 Spatial Scan Statistics – Literature Review

Kulldorff's original spatial scan (Kulldorff and Nagarwalla, 1995) represents the breakthrough in the field of spatial scan statistics in part because the method could identify and evaluate *circular* clusters in a methodologically rigorous way. The use of only circular clusters in conjunction with the proof that the likelihood ratio was the individually most powerful test resulted in a method that was computationally feasible and theoretically sound. The restriction to compact circular clusters is arbitrary and limiting in a field as diverse as public health, however. Clusters of health events can be conceived to occur in any configuration – for example, long and irregular following a river or other geographic feature; starting small and growing outward following an air borne dispersal pattern; compact following a point source exposure; completely irregular following transmission along a social network. This implies that spatial scan techniques applied in public health should be able to detect both irregular and regularly shaped clusters. The FlexScan method proposed by Tango and Takahashi (Tango and Takahashi, 2005) attempts to do this by performing an exhaustive search to detect any possible shape of cluster. But exhaustive searching across a large number of geographies quickly becomes computationally infeasible (Neill and Moore, 2004; Yiannakoulias, Rosychuk and Hodgson, 2007). Thus developments in spatial scans focus on algorithms or approaches that increase the computational speed. Increases in speed may trade off with

loss of accuracy in the detection of clusters, however. That is, as searches become more restrictive, the methods usually become less accurate at cluster detection. Attempts to improve spatial scan methods can be categorized into methods that restrict the shape of the cluster, restrict the search by penalizing the likelihood ratio, or increase the efficiency of the algorithm.

Shape restrictions have been straightforward to develop. Kulldorff's original circular spatial scan (Kulldorff and Nagarwalla, 1995) was later extended to an elliptical scan (Kulldorff *et al.*, 2006). Rectangular scans have also been considered (Neill and Moore, 2004). The original circular scan performs quite well when the underlying cluster is compact or close to compact, but performance (both ability to detect and accuracy) drops off dramatically when the underlying cluster is no longer compact [for example (Goujon-Bellec *et al.*, 2011)]. Fundamentally, when detecting a cluster known *a priori* to be compact, straightforward methods exist today that take advantage of that prior knowledge.

A wide variety of penalization approaches have been proposed. Non-compactness penalties on the likelihood ratio that limit the growth of or absolute size of the cluster have been proposed (Duczmal, Kulldorff and Huang, 2006; Yiannakoulias, Rosychuk and Hodgson, 2007). Depth limiting search rules, restricting how many additional geographies would be included in a search, have also been proposed (Yiannakoulias, Rosychuk and Hodgson, 2007; Tango, 2008). Search rules can consider the significance of rates in individual geographies prior to inclusion in a potential cluster (Tango and Takahashi, 2005) or the search can be restricted to a local (rather than global) maximum of the likelihood ratio with a length based stopping rule (Yiannakoulias, Rosychuk and Hodgson, 2007). All of these methods rely upon arbitrary tuning parameters.

There have been a number of different algorithms proposed to detect irregularly shaped clusters. The upper level set approach (Patil and Taillie, 2004) sorts the data by the observed rates and proceeds within each level set (the set of rates higher than a value) to evaluate clusters of connected geographies. This can be seen in Figure 1-4 following down the dashed lines from the largest to ever smaller observed rates. The method was generalized as a minimum spanning tree (Assunção *et al.*, 2006) and is related to the

weak version of the linear time subset scanning algorithm (Neill, 2012) that proved the optimality of this approach under certain restrictive conditions. An order based search sorts the data by the observed z-scores and proceeds to create clusters which are then evaluated (Que and Tsui, 2011). A simulated annealing approach (Duczmal and Assuncao, 2004) that begins with the most likely cluster from Kulldorff's circular scan has also been proposed. A genetic algorithm with similar performance but improved speed compared to the simulated annealing approach has also been developed (Duczmal *et al.*, 2007). More recently, Neill (Neill, 2012) has provided a theoretical result proving that the likelihood can be maximized in a computationally efficient manner. This approach has been implemented as GraphScan (Speakman, McFowland and Neill, 2015).

The methods described thus far focus on detection of a single cluster. Theoretical approaches to multiple cluster detection have been pursued in the literature in a limited number of forms. A sequential test to evaluate cluster significance for up to three distinct clusters has been developed (Li *et al.*, 2011). This method approaches the problem of multiple clusters as an issue of p-value evaluation, and assumes the underlying scan is appropriate and can identify non-overlapping clusters. The method is restricted to a small, fixed maximum number of clusters (three in the original paper). A theoretically elegant approach to multiple cluster detection is to place a prior distribution on the expected number of clusters (Wakefield and Kim, 2013). This method then simultaneously estimates the number of clusters and their locations and rates. Unfortunately, this method suffers from sensitivity to the choice of prior distribution and other tuning parameters. Due to their computational complexity and non-standard software requirements, neither of these approaches are commonly used in public health surveillance. Current practice involves a variation of detecting clusters (Zhang, Assuncao and Kulldorff, 2010) and then evaluating each as though it were the only cluster detected. Some approaches list all "significant" clusters that are non-overlapping, evaluated in a ordered manner based on the likelihood ratio (Kulldorff, 1997); other approaches detect the cluster, remove it from the data, and repeat this process until no further significant clusters are detected (Zhang, Assuncao and Kulldorff, 2010). These approaches are generally believed to have conservative properties, but the performance characteristics are unknown [for example (Wakefield and Kim, 2013)] and lack a sound theoretical foundation.

Figure 1-5 summarizes the substantive contributions following Kulldorff's introduction of the circular spatial scan statistic. Details of the statistical basis for spatial scans are reserved until Chapter 3 when the discussion can involve the new MultScan spatial scan developed.

Figure 1-4. Funnel plot and level sets

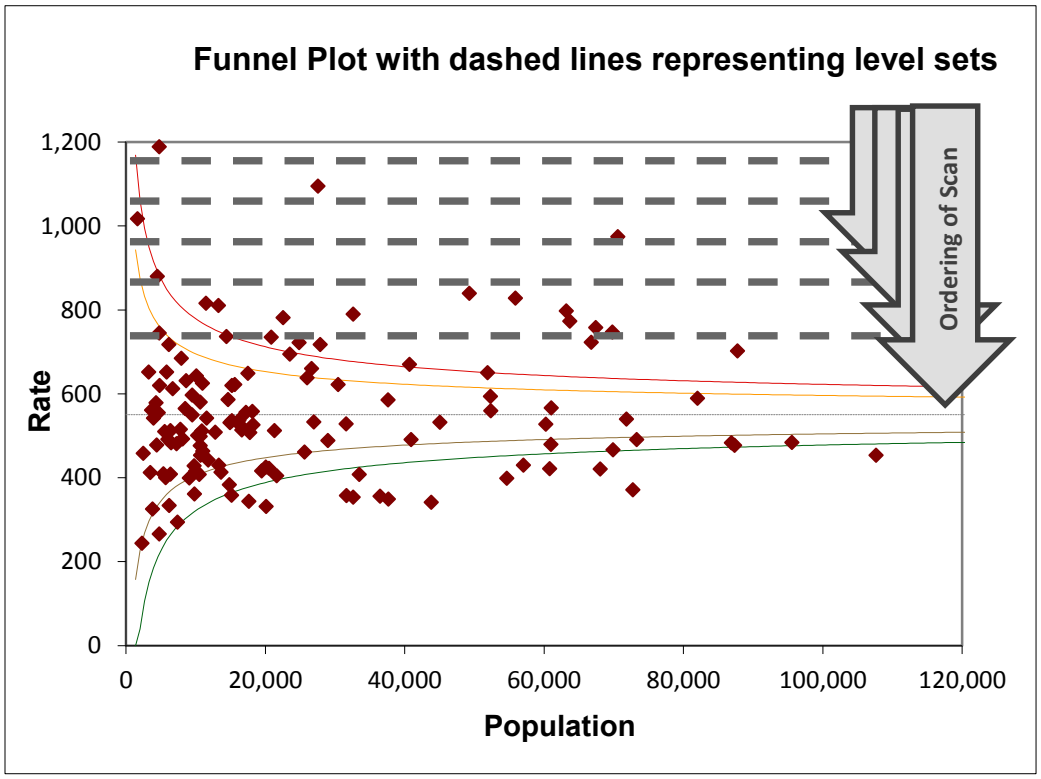
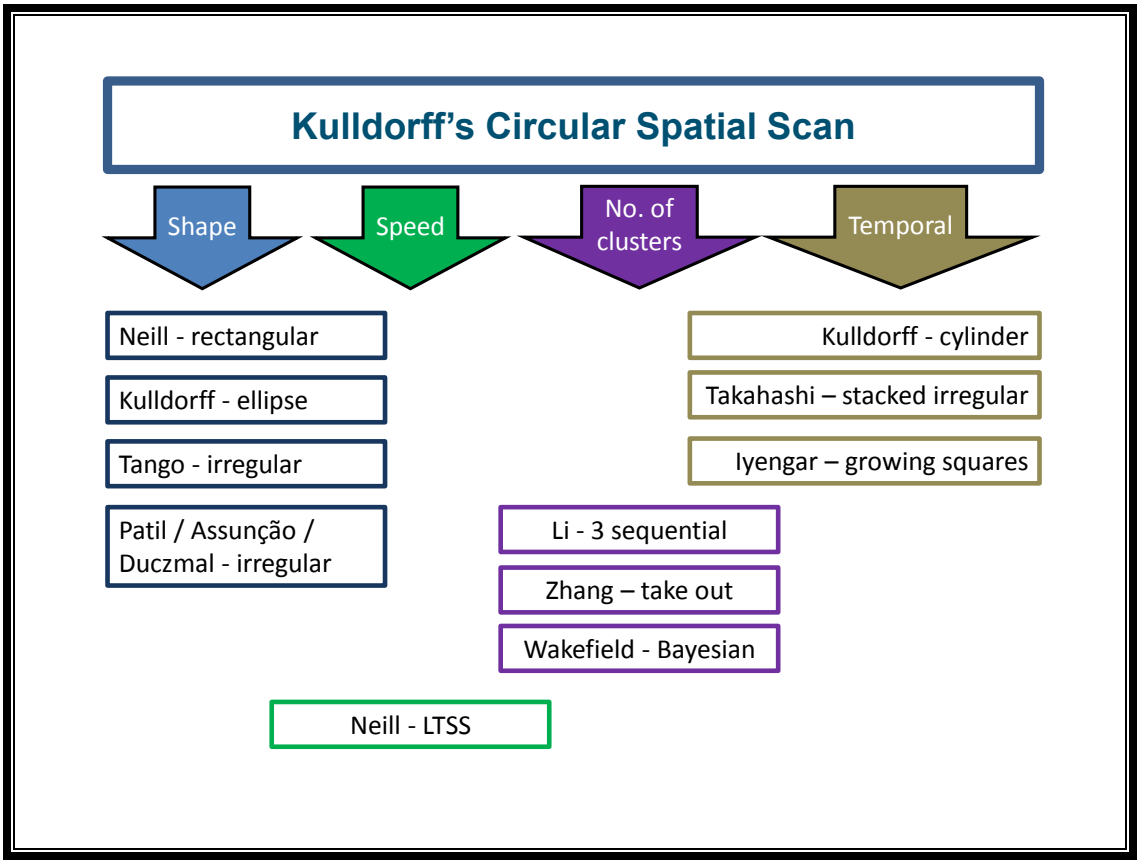


Figure 1-5. Evolution of Spatial-Temporal Scan Statistics



1.7 Childhood Immunization Surveillance

This portion of the literature review will provide an overview of childhood immunization in Alberta. It will provide background regarding the delivery of this public health intervention in Alberta and summarize key literature regarding factors affecting childhood immunization uptake.

1.7.1 Background - Childhood Immunization in Alberta

The routine childhood immunization schedule for Alberta, up to the age of two, is shown in Table 1-1. The routine schedule changes over time with the introduction of new vaccines or improvements in the timing of the schedule. The schedule in Table 1-1 is as of February 2011, to coincide with the study period in Chapter 3. The temporal spacing between administrations of the same vaccine and the age at which vaccines are administered represent an evidence-based attempt to optimize the child's immune response and minimize adverse reactions within the immunization delivery system. For example, MMRV, protecting against measles, mumps, rubella, and varicella (MMRV), is given at 12 months of age. MMRV vaccination given prior to 12 months of age does not result in a sufficient antibody response (Goh *et al.*, 2007). These immunological considerations in conjunction with the practical considerations surrounding vaccine delivery in a particular context result in the routine immunization schedule (Institute of Medicine, 2013)

Table 1-1. Alberta's Routine Childhood Immunization Schedule to Two Years of Age

2 months	• DTaP-IPV-Hib[1] • Pneumococcal conjugate (PCV13) • Meningococcal conjugate (Men C)
4 months	• DTaP-IPV-Hib • Pneumococcal conjugate (PCV13) • Meningococcal conjugate (Men C)
6 months	• DTaP-IPV-Hib • Pneumococcal conjugate (PCV13) (for high risk children only)
6 months and older	• Influenza[2]
12 months	• MMRV[3] • Meningococcal conjugate (Men C) • Pneumococcal conjugate (PCV13)
18 months	• DTaP-IPV-Hib

Effective: February 3, 2011

Note: Each bullet represents one vaccine/injection unless otherwise noted.

- [1] Diphtheria, tetanus, acellular pertussis, polio, haemophilus influenzae type b
- [2] Annually, during influenza season
- [3] Measles, mumps, rubella, and varicella

Source: http://www.health.alberta.ca/health-info/imm-routine-schedule.html Accessed February 3, 2012

1.7.2 Childhood Immunization Coverage Surveillance

Immunization has been extremely successful in reducing the burden of communicable disease morbidity and mortality and is regarded as one of the top ten achievements in public health in the 20th century (Center for Disease Control (CDC), 2011). The ability to have accurate and timely data on immunization coverage in a population is important for two reasons. First, accurate information on coverage allows the identification of areas that have low coverage and hence should be targeted for interventions. Second, knowledge of immunization coverage can highlight areas at risk of a disease outbreak thus permitting informed planning of the public health response. This highlights the importance of immunization surveillance in protecting the population's health. To illustrate this importance, I briefly review the potential impact that immunization surveillance can have in public health using the recent 2013 measles outbreak in Alberta as an example (Suttorp, 2014). Surveillance data and local knowledge were key components to planning for the anticipated measles outbreak. Immunization surveillance data showed that certain communities had persistently low measles immunization coverage. This information, in conjunction with local knowledge and knowledge of coverage in surrounding communities, lead to evidence informed planning by the regional Medical Officer of Health. When in 2013 the outbreak did in fact occur, it was well contained and resulted in a total of only 42 confirmed measles cases (Suttorp, 2014). This can be compared to recent outbreaks in Quebec (De Serres *et al.*, 2013) and the Netherlands (Woudenberg *et al.*, 2017), where there were 725 and 2,700 cases, respectively. Such rapid response and excellent epidemic containment were the product of using surveillance data to identify public health risks and plan the public health response.

Public health officials are responsible for protecting the health of the population - this mandate clearly extends to protecting the population from vaccine preventable diseases. To effectively do this, public health officials must understand immunization coverage and determine what interventions are required in which areas. I now review the state of the art in immunization surveillance practice.

Currently, only the most basic choropleth maps of immunization coverage rates are used to inform public health officials (Walter, 2000). While examples of more complex analyses of immunization coverage exist in the research literature [for example, (Lieu *et al.*, 2015)], a robust methodology meeting the complex, population based, ongoing, systematic surveillance needs of public health has not yet been developed.

To inform the choice of intervention, a surveillance system must assist in explaining immunization uptake in the population across geographic administrative areas. Knowledge of factors influencing immunization behaviour must be included to intelligently inform the type of intervention required. A surveillance system must also be able to identify areas where the population is at unusually high risk and simultaneously most amenable to intervention. The optimal approaches and methods for explaining immunization behaviour, or any health outcome in fact, and identifying spatial clusters will come from the research domain. The objective in designing a surveillance system is to evaluate these optimal approaches, modifying them as necessary for incorporation into a surveillance system.

1.7.3 Literature Review - Determinants of Childhood Immunization Uptake

1.7.3.1 Behavioural Model of Health Services Use

I use the Behavioural Model of Health Services Use (Andersen and Newman, 1973; Aday and Andersen, 1974) to assist in framing the large number of determinants of immunization uptake. This framework was originally developed for the study of families' use of health services. It characterizes relationships between health policy, the health care system, characteristics of the population and health service utilization and satisfaction. Briefly, the model identifies Predisposing Characterises, Enabling Resources, and Health Care System Characteristics that influence Health Care Service Utilization and Satisfaction with services. This thesis focuses on the utilization outcome of immunization uptake, so there will be no further exploration of the satisfaction element. Figure 1-6 shows the hypothesized interrelationships between the factors. The Predisposing Characteristics represent those characteristics associated with an individual choosing to

utilize health care resources or not (Andersen and Newman, 1973). Enabling Resources represent those conditions that support the utilization of health care resources when an individual is predisposed to do so (Andersen and Newman, 1973). Health Care System Characteristics refer to what resources the system has and how it distributes and structures those resources for the delivery of health services (Aday and Andersen, 1974). This framing of influencing characteristics as Predisposing Characteristics, Enabling Resources or Health Care System Characteristics has been useful in many areas [see for example (Phillips *et al.*, 1998) for a review], including childhood immunization (Acosta-Ramírez *et al.*, 2005). Each component is now reviewed, using the model to organize the known determinants to childhood immunization (refer to Table 1-2 for a summary).

1.7.3.2 Predisposing Characteristics

Predisposing characteristics are those individual or family characteristics that influence the decision to immunize. Beliefs – religious, moral, and philosophical – play an important part in choosing to attempt to access immunization services (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016). These beliefs may include the belief that immunization is either not allowed or is harmful. Beliefs have been shown to have quite large effects in multivariate models (van der Wal *et al.*, 2005).

The perceptions around risks and benefits are some of the most influential determinants of immunization uptake. Perceived vaccine safety (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; MacDonald *et al.*, 2014; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016), vaccine effectiveness (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; MacDonald *et al.*, 2014; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016), and social responsibility (Mills *et al.*, 2005; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016) must be weighed against, for instance, perceived susceptibility to disease (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; MacDonald *et al.*, 2014; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016), disease severity (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; MacDonald *et al.*, 2014; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016),

perceptions of fear around needles and the number of needles the child is receiving, and perceptions of pain during vaccination (Mills *et al.*, 2005; MacDonald *et al.*, 2014; Larson *et al.*, 2014). Illustrating this, "too many needles at once" has been documented as having an adjusted odds ratio (OR) of 7.7 (MacDonald *et al.*, 2014) in an Alberta multivariate study examining incomplete immunization; this odds ratio was almost twice any other odds ratio in the study. Risk-benefits have also been seen as the most influential construct in the Americas (Larson *et al.*, 2014). Perceived contraindications to immunization (such as recent illness) (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Pearce *et al.*, 2008) and past immunization experiences (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; MacDonald *et al.*, 2014; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016) are also predisposing factors.

In households with a single child, immunization is but one of many child health related activities that must be attended to. There are additional demands in running a house and family that also require attention. In all of these competing demands, time and resources must be allocated to immunization. In households with more demands, the likelihood of immunization occurring can decline. These competing pressures are magnified when there are multiple children in the household (Pearce *et al.*, 2008; MacDonald *et al.*, 2014; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016). This is seen in multivariate models capturing this "household chaos" using an indicator of being a second or later child. This has a consistent and sizable effect on immunization (Pearson *et al.*, 1993; Angelillo *et al.*, 1999; Trauth *et al.*, 2002; Dombkowski, Lantz and Freed, 2004; Matsumura *et al.*, 2005; Kim *et al.*, 2007; Pearce *et al.*, 2008). The effects of competing pressures have the potential to be ameliorated through, for example, access to child care (Falagas and Zarkadoulia, 2008; Pearce *et al.*, 2008; MacDonald *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016), hence increasing the likelihood of accessing immunization.

1.7.3.3 Enabling Resources

Socio-economic resources are enabling resources that facilitate access to immunization services. Material resources (Mills *et al.*, 2005; Pearce *et al.*, 2008; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016) are usually measured based on an income

related variable. Results of the effects of material resources on immunization are mixed within Canada [(Lemstra *et al.*, 2007) finding an effect of income while (MacDonald *et al.*, 2014) did not] and elsewhere [(Suarez, Simpson and Smith, 1997; Christakis *et al.*, 2000; Dombkowski, Lantz and Freed, 2004; Zhao, Mokdad and Barker, 2004; Smith and Singleton, 2008) found effects of material resources variables while (Pearson *et al.*, 1993; Trauth *et al.*, 2002; van der Wal *et al.*, 2005; Kim *et al.*, 2007; Pavlopoulou *et al.*, 2013) did not]. Education (Pearce *et al.*, 2008; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016) as a measure of socio-economic resources does not appear empirically as significant in most studies considering several factors simultaneously (multivariate studies) [(Suarez, Simpson and Smith, 1997; Zhao, Mokdad and Barker, 2004; Pearce *et al.*, 2008) found education effects while (Angelillo *et al.*, 1999; Trauth *et al.*, 2002; Dombkowski, Lantz and Freed, 2004; Kim *et al.*, 2007; Lemstra *et al.*, 2007; Smith and Singleton, 2008; MacDonald *et al.*, 2014; Pavlopoulou *et al.*, 2013) did not].

Social supports can provide the opportunity to access immunization services even when socio-economic resources are lacking (Pearce *et al.*, 2008; MacDonald *et al.*, 2014; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016). Support available through a spouse (Pearce *et al.*, 2008; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016), measured as living in a two parent family, is only inconsistently related to immunization [(Pearson *et al.*, 1993; Dombkowski, Lantz and Freed, 2004; Kim *et al.*, 2007; Pearce *et al.*, 2008) not in (Lemstra *et al.*, 2007; Smith and Singleton, 2008; MacDonald *et al.*, 2014)] in multivariate studies. Direct measures of self-reported need for social support have not achieved significance (MacDonald *et al.*, 2014) in relation to immunization outcomes. Supports through ethnic community (Pearce *et al.*, 2008; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016), healthy families (Pearce *et al.*, 2008), and caring about at risk household members (MacDonald *et al.*, 2014) can all facilitate immunization. The broad social context within which immunization is provided and discussed (Dubé *et al.*, 2013; MacDonald *et al.*, 2014; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016) also can support the choice to immunize through, for instance, the influence of social norms on behaviours.

Direct measures of specific immunization knowledge (Mills *et al.*, 2005; Dubé *et al.*, 2013; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016) and beliefs around vaccine safety (Dubé *et al.*, 2013; Larson *et al.*, 2014) have consistently been found an important factor in immunization uptake (Matsumura *et al.*, 2005; MacDonald *et al.*, 2014). When knowledge was restricted to knowing the immunization schedule (Falagas and Zarkadoulia, 2008), this result was no longer important (Angelillo *et al.*, 1999).

Transportation or the location of immunization clinics can present a barrier to immunization uptake (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Pearce *et al.*, 2008; MacDonald *et al.*, 2014; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016). Results of studies controlling for a variety of factors do not support the importance of location and transportation in Alberta (MacDonald *et al.*, 2014).

The "interaction between patients and providers is the cornerstone of maintaining confidence in vaccination" (Dubé *et al.*, 2013) and appears in most reviews of determinants (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; MacDonald *et al.*, 2014; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016). Having a regular medical provider is statistically significant in studies considering several factors simultaneously (Christakis *et al.*, 2000; Dombkowski, Lantz and Freed, 2004; MacDonald *et al.*, 2014).

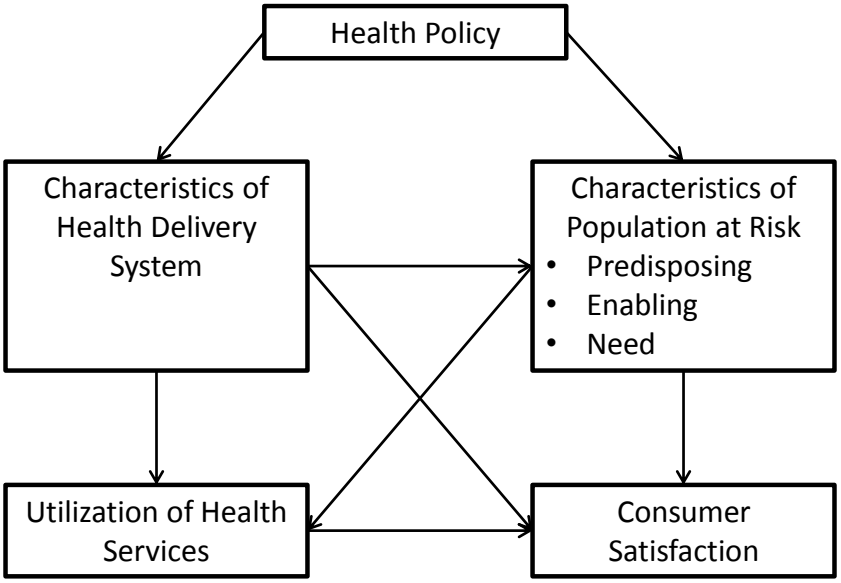
1.7.3.4 Health Care System Characteristics

Access is regularly cited as a factor in health service utilization, including immunization (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Pearce *et al.*, 2008; MacDonald *et al.*, 2014; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016; Thomson, Robinson and Vallée-Tourangeau, 2016). There is however, limited empirical evidence supporting that this affects immunization in Alberta (MacDonald *et al.*, 2014).

Many other health care system characteristics have been related to immunization uptake. Wait times (Falagas and Zarkadoulia, 2008; MacDonald *et al.*, 2014), integration of individuals into the health care system (specially, the immunization stream) (Falagas and Zarkadoulia, 2008; MacDonald *et al.*, 2014), prompts to action provided by the system

(Falagas and Zarkadoulia, 2008; de Cantuária Tauil, Sato and Waldman, 2016; Thomson, Robinson and Vallée-Tourangeau, 2016), provider knowledge (Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013) and attitudes (Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013) towards immunization, the communication skills and interpersonal interactions with the health care system providers (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; Larson *et al.*, 2014), and differences in the organization and implementation of different systems to deliver immunization (Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016) have all been cited. For example, having recently moved, which may represent the ability to integrate into the local health care system after moving from one provider's area to another (Falagas and Zarkadoulia, 2008; MacDonald *et al.*, 2014), shows substantive significant results. Comparisons of health care systems and providers are rare in the literature with mixed results [(Pavlopoulou *et al.*, 2013) finding an association with immunization and (Kim *et al.*, 2007) finding some differences between providers].

Figure 1-6: Behavioural Model of Health Services Use



From Aday and Andersen (Aday and Andersen, 1974).

Table 1-2. Determinants of Childhood Immunization

Health Behaviours Model Category		*Determinant*		Reviews						Multivariate		
				Larson	Falagas	Tauil	Thomson	Mills	Dube	MacDonald	Pearce	
Pre-Disposing Characteristics		Beliefs	Religion based beliefs	✓	✓				✓			
			Prevention/Health related beliefs	✓	✓		✓	✓	✓			
	represent those characteristics associated with an individual making them more or less likely to utilize health care resources (Andersen, Newman, 1973)	Risk - Benefit	Perceived Safety	✓	✓		✓	✓	✓	✓		
			Perceived vaccine effectiveness	✓	✓		✓	✓	✓	✓		
			Perceived disease severity	✓	✓		✓	✓	✓	✓		
			Perceived disease susceptability	✓	✓		✓	✓	✓	✓		
			Percevied social responsibility	✓			✓	✓				
			Fear of needles / pain	✓				✓		✓		
		Perceived Contraindications			✓			✓			✓	
		Past Experiences		✓	✓		✓	✓	✓	✓		
		Competing Pressures	Household chaos	✓		✓				✓	✓	
			Childcare		✓	✓				✓	✓	

Table 1-2. –continued-

Health Behaviours Model Category		*Determinant*		Reviews							Multivariate		
				Larson	Falagas	Tauil	Thomson	Mills	Dube		MacDonald	Pearce	
Enabling Resources		Socio-economic resources	Material resources	✓		✓		✓				✓	
			Education	✓		✓						✓	
	represent those conditions that support the utilization of health care resources when an individual is predisposed to do so (Andersen, Newman, 1973)	Social Supports	Social Support	✓			✓				✓	✓	
			Single Parent	✓		✓						✓	
			Ethnicity	✓		✓						✓	
			Psychological wellbeing of parent									✓	
			Household member at high risk								✓		
`		Social Context	Social Norms	✓			✓		✓				
			Communication/media environment	✓					✓		✓		
		Knowledge	of schedule		✓								
			of vaccines	✓		✓	✓		✓				
			knowledge of safety monitoring	✓					✓				
		Transportation		✓	✓		✓	✓			✓	✓	
		Relationship with Provider		✓	✓	✓		✓	✓		✓		

Table 1-2. –continued-

Health Behaviours Model Category		*Determinant*		Reviews							Multivariate		
				Larson	*Falagas*	*Tauil*	*Thomson*	*Mills*	*Dube*		*MacDonald*	*Pearce*	
Healthcare System Characteristics		**Accessible**	**Convenience of access**	✓	✓	✓	✓	✓			✓	✓	
		Specific Characteristics	**Wait times**		✓						✓		
	refers to what resources the system has and how it distributes and structures those resources for the delivery of health care (Aday, Andersen 1974)		**Integration**		✓						✓		
			Prompts		✓	✓	✓						
			Provider attitudes towards imm.		✓				✓				
			Provider knowledge		✓				✓				
			Provider communication skills	✓	✓			✓	✓				
			System	✓		✓							

The multivariate studies are select studies that capture elements across the spectrum of Behavioural Model of Health Service Utilization categories. MacDonald is an Alberta study.

1.8 Summary of Limitations

1.8.1 Funnel Plots – Limitations to their Use in Public Health Surveillance

Public health surveillance distinguishes itself from institutional performance by the type of process being monitored and hence clear recommendations surrounding the use of funnel plots are required. Health outcomes are the results of complex causal processes where the risk factor prevalence in those processes usually varies across geographies. These processes are an integral part of epidemiologic thinking and are used to inform public health interventions. An ideal model for routine monitoring of a health outcome of interest in health surveillance would begin with a model that fit the data, that is, where the rate generation process could be considered to be understood or in control. Subsequent monitoring over time could focus on whether the rate generation process could be considered to be remaining in control. However, overdispersion is commonly observed in geographically aggregated health data. In a funnel plot, overdispersion manifests as a relatively large proportion of points outside the control limits. Much of the performance measurement literature has focused on statistical adjustments for overdispersion [see for example (Spiegelhalter, 2005b; Ohlssen, Sharples and Spiegelhalter, 2007)]. These statistical adjustments often take the form of random effects models or of variance inflation parameters to widen the control limits. Clear recommendations around the use of funnel plots and the associated approach to dealing with overdispersion need to be developed for the public health surveillance discipline.

1.8.2 Spatial Scan Statistics

Spatial scan statistics today are limited in their ability to deal with multiple, irregularly shaped clusters in space and time. Methods that can balance computational feasibility while attempting to improve upon the positive predictive value of identified clusters are needed. Ideally, these methods could detect clusters that have irregular shapes in both the space and time dimensions. Statistically sound and robust methods to simultaneously identify multiple clusters would improve current practice.

1.8.3 Childhood Immunization Surveillance

To date, no childhood immunization surveillance systems account for the complex determinants of childhood immunization. This complexity must be accounted for in surveillance activities for the information from a surveillance system to be useful to public health professionals.

1.9 Overview and Context for Study Methods

The section provides contextualization and direction for the study methods used in each chapter. The approach to developing the funnel plot methodology, the direction and motivation behind the creation and evaluation of a spatial scan statistic, and integrating these methods in childhood immunization in Alberta are discussed in this section. Further specific details of the methods used are included in their respective chapters.

1.9.1 Funnel Plot Methodology

Control charts have been adapted to health system performance in various jurisdictions where it is assumed that administrators within the health system or institution can exercise control over a health event-related process (Spiegelhalter, 2005a). Many of the issues in institutional performance monitoring are shared by health surveillance in support of public health. Both activities deal with small domains, highly variable rates, large differences in population sizes, multiple testing issues, ongoing monitoring activities, and dissemination of results to interested parties invested with the authority or responsibility to affect change. Public health surveillance distinguishes itself from institutional performance by the type of process being monitored. Health outcomes are the results of complex causal processes where the risk factor prevalence in those processes usually varies across geographies. These processes are an integral part of epidemiologic thinking and are used to inform public health interventions.

An ideal model for routine monitoring of a health outcome of interest in health surveillance would begin with a model that fit the data, that is, where the rate generation process could be considered to be understood or in control. Subsequent monitoring over

time could focus on whether the rate generation process could be considered to be remaining in control. However, overdispersion is commonly observed in geographically aggregated health data. In a funnel plot, overdispersion manifests as a relatively large proportion of points outside the control limits. Much of the performance measurement literature has focused on statistical adjustments for overdispersion [see for example (Spiegelhalter, 2005b; Ohlssen, Sharples and Spiegelhalter, 2007)]. These statistical adjustments often take the form of random effects models or of variance inflation parameters to widen the control limits.

The implications of different statistical adjustments and their logical implications in the complex causal structures seen in epidemiologic data will be investigated. Different approaches to overdispersion will have different interpretations and implications and whether a statistical approach has a meaningful interpretation within a discipline defines its utility within that discipline. The appropriateness of statistical adjustment and other approaches in a public health surveillance setting from both an "epidemiologic thinking" (Rothman, 2002) perspective and a methodology perspective are examined and analytical approaches are recommended. These issues are explored in depth in Chapter 2.

1.9.2 Developing a Spatial Scan Statistic

Maximizing the log likelihood ratio is the commonly accepted gold standard in spatial scan statistics. Likelihood ratio based spatial scan techniques appear to suffer from low positive predictive values and tend to include areas near the true clusters with even slightly elevated rates as a part of the cluster (Tango and Takahashi, 2005; Goujon-Bellec *et al.*, 2011). This characteristic appears to be accentuated in many irregularly shaped cluster detection techniques.

I will also consider whether or not a statistic can be motivated from funnel plots. Funnel plots were designed to deal with statistical process control and the identification of change in a process. Historically, statistical process control methods have performed well in small sample situations, suggesting they might provide insights into objective

functions for spatial scan statistics that could improve upon the likelihood ratio finite sample characteristics.

The computational issues associated with exhaustive spatial scan techniques are noted throughout the literature [see (Yiannakoulias, Rosychuk and Hodgson, 2007; Neill, 2012) for recent discussions of the magnitude of the issue]. These issues result in the practical inability to use a technique in public health surveillance, although it may remain of interest in theoretical research. Many of the methods developed to deal with the issue of speed rely on either ad hoc parameters that limit the scope of the search required or restrict the search to relatively compact clusters by use of a non-compactness penalty. The tuning of the ad hoc parameters would be geography-specific or completely arbitrary, resulting in what could be described as a non-reproducible method. The use of a non-compactness penalty or pre-specified shape negates the philosophical public health underpinnings of attempting to identify possibly irregularly shaped clusters. Methods to detect irregularly shaped clusters that do not rely upon arbitrary tuning parameters and yet are computationally feasible are needed.

Spatial scan statistics make explicit use of geographic location through adjacency or distance in the determination of a cluster. Funnel plots are fundamentally a-spatial in nature while potentially identifying specific areas with elevated rates. I now discuss a relationship between these approaches to dealing with spatial variability in aggregate data.

Algorithms in spatial scan statistics have the goal of reducing a computationally infeasible exhaustive search to a computationally feasible search while maintaining the ability to accurately detect the true cluster. The upper level set (Patil and Taillie, 2004) and weak linear time subset scan (Neill, 2012) approach this by ordering the data and searching along this ordering followed by the application of a spatial constraint. Both order the data by the observed rate in the geographic area. This ordering can be seen visually on a funnel plot as following sequentially lower horizontal lines. Figure 1-4 illustrates a funnel plot with the horizontal lines represent the ordering followed by the upper level set and weak linear time subset scanning algorithms. The upper level sets and funnel plot contours both describe an ordering of geographies.

Drawing upon the concept of ordering and then searching, in particular ordering along the data based contours of a funnel plot, I develop a spatial scan statistic. The performance characteristics will be evaluated via simulation with details provided in Chapter 3.

1.9.3 Application to Childhood Immunization Surveillance

It is clear from this review that the factors that influence immunization uptake are many, varied, and likely display complex interactions. Contrast this complexity with the current surveillance activities. These activities generally produce only a single summary measure, such as coverage by age 2 without any accounting for these known determinants. Much of the underlying process cannot be examined or understood using the current suite of surveillance indicators.

I argue that it is necessary for public health practitioners to understand the more complex and subtle influences on coverage in the geographic areas that they are responsible for. Without a full understanding of the factors and levels of factors affecting uptake in their geographies, it will be difficult to appropriately target programmes addressing the relevant barriers to uptake. It will be impossible to monitor performance. If the performance indicators are unable to account for the underlying distribution of immunization related factors, understanding the level of performance will be, literally, meaningless. Consider an example with two geographic areas, one with a large number of religious refusers of immunization and one with no prevalent barriers to immunization. Suppose further that both areas reported 70% coverage at age 2 for a childhood vaccine. How well is the immunization delivery system performing in these geographic areas? For the first area with refusers, 70% could be phenomenal, immunizing everyone who would allow it. In contrast, immunizing only 70% of children in the "average" group might be considered a very poor outcome. Understanding the levels of performance vis-à-vis a rational expectation is important for a performance measure. The type of intervention or public health actions (taken or possible) will also differ significantly between these types of areas – the first area is likely to be a long term investment in building community trust while the second area could benefit from broad campaigns. Understanding the shifts in underlying factors and their impact on uptake and effects on interventions is a key

surveillance objective. All of this information ultimately leads to improved and better public health interventions.

1.9.3.1 An Approach to Immunization Surveillance

The study of immunization coverage in Alberta carried out in this thesis should have characteristics that make the method portable to the public health surveillance environment. Attributes desirable in a surveillance system include simplicity, flexibility, high data quality, acceptability, representativeness, timeliness, stability, and cost effectiveness (Lee *et al.*, 2010). An approach based on aggregate data from existing sources is proposed. Aggregate data offer high quality data themselves, especially from existing public health and other related data systems. Many attributes of the underlying data will be inherited in the aggregate data, for example, timeliness, representativeness, and stability. The use of multiple existing data sources facilitates a simple, acceptable, and flexible system. Finally, use of aggregate data is quite cost effective when compared to primary data collection.

The primary outcome of interest, immunization coverage at age 2, is publicly available and currently part of an operational surveillance system in Alberta. Coverage by age 2 for all of the routine childhood vaccines listed in Table 1-1 are publicly available, at the provincial, zone and local area levels of geography since 2008 in the Alberta Ministry of Health Interactive Health Data Application (IHDA).

Going beyond simple measures of coverage, this thesis focuses on adding factors related to immunization uptake to the surveillance system in addition to improving the analysis of geographic variation. Aggregate data from the Canadian Census and the Canadian Community Health Survey (CCHS) have these desirable attributes. Using aggregate data to combine outcomes with factors related to utilization at small geographies fits well within resource constraints, as it imposes a small marginal cost to the system in terms of analysis but no additional costs in terms of primary collection.

I propose an approach to immunization surveillance that both fits within the resource constraints of surveillance systems and simultaneously provides substantially improved information for public health decision makers. I propose complementing the funnel plot

based approach to understanding geographic variability presented in Chapter 2 with the ability to detect multiple, irregularly shaped spatial-temporal clusters using MultScan from Chapter 3. This approach is illustrated using Alberta childhood immunization coverage data for the four routine doses of DTaP-IPV-Hib measuring coverage at age 2 in 2012 for the 2010 birth cohort.

1.9.3.2 Available Data on Determinants of Immunization Uptake

I now review the factors identified in Table 1-2 that are routinely available in publicly accessible data sources. Direct measurement of the concepts is, of course, preferred, but proxies or highly correlated factors are also effective at accounting for between area variability as described in Chapter 2. The interest is in identifying data that is (a) available for the entire province and broken down at the local geographic level to match the finest geographic level that the outcome measure DTaP-IPV-Hib coverage at age 2, (b) routinely and predictably published at least annually to match the current reporting cycle of immunization coverage, (c) of high quality, and (d) already publicly available to contain costs and for ease of access. Information on the determinants of immunization was identified stemming from three sources: the Canadian Community Health Survey (CCHS), the National Household Survey (NHS), and the Alberta Notice of Birth database. While these databases have been described in details elsewhere, a brief overview of each is now provided.

The CCHS is a Statistics Canada health survey that has been ongoing since 2001. The survey provides provincial samples representative of the age 12 and over population. The survey provides data on many measures of health and determinants of health. Modelled indicators at the local area level are provided by the Alberta Ministry of Health on their public website (Interactive Health Data Application – IHDA). The NHS was the survey replacing the long form census in Canada in 2011. Data aggregated to the local area level have also been released on the IHDA. Finally, the Alberta Notice of Birth data contain information collected on births in the province, including child and maternal outcomes and risk factors. Data aggregated to the local area level have also been released on the IHDA. Details of which determinants of immunization are available from each source are now given.

Of the Predisposing Characteristics, competing pressures, measured by the proportion of households with 3 or more children, is available at the local area level based upon National Household Survey data (further detail on data sources is provided in the next section). For beliefs, a proxy measure of Vitamin K uptake, originally from the Alberta Notice of Birth forms, is available from the IHDA. Refusing Vitamin K prophylaxis has been shown to be strongly associated with under-immunization (Sahni, Lai and MacDonald, 2014) and is conjectured to be related to personal or religious beliefs.

Of the Enabling Resources, the relationship with the provider, measured as having a regular family doctor, is obtained from CCHS data. Income quintiles used to measure material resources are available from NHS data on the IHDA. While education could be used, the consistent insignificant empirical results in multivariate models in Alberta and other jurisdictions previously cited lead to not including it in the model.

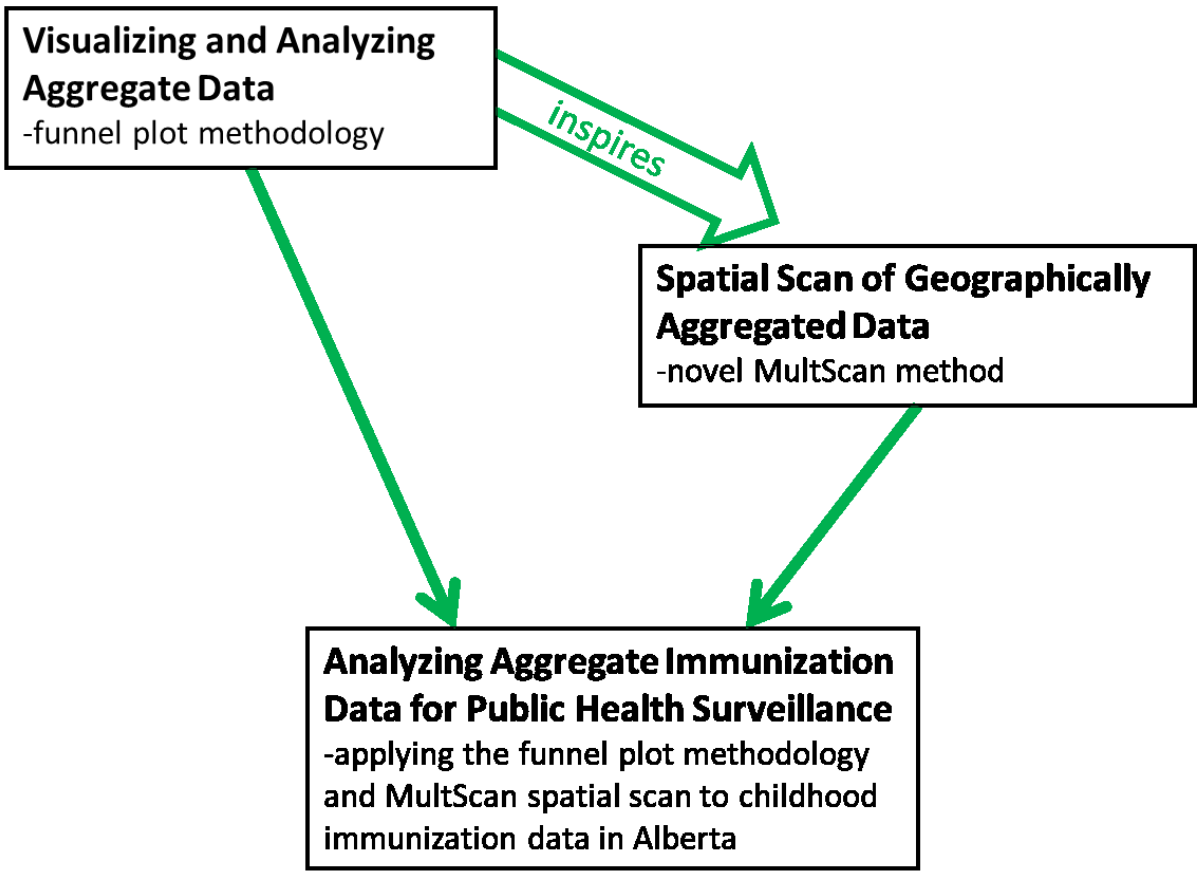
Integration with the Health Care System is measured by the proportion of families having moved within the last 5 years, available from NHS data on the IHDA. The provision of childhood immunization is carried out by public health nurses in Alberta. At the time of the study, Alberta Health Services had been newly formed in 2008 and the former nine Regional Health Authorities (RHA) were transforming into the current five Zones. To best capture the differences in practices, recognizing that organizational change is a slow process, differences in system organization and resources use indicators for each of the former RHAs.

1.10 Interrelationships between the Manuscripts

The first aim focuses on visualizing and analyzing aggregate data to identify areas with high or low rates of disease. Chapter 2 contains my paper clearly delineating a methodology for using funnel plots in public health surveillance. The second aim is to identify specific geographic clusters of high (or low) rates. The spatial scan method that is developed for aim two was inspired by the previous work on funnel plots. Chapter 3 contains my paper developing and evaluating my novel MultScan method for identifying multiple, irregularly shaped clusters in space and time. The funnel plot and spatial scan

methods are then brought together to examine childhood immunization from a surveillance perspective. Chapter 4 contains my paper analyzing childhood immunization using the funnel plot methodology and the MultScan spatial scan statistic, illustrating the utility of these methods in public health surveillance. The interrelationships between these three papers are explicitly laid out in Figure 1-7.

Figure 1-7. Interrelationships between Papers



1.11 REFERENCES

Acosta-Ramírez, N. *et al.* (2005) 'Determinants of vaccination after the Colombian health system reform', *Revista de Saude Publica*, 39(3), pp. 421–429.

Aday, L. A. and Andersen, R. (1974) 'A Framework for the Study of Access to Medical Care', *Health services research*, 9(3), pp. 208–220.

Alberta Health (2017) *2016/17 Influenza Surveillance Report*, *2016/17 Influenza Surveillance Report*. Available at http://www.health.alberta.ca/documents/Influenza-Surveillance-2017-01.pdf.

Andersen, R. and Newman, J. F. (1973) 'Societal and Individual Determinants of Medical Care Utilization in the United States', *The Milbank Memorial Fund Quarterly. Health and Society*, 51(1), p. 95.

Angelillo, I. F. *et al.* (1999) 'Mothers and vaccination: knowledge, attitudes, and behaviour in Italy.', *Bulletin of the World Health Organization*, 77(3), pp. 224–229.

Assunção, R. *et al.* (2006) 'Fast detection of arbitrarily shaped disease clusters.', *Statistics in medicine*, 25, pp. 723–742.

Byrne, B. E. *et al.* (2013) 'Population-based cohort study comparing 30- and 90-day institutional mortality rates after colorectal surgery', *British Journal of Surgery*, 100(13), pp. 1810–1817.

de Cantuária Tauil, M., Sato, A. P. S. and Waldman, E. A. (2016) 'Factors associated with incomplete or delayed vaccination across countries: a systematic review', *Vaccine*. Elsevier, 34(24), pp. 2635–2643.

Center for Disease Control (CDC) (1986) *Comprehensive plan for epidemiologic surveillance.* Atlanta , GA.

Center for Disease Control (CDC) (2011) 'Ten Great Public Health Achievements—United States , 2001-2010', *JAMA*, 306(1), pp. 36–38.

Choi, B. C. K. (2012) ‘The past, present, and future of public health surveillance’, *Scientifica*. Hindawi Publishing Corporation, 2012.

Choi, B. C. K. and Pak, A. W. P. (2001) ‘Lessons for surveillance in the 21 st century: A historical perspective from the past five millennia’, *Sozial-und Pr{ä}ventivmedizin/Social and Preventive Medicine*. Springer, 46(6), pp. 361–368.

Christakis, D. A. *et al.* (2000) ‘The association between greater continuity of care and timely measles-mumps-rubella vaccination.’, *American journal of public health*, 90(6), pp. 962–965.

Declich, S. and Carter, A. O. (1994) ‘Public health surveillance: historical origins, methods and evaluation.’, *Bulletin of the World Health Organization*. World Health Organization, 72(2), pp. 285–304.

Dombkowski, K. J., Lantz, P. M. and Freed, G. L. (2004) ‘Risk factors for delay in age-appropriate vaccination.’, *Public health reports*, 119(2), pp. 144–155.

Dubé, E. *et al.* (2013) ‘Vaccine hesitancy: an overview’, *Human vaccines & immunotherapeutics*. Taylor & Francis, 9(8), pp. 1763–1773.

Duczmal, L. *et al.* (2007) ‘A genetic algorithm for irregularly shaped spatial scan statistics’, *Computational Statistics and Data Analysis*, 52, pp. 43–52.

Duczmal, L. and Assuncao, R. (2004) ‘A simulated annealing strategy for the detection of arbitrarily shaped spatial clusters’, *Computational Statistics and Data Analysis*, 45, pp. 269–286.

Duczmal, L., Kulldorff, M. and Huang, L. (2006) ‘Evaluation of Spatial Scan Statistics for Irregularly Shaped Clusters’, *Journal of Computational and Graphical Statistics*, pp. 428–442.

Falagas, M. E. and Zarkadoulia, E. (2008) ‘Factors associated with suboptimal compliance to vaccinations in children in developed countries: a systematic review’, *Current medical research and opinion*. Taylor & Francis, 24(6), pp. 1719–1741.

Goh, P. *et al.* (2007) ‘Safety and immunogenicity of early vaccination with two doses of tetravalent Measles-Mumps-Rubella-Varicella (MMRV) vaccine in healthy children from 9 months of age’, *Infection*, 35(5), pp. 326–333.

Goujon-Bellec, S. *et al.* (2011) ‘Detection of clusters of a rare disease over a large territory: performance of cluster detection methods’, *International Journal of Health Geographics*, p. 53.

Gower, J. C. and Ross, G. J. S. (1969) ‘Minimum Spanning Trees and Single Linkage Cluster Analysis’, *Applied Statistics*. WileyRoyal Statistical Society, 18(1), p. 54.

Institute of Medicine (2013) *The Childhood Immunization Schedule and Safety: Stakeholder Concerns, Scientific Evidence, and Future Studies*. Washington, D.C.: National Academies Press.

Kim, S. S. *et al.* (2007) ‘Effects of maternal and provider characteristics on up-to-date immunization status of children aged 19 to 35 months.’, *American journal of public health*, 97(2), pp. 259–266.

Kirkham, J. J. and Bouamra, O. (2008) ‘The use of statistical process control for monitoring institutional performance in trauma care’, *Journal of Trauma-Injury Infection & Critical Care*, 65(6), pp. 1494–1501.

Kulldorff, M. (1997) ‘A spatial scan statistic’, *Communications in Statistics - Theory and Methods*, 26(6), pp. 1481–1496.

Kulldorff, M. *et al.* (2006) ‘An elliptic spatial scan statistic.’, *Statistics in medicine*, 25, pp. 3929–3943.

Kulldorff, M. and Nagarwalla, N. (1995) ‘Spatial disease clusters: detection and inference.’, *Statistics in medicine*, 14, pp. 799–810.

Kunadian, B. *et al.* (2009) ‘Funnel plots for comparing performance of PCI performing hospitals and cardiologists: demonstration of utility using the New York hospital mortality data’, *Catheterization and Cardiovascular Interventions*. Wiley Online Library, 73(5), pp. 589–594.

Langmuir, A. D. (1976) ‘William Farr: founder of modern concepts of surveillance’, *International Journal of Epidemiology*. IEA, 5(1), pp. 13–18.

Larson, H. J. *et al.* (2014) ‘Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007--2012’, *Vaccine*. Elsevier, 32(19), pp. 2150–2159.

Last, J. M. (1998) ‘A dictionary of epidemiology, third edition’, *American Journal of Preventive Medicine*, p. 242.

Lee, L. *et al.* (2010) *Principles and Practice of Public Health Surveillance*. 3rd edn. Oxford University Press, USA.

Lemstra, M. *et al.* (2007) ‘Disparity in childhood immunizations.’, *Paediatrics & child health*, 12(10), pp. 847–852.

Li, X. Z. *et al.* (2011) ‘A spatial scan statistic for multiple clusters’, *Mathematical Biosciences*, 233(2), pp. 135–142.

Lieu, T. *et al.* (2015) ‘Geographic clusters in underimmunization and vaccine refusal.’, *Pediatrics TA* -, 135(2), pp. 280–289.

Light, R. J. and Pillemer, D. B. (1984) *Summing up: the science of reviewing research*, *Summing up: the science of reviewing research*. Cambridge, Mass., Harvard University Press.

MacDonald, S. E. *et al.* (2014) 'Parental concern about vaccine safety in Canadian children partially immunized at age 2: A multivariable model including system level factors', *Human Vaccines & Immunotherapeutics,* 10(9), pp.2603–2611.

Marshall, C. E. and Spiegelhalter, D. J. (1998) ‘Reliability of league tables of in vitro fertilisation clinics: retrospective analysis of live birth rates’, *BMJ*, 316, pp. 1701–1705.

Marshall, T., Mohamnmed, M. A. and Rouse, A. (2004) ‘A randomized controlled trial of league tables and control charts as aids to health service decision-making’, *Int J Qual Health Care*, 16(4), pp. 309–315.

Massat, N. J. *et al.* (2015) 'Variation in cervical and breast cancer screening coverage in England: a cross-sectional analysis to characterise districts with atypical behaviour', *BMJ Open*, 5(7), p. e007735.

Matsumura, T. *et al.* (2005) 'Measles vaccine coverage and factors related to uncompleted vaccination among 18-month-old and 36-month-old children in Kyoto, Japan.', *BMC public health*, 5, p. 59.

Mazzucco, W. *et al.* (2017) 'Funnel plots and choropleth maps in cancer risk communication: a comparison of tools for disseminating population-based incidence data to stakeholders', *BMJ open*. British Medical Journal Publishing Group, 7(3), p. e011502.

Mills, E. *et al.* (2005) 'Systematic review of qualitative studies exploring parental beliefs and attitudes toward childhood vaccination identifies common barriers to vaccination', *Journal of clinical epidemiology*. Elsevier, 58(11), pp. 1081–1088.

Neill, D. B. (2012) 'Fast subset scan for spatial pattern detection', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74, pp. 337–360.

Neill, D. B. and Moore, A. W. (2004) 'Rapid detection of significant spatial clusters', *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*, p. 256.

Neutra, R. R. (1990) 'Counterpoint from a cluster buster.', *American journal of epidemiology*, 132(1), pp. 1–8.

Nur, U. *et al.* (2015) 'Inequalities in non-small cell lung cancer treatment and mortality', *Journal of Epidemiology and Community Health*, 69(10), pp. 985–992.

Ohlssen, D. I., Sharples, L. D. and Spiegelhalter, D. J. (2007) 'A hierarchical modelling framework for identifying unusual performance in health care providers', *J R Statist Soc A*, 170, pp. 865–890.

Patil, G. and Taillie, C. (2004) 'Upper level set scan statistic for detecting arbitrarily shaped hotspots', *Environmental and Ecological statistics*, 11, pp. 189–197.

Pavlopoulou, I. D. *et al.* (2013) 'Immunization coverage and predictive factors for complete and age-appropriate vaccination among preschoolers in Athens, Greece: a cross--sectional study.', *BMC public health*. BMC Public Health, 13(1), p. 908.

Pearce, A. *et al.* (2008) 'Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study.', *BMJ (Clinical research ed.)*, 336(7647), pp. 754–7.

Pearce, N. (2000) 'The ecological fallacy strikes back.', *Journal of epidemiology and community health*. BMJ Publishing Group, 54(5), pp. 326–327.

Pearson, M. *et al.* (1993) 'Primary immunisations in Liverpool. II: Is there a gap between consent and completion?', *Archives of disease in childhood*, 69(1), pp. 115–119.

Phillips, K. A. *et al.* (1998) 'Understanding the context of healthcare utilization: assessing environmental and provider-related variables in the behavioral model of utilization.', *Health services research*. Health Research & Educational Trust, 33(3 Pt 1), p. 571.

Que, J. and Tsui, F.-C. (2011) 'Rank-based spatial clustering: an algorithm for rapid outbreak detection.', *Journal of the American Medical Informatics Association : JAMIA*, 18, pp. 218–224.

Rogerson, P. and Yamada, I. (2008) *Statistical Detection and Surveillance of Geographic Clusters*, *Statistical Detection and Surveillance of Geographic Clusters*. Boca Raton, FL, Chapman and Hall.

Rothman, K. J. (2002) *Epidemiology: An Introduction*. 1st edn. Oxford University Press, USA.

Sahni, V., Lai, F. Y. and MacDonald, S. E. (2014) 'Neonatal vitamin K refusal and nonimmunization', *Pediatrics*. Am Acad Pediatrics, pp. 497-503.

De Serres, G. *et al.* (2013) 'Largest measles epidemic in North America in a decade-Quebec, Canada, 2011: Contribution of susceptibility, serendipity, and superspreading events', *Journal of Infectious Diseases*, 207(6), pp. 990–998.

Shen, W. and Louis, T. A. (2000) 'Triple-goal estimates for disease mapping', *Statistics in Medicine*, 19, pp. 2295–2308.

Smith, P. J. and Singleton, J. A. (2008) 'Vaccination Coverage Estimates for Selected Counties: Achievement of Healthy People 2010 Goals and Association with Indices of Access to Care, Economic Conditions, and Demographic Composition', *Public health reports*. Association of Schools of Public Health, 123(2), pp. 155–172.

Speakman, S., McFowland, E. and Neill, D. B. (2015) 'Scalable Detection of Anomalous Patterns With Connectivity Constraints', *Journal of Computational and Graphical Statistics*, 24(4), pp. 1014–1033.

Spiegelhalter, D. J. (2005a) 'Funnel plots for comparing institutional performance', *Statistics in Medicine*, 24, pp. 1185–2102.

Spiegelhalter, D. J. (2005b) 'Handling over-dispersion of performance indicators', *Qual Saf Health Care*, 14, pp. 347–351.

Sterne, J. A. C., Egger, M. and Smith, G. D. (2001) 'Investigating and dealing with publication and other biases in meta-analysis', *BMJ*, 323, pp. 101–105.

Suarez, L., Simpson, D. M. and Smith, D. R. (1997) 'The impact of public assistance factors on the immunization levels of children younger than 2 years.', *American journal of public health*, 87(5), pp. 845–848.

Sui, D. Z. and Holt, J. B. (2008) 'Visualizing and Analysing Public-Health Data Using Value-by-Area Cartograms: Toward a New Synthetic Framework', *Cartographica*, 43(1), pp. 3–20.

Suttorp, V. (2014) 'Outbreaks in Rural Communities with Low Immunization Rates', *Western Canadian Immunization Forum*. Edmonton, AB.

Tango, T. (2008) 'A spatial scan statistic with a restricted likelihood ratio', *Japanese Journal of Biometrics*, 29(2), pp. 75–95.

Tango, T. and Takahashi, K. (2005) 'A flexibly shaped spatial scan statistic for detecting

clusters', *International Journal of Health Geographics*, 4(1), p. 11.

Thomson, A., Robinson, K. and Vallée-Tourangeau, G. (2016) 'The 5As: A practical taxonomy for the determinants of vaccine uptake', *Vaccine*. Elsevier, 34(8), pp. 1018–1024.

Trauth, J. M. *et al.* (2002) 'Do beliefs of inner-city parents about disease and vaccine risks affect immunization?', *Journal of the National Medical Association*, 94(9), pp. 820–832.

Tunstall, H. V. Z., Shaw, M. and Dorling, D. (2004) 'Places and health.', *Journal of epidemiology and community health*, 58(1), pp. 6–10.

Wakefield, J. and Kim, A. (2013) 'A Bayesian model for cluster detection.', *Biostatistics (Oxford, England)*, 14, pp. 752–65.

van der Wal, M. F. *et al.* (2005) 'Vaccination rates in a multicultural population.', *Archives of disease in childhood*, 90(1), pp. 36–40.

Waller, L. A. and Gotway, C. A. (2004) 'Applied Spatial Statistics for Public Health Data', *Environmental Health*, 100, pp. 702–703.

Walter, S. D. (2000) 'Disease mapping: a historical perspective', in *Spatial epidemiology: methods and applications*, pp. 223–239.

Woodall, D. H. (2006) 'The Use of Control Charts in Health-Care and Public-Health Surveillance', *J Qual Technol*, 38(2), pp. 89–104.

Woudenberg, T. *et al.* (2017) 'Large measles epidemic in the Netherlands, May 2013 to March 2014: changing epidemiology', *Eurosurveillance*. European Centre for Disease Prevention and Control, 22(3).

Yiannakoulias, N. (2011) 'Spatial aberration vs. geographical substance: Representing place in public health surveillance', *Health & Place*, 17(6), pp. 1242–1248.

Yiannakoulias, N., Rosychuk, R. J. and Hodgson, J. (2007) 'Adaptations for finding irregularly shaped disease clusters.', *International journal of health geographics*, 6, p. 28.

Zhang, Z., Assuncao, R. and Kulldorff, M. (2010) ‘Spatial scan statistics adjusted for multiple clusters’, *Journal of Probability and Statistics*, 2010, pp. 1–11.

Zhao, Z., Mokdad, A. H. and Barker, L. (2004) ‘Impact of health insurance status on vaccination coverage in children 19-35 months old, United States, 1993-1996.’, *Public health reports*, 119(2), pp. 156–162.

CHAPTER 2: FUNNEL PLOTS IN PUBLIC HEALTH SURVEILLANCE

The following paper has been published as:

> Dover DC, Schopflocher DP. Using funnel plots in public health surveillance. Population health metrics. 2011;9(1):1-11. doi:10.1186/1478-7954-9-58.

This paper develops a deep understanding of the potential utility and methodologic issues in the use of funnel plots in public health surveillance. Funnel plots are examined as a visualization technique. The concept of overdispersion is then scrutinized and an adjustment methodology developed.

2.1 ABSTRACT

2.1.1 BACKGROUND

Public health surveillance is often concerned with the analysis of health outcomes over small areas. Funnel plots have been proposed as a useful tool for assessing and visualizing surveillance data, but their full utility has not been appreciated (for example, in the incorporation and interpretation of risk factors).

2.1.2 METHODS

We investigate a way to simultaneously focus funnel plot analyses on direct policy implications while visually incorporating model fit and the effects of risk factors. Health survey data representing modifiable and nonmodifiable risk factors are used in an analysis of 2007 small area motor vehicle mortality rates in Alberta, Canada.

2.1.3 RESULTS

Small area variations in motor vehicle mortality in Alberta were well explained by the suite of modifiable and nonmodifiable risk factors. Funnel plots of raw rates and of risk adjusted rates lead to different conclusions; the analysis process highlights opportunities for intervention as risk factors are incorporated into the model. Maps based on funnel plot methods identify areas worthy of further investigation.

2.1.4 CONCLUSIONS

Funnel plots provide a useful tool to explore small area data and to routinely incorporate covariate relationships in surveillance analyses. The exploratory process has at each step a direct and useful policy-related result. Dealing thoughtfully with statistical overdispersion is a cornerstone to fully understanding funnel plots.

2.2 BACKGROUND

According to a widely cited definition proposed by the CDC, "Public Health Surveillance is 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 to those who need to know" [1]. The results of analyses conducted on data collected within a surveillance system can be used to inform public health policy and planning, to monitor the health status of a population, and to stimulate research. A functional surveillance system will provide information about the number of health events of specified types that occur within specified populations on an ongoing basis and can therefore be used to derive disease and health event rates over time in different areas (or subpopulations of other types).

One routine surveillance activity may be to monitor rates of disease occurrence in small areas in order to identify anomalies that might have a geographic basis and to enable the reporting of such anomalies to authorities in these areas. Substantial variability in population sizes in small areas introduces some challenges in the comparisons of rates, however, because the precision of estimation of these rates depends on the size of the population over which they are measured.

Several graphical procedures have been proposed for displaying small area rates to support the location of anomalous patterns. League plots [2] and choropleth maps [3] are two common approaches. League plots display observed rates (with confidence intervals) ordered by those rates. These plots are difficult to interpret [4] because they encourage interpretation as a rank ordering, and rank orderings are known to have extremely poor statistical properties [see for example, 5, 6]. Choropleth maps of rates apply differential color schemes to chosen categorizations (often quintiles) of observed rates and color each area on a map according to the category of its observed rate. These are also easy to misinterpret because the map reflects geographic area rather than population density and because the same data may result in maps with very different appearances, since the choice of category is arbitrary. Cartogram versions [7] attempt to redraw areas in proportion to populations but are often difficult to reconcile to geographies and still suffer from the arbitrary category problem.

Funnel plots are an alternative to both league plots and choropleth maps. Funnel plots are a form of scatter plot in which observed area rates are plotted against area populations. Control limits are then overlaid on the scatter plot. The control limits represent the expected variation in rates assuming that the only source of variation is stochastic. The control limits are computed in a fashion very similar to confidence limits and exhibit the distinctive funnel shape as a result of smaller expected variability in larger populations.

Funnel plots were first introduced in meta-analyses, where they are often used to determine whether a lack of a particular type of published findings demonstrates the presence of a publication bias [8, 9]. This would be indicated by the absence of points in a particular region of the funnel (especially an absence of studies with a small sample size and a negative result).

The funnel plot can also be considered a form of control chart [2]. Control charts monitor whether a manufacturing or business process is under control. If analysis indicates that the process is currently stable, with only stochastic variation, then data from the process will vary within known limits and can be used to predict the future performance of the process. If the chart indicates that the data from the process being monitored are too variable, analysis of the chart can help determine the sources of variation, which might then be eliminated to bring the process back into control. In a funnel plot, if rate variation is only random and stochastic, then an appropriate proportion of the points representing area rates will tend to fall within the funnel, and importing control chart terminology, we might consider the (rate generation) process to be "under control." We can also revert to statistical terminology and note that the model fit is adequate (where, in this simple case, the model is of a single stable rate). When many rates fall outside the funnel, the plot can be described as "overdispersed," and it can be said that the process is not in control or the model does not fit the data well. Control chart terminology has been adapted to health system performance in various jurisdictions where it is assumed that managers within a health system can exercise control over a health event-related process [11]. Many of the issues in institutional performance monitoring are shared by health surveillance in support of public health. Both activities deal with small domains, highly variable rates,

large differences in population sizes, multiple testing issues, ongoing monitoring activities, and dissemination of results to interested parties invested with the authority or responsibility to effect change.

It should also be noted that funnel plots are not limited to representing the model of a single stable rate; more complex models can underlie the estimation of the rate or quantity of interest [10]. For example, plotted rates can represent the residuals that remain after a rate, predicted from the values of relevant covariates using a regression model, has been subtracted from the observed rate. In health services research this process is typically called risk adjustment [2, 10, 12].

An ideal model for routine monitoring in health surveillance would begin with a model that fit the data, that is, where the rate generation process could be considered to be under control. Subsequent monitoring over time could focus on whether the rate generation process could be considered to be remaining under control. As well, funnel plots provide a natural, graphical method of assumption checking and model diagnostics during the model development process itself. At any stage, funnel plots may also locate areas with unusually high or low rates (outliers) and this might justify further field epidemiologic or research investigations.

In this paper we demonstrate the use of funnel plots for model development using motor vehicle mortality data in Alberta, Canada. We begin by constructing a funnel plot under the simple model of a single provincial rate and observe that it shows overdispersion. Then we demonstrate a risk adjustment process that largely eliminates this overdispersion. Finally, we discuss steps that emerge from the model that might be taken by public health decision-makers and discuss its use for routine monitoring.

We will speak in terms of small geographies, counts, and rates, and comparisons to an overall rate as these terms are commonly used in health surveillance. However, it should be noted that funnel plots are quite general and can be used for any domain where multiple estimates have been made using varying sample sizes.

2.3 METHODS

2.3.1 DATA

Data are from the province of Alberta, Canada. Alberta is located in Western Canada and has a population of 3,600,000 in 2009. The province maintains a publicly-funded universally-available health care system. All residents of the province (except the military, the Royal Canadian Mounted Police, and federal inmates) are registered with the Alberta Health Care Insurance Plan (AHCIP). This Stakeholder Registry contains demographic information including addresses and therefore provides a source of population estimates by temporal and spatial boundaries.

Maps are based on the Alberta Regional Health Authorities (RHA), reflecting boundary changes introduced in December 2003 and in force until 2009. The small areas analyzed are 70 subregional boundaries created specifically for the analysis of health data [13]. RHA officials were engaged in the process to insure that the subregions would have operational relevance. A population of 20,000 was chosen as a minimum target within each subregion in order to ensure that rates would be relatively stable and this target was met in almost all cases.

The Alberta Vital Statistics Death Registry provides demographic information about each death in Alberta as well as the cause of death according to International Classification of Diseases, 10th revision (ICD-10) codes. The current analysis reports motor vehicle traffic death rates during 2007. Motor vehicle traffic deaths were identified as ICD-10 codes V30-V89 with .5, V39-V79.4, V86.00, and V86.08.

Covariates for risk adjustment (seat belt use; drinking and driving; road type and utilization) are derived from the 2007 cycle of the Canadian Community Health Survey (CCHS), a self-report survey administered annually to approximately 65,000 Canadians (5,000 Albertans) by Statistics Canada [14]. Provincial health ministries are granted special access to location information for respondents in the CCHS sharefile, making it possible to estimate rates at the subregion level by linking CCHS postal codes to subregion boundary files and utilizing the CCHS survey weights.

Drinking and driving is the self-reported proportion of respondent drivers having driven a vehicle after two or more drinks; *seat belt use* is the self-reported proportion of drivers "always" wearing seat belts or passengers "always" wearing seat belts while in the front seat. A proxy for *road type and utilization* was based on the Statistics Canada Metropolitan Influence Zone (MIZ), a measure of the influence a major urban center has upon outlying areas substantially based upon the percent of the population that commutes daily to an urban center. Subregions were assigned the modal MIZ score.

The population, mortality, and survey data are all aggregated and analyzed at the 70 subregional boundaries.

2.3.2 FUNNEL PLOTS

The funnel plots use binomial control limits given by $\hat{p} \pm \Phi(1-\alpha/2)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ where $\Phi(\bullet)$ is the cumulative inverse normal distribution evaluated for 1-α % control limits. Other methods for control limit generation could be used, see [10] for a comprehensive review. To emphasize, $\hat{p}$ is fixed at the overall provincial rate as estimated from the data and n varies freely. The rate for each subregion is then overlaid on the plot at their actual population size and rate.

The funnel plot control limits are set at 95% and 99.8%. These correspond conceptually to the 95% confidence level often used in health services research and to the 3-sigma limits commonly used in process control.

Funnel plots for survey-based measures require a slight modification to account for the complex survey design. The population values are scaled by the particular survey question design effect to account for the additional variability due to the complex survey design [15].

2.3.3 FUNNEL PLOT PRINCIPLES FOR MAPPING

Funnel plots are adapted to mapping through the use of z-scores [3, 16]. The funnel plot based z-scores are computed as $\frac{p_i - \hat{p}}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_i}}}$ where $\hat{p}$ is the provincial rate, p_i is the i^{th} subregion rate, and n_i is the i^{th} subregion population. Values greater than 2 are color-coded orange, values greater than 3 are red, values less than -2 are green, and values less than -3 are dark green. All other values are color-coded yellow. These z-score cut-offs correspond to the 95% and 99.8% control limits in the funnel plot.

2.3.4 RISK ADJUSTMENT

Risk adjustment was carried out using a judgment-based modeling procedure. Covariates that may explain between-region variability in rates were selected *a priori*. Poisson regression on mortality counts with a log(population) offset, a standard method for regression on rates, was carried out sequentially including demographic factors (age, sex), behavioral risk factors (seat belt use, drinking and driving) and environmental factors (proxy for road type and utilization). The adjusted rate is derived using indirect standardization as the product of the provincial crude rate and the ratio of observed to expected values from the relevant regression model [17]. Poisson regression methods are not discussed in any further detail as the focus of this paper is on the use of funnel plots; other sources offer complete discussions of risk adjustment and regression methods [10, 12]. Pearson goodness-of-fit statistics, in addition to the number of small areas outside the control limits, are reported at each stage in the modeling process.

All analyses were carried out in SAS 9.2. The macro code used to create the funnel plots is freely available from the authors.

2.4 RESULTS

Figure 1 shows the funnel plot of crude motor vehicle mortality rates in Alberta in 2007. Of the 70 subregions, 16/70 (23%) fall outside the 95% funnel plot control limits and

6/70 (9%) fall outside the 99.8% limits. If the model of a single provincial rate were correct (i.e., the rate generation process assumed by the model underlying this plot were in control), a funnel plot with 95% control limits could be expected to have three or four (out of 70) areas fall outside the limits. The figure shows substantial overdispersion, and we can conclude that the model does not fit well. That is, there might be additional factors (unmeasured covariates) that differ between subregions contributing to the increased variability of rates between subregions.

In searching for a model with a better fit to the data, we begin by adjusting for demographic factors, age distribution, and sex. Then, we adjust for two well-known behavioral risk factors of motor vehicle mortality for which health surveillance data are regularly available: seat belt use and drinking and driving [18]. Finally, the model is adjusted for the proxy for road type and utilization.

Adjusting for age and sex differences between small areas has improved the model fit substantially, with the Pearson chi-square goodness of fit now down to 1.55 from 3.78 in the raw data. Age distribution and sex are maintained in the model and drinking and driving is included. Drinking and driving has little effect on the model: the goodness of fit is unchanged and the model p-value for drinking and driving is not significant (p=0.41). Removing drinking and driving from the model, the inclusion of seat belt use offers a very slight improvement in fit (goodness of fit 1.52), even though it is not significant at the alpha=0.05 level in the model (p=0.07). Finally, the proxy for road utilization and type is added to the model. There is a substantial improvement in fit and seat belt use is now suggestive in the model (p=0.09). Seat belt use is maintained in the model based on a combination of its public health policy relevance and its suggestive significance level. A final model including age, sex, seat belt use, and proxy for road type and utilization is kept, with a goodness of fit of 1.34, 5/70 subregions outside the 95% limits, and no subregions outside the 99.8% control limits. A funnel plot of age-sex-seat belt use-road adjusted rates is shown in Figure 2. The sequence of models, their goodness of fit statistics, and the number of points outside the 95% and 99.8% limits is shown in Table 1. Figures 3 through 5 show z-score maps of the unadjusted rates, and Figures 6

through 8 show the adjusted rates based on the funnel plot methodology. Figures 3 and 6 show the entire province of Alberta. Figures 4 and 7 show expanded views of the areas within Edmonton, the major population center in the north of the province. Figures 5 and 8 show expanded views of the areas within Calgary, the major population center in the south.

We note here that the unadjusted funnel plot in Figure 1 has three points, at (approximate) populations 15,000, 25,000 and 85,000, that appear to have very large rates as they lie far outside even the 99.8% limits. However, examining the adjusted funnel plot in Figure 2, these points are no longer anomalous after adjusting for age, sex, seat belt use, and road type and utilization.

Adjusting for *modifiable* risk factors, like seat belt use and drinking and driving, leads to further applications of funnel plots. Funnel plots showing these risk factors are also possible, again opening up for their use in anomaly detection and informing the focus of public health interventions or prevention activities. Figure 9 shows the application of the funnel plot to the survey results on seat belt use. This funnel plot focuses attention on areas that may be different from the provincial average seat belt use rate of 88.9%. Figure 10 shows the same points, but the funnel plot limits are adjusted to a target seat belt use rate of 95% [19] as an example. This visually emphasizes the proportion of subregions in the province reaching or not reaching the target level.

Adjusting for *nonmodifiable* risk factors also leads to the ability to clearly explain the difference between the crude and adjusted rates, and hence the potential impact programs could have to those interested in community-level policy and interventions. For example, the crude rate of 66 per 100,000 (population near 15,000 in Figure 1) appears quite high. Figure 11 shows motor vehicle mortality rate adjusted for only the nonmodifiable risk factors age, sex, and road type and utilization. The adjusted rate for the particular subregion is now 29 per 100,000 and inside the 95% funnel plot limits. This subregion has, compared to other parts of the province, a younger, more male population and rural roads with combined effect on motor vehicle mortality rates that can be readily seen. These factors are not likely to be affected by policy or intervention.

Since the CCHS implements a complex survey design, the funnel plots have been adjusted for the design effect of 3.8 for seat belt use in 2007 by multiplying the variance term in the control limits by the design effect. All survey related points are randomly jiggled in the figures and the axis has been suppressed to protect confidentiality as required by Statistics Canada, the statistical agency that owns the data.

2.5 Discussion

One interesting aspect of the funnel plot in Figure 1 is the substantial number of rates for small areas falling outside the funnel plot limits. This overdispersion is not an unusual phenomenon in health data [20]. The ability of the funnel plot to clearly show overdispersion is, we feel, one of the most useful aspects of the funnel plot. We can immediately and visually see that we don't fully understand the disease process. This judgment should be considerably easier than judgments of the presence or absence of publication bias when considering funnel plots of effect sizes from a meta-analysis, which depends upon the distribution of points within the funnel limits and is therefore quite error prone [21].

Funnel plots are therefore extremely useful in focusing analysts' attention on model misspecification. When overdispersion is observed, the key question becomes what to do with the apparent overdispersion. Some have advocated the use of statistical correction [22, 23] to adjust for overdispersion, either through random effects models, via an overdispersion parameter, or both. We feel this should be an approach of last resort only. If there is large variability in a health variable being monitored, adjustment via the inclusion of missing covariates should be the first line of attack. We note that this adjustment need not be directly causally based. For example, if seat belt use data were not available, but a similar risk taking behavior variable was, that proxy variable could still have served to substantially explain the variability in motor vehicle mortality rates. With the plethora of survey and administrative data available today, there is no reason not to attempt to understand and model the factors affecting between-region variability before resorting to random effects-type models. Also, these blind approaches to

overdispersion carry substantial risk in the surveillance arena. In the case of misspecification due to a missing covariate, random effects models make the strong assumption that the missing covariate value is essentially proportional to the observed rate [24]. It is very easy for this not to be the case in practice, as illustrated in our example. In fact, had further attempts at adjustment not been made, the interpretation of the funnel plot would have pointed public health epidemiologists to the wrong area. The purported statistical approach to fixing overdispersion must be used with great caution. Echoing Berk et al [25], "one risks an arbitrary correction leading to arbitrary results." In the motor vehicle mortality example, one small area would still have been outside the 99.8% controls limits if an overdispersion factor had been included, even though our analysis shows that this was not any sort of outlier but simply has a poor combination of age, sex, seat belt use, and road type. In the analysis presented, the final model shows good fit. Had this not been the case, it would have been possible to include random effects or an overdispersion factor in the final model. Future surveillance and monitoring efforts could continue, keeping the random effects and/or overdispersion value fixed. Attempts to dynamically alter the overdispersion parameter or re-predict random effects might only mask any real changes over time.

The funnel plot methodology encourages the use of data from multiple sources. Funnel plots in the surveillance domain can rely on aggregate data, making the linking process between data sources much easier to facilitate. In our example, we were able to seamlessly integrate survey and administrative data sources because they are only required to be available at the aggregate subregion level. This also has implications for ongoing monitoring: with systems in place to create small area estimates from a variety of data sources, ongoing monitoring and creation of future funnel plots should be possible.

Underlying the outlined funnel plot methodology is the choice of method for creating the limits. The asymptotic normal approximation was used. This choice may appear unusual in light of the fact that, for binary confidence intervals, the use of the asymptotic normal approximation is generally not recommended as it can have very poor performance characteristics [see 26 for a recent review]. Ongoing research by the authors suggests that

it is the opposite case for funnel plots and the asymptotic normal approximation outperforms other methods for creating limits.

The funnel plot methodology was also successfully adapted to a mapping framework. The ability to display surveillance data in a geographic context can aid in the understanding of that data. The maps, combined with expert knowledge of the areas, can generate suggestions as to what factors may explain any residual overdispersion.

Following the institutional performance literature, funnel plots of disease rates, risk factors, or changes in these could also be used as performance measurement tools [10]. Using target rates as the funnel plot center line and placing the funnel around them gives an indication of how many small areas are likely achieving a public health target. The funnel plot of seat belt use rates around a target of 95% in Figure 10 gives a visual representation of both the range of seat belt use rates and the number of areas where seat belt use is below the target.

The recommended analysis process employing sequential funnel plots and multiple covariates lends itself to identifying opportunities for policy recommendations. For a covariate that does enter the model, there is evidence that the covariate varies across the province, naturally suggesting that further analysis of this covariate might identify local area level intervention and policy opportunities. If a known risk factor does not enter the model, a global policy level recommendation may be in order. In our example, drinking and driving did not enter the final model, suggesting that policy recommendations could be made at the provincial level; while seat belt use, which did enter the final model, lends itself to local level interventions. Further consideration of factors as modifiable or nonmodifiable facilitates the interpretation of individual small area rates. Adjusting for nonmodifiable risk factors allows a clear comparison to crude rates and highlights the potential for improvement through modifiable factors. Assessing the modifiable factors through their own funnel plots can help target local area level interventions and policy initiatives.

The use of funnel plots and modeling to assess the relationships between potential risk factors and outcomes must always be carried out with care. The process described

employs an ecological model and carries with it the potential limitations and cautions of this type of design [see for example, 27]. Particular care should be taken in interpreting the meaning of any coefficients in the model to avoid the ecological fallacy. We have framed the process as a surveillance activity where there is usually an evidential basis for inclusion of risk factors or proxies for risk factors. Clearly any single ecological correlation would be insufficient evidence to justify public health action, but when noted in the context of established risk factors, public health activities may be reasonable.

We envision three key areas for the evolution of the funnel plot in public health surveillance. The first area is the integration of the funnel plot into ongoing monitoring activities over time. We have touched on issues regarding the use of random effects and overdispersion parameters as they relate to repeated applications of a funnel plot over time. The questions related to incorporating modeling into a funnel plot-based surveillance process (Re-run the model each year with additional data? Hold coefficients constant over time? How best to display multiple years of data?) are an area of active inquiry. A related area for evolution of the funnel plot is how to appropriately incorporate funnel plots into a multilevel model framework. As multiple levels of data are becoming available for analysis in surveillance, multilevel models will become more common. Finally, funnel plots have a close link to spatial data as they are currently used in public health surveillance. The ties, theoretical and applied, to spatial methods provide a large area for future contributions.

2.6 Conclusions

Funnel plots and their cartographic equivalents provide visually attractive means of displaying small area data in health surveillance and other disciplines for the purposes of anomaly detection and ongoing monitoring, while accounting for variation in small samples. Overdispersion, readily apparent when present in funnel plots, needs to be dealt with thoughtfully in the analysis and modeling stages of surveillance to ensure that the interpretation of the surveillance data is appropriate. The use of funnel plots in health

surveillance modeling activities naturally focuses attention to the level that policy recommendations should be made.

2.6.1 Competing interests

The authors declare they have no competing interests.

2.6.2 Authors' contributions

DCD and DS contributed to the design and conceptualization of the study, the interpretation of the data, and the writing of the manuscript. DCD performed the statistical analysis. All authors have approved of the final draft of the manuscript.

2.6.3 Acknowledgments

This work was made possible by a grant from Alberta Health and Wellness to DS. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of Alberta Health and Wellness or the University of Alberta.

2.7 REFERENCES

1. CDC: Guidelines for evaluating surveillance systems. *MMWR*, 1988, 37(S-5).

2. Woodall, DH: The Use of Control Charts in Health-Care and Public-Health Surveillance. *J Qual Technol* 2006, 38(2):89-104

3. Rogerson P, Yamada I: *Statistical detection and surveillance of geographic clusters.* Hoboken, NJ, Taylor & Francis, 2008

4. Marshall T, Mohamnmed MA, Rouse A: A randomized controlled trial of league tables and control charts as aids to health service decision-making. *Int J Qual Health Care*, 2004, 16(4):309-315

5. Marshall CE, Spiegelhalter DJ: Reliability of league tables of in vitro fertilisation clinics: retrospective analysis of live birth rates. *BMJ* 1998, 316:1701-1705

6. Shen W, Louis TA: Triple-goal estimates for disease mapping. *Statistics in Medicine* 2000, 19: 2295–2308

7. Sui DZ, Holt JB: Visualizing and Analysing Public-Health Data Using Value-by-Area Cartograms: Toward a New Synthetic Framework. *Cartographica* 2008, 43(1):3-20

8. Light RJ, Pillemer DB: *Summing up: the science of reviewing research.* Cambridge, Mass., Harvard University Press, 1984

9. Sterne JAC, Egger M, Smith GD: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001, 323:101–105

10. Spiegelhalter, DJ: Funnel plots for comparing institutional performance. *Statistics in Medicine* 2005, 24:1185-2102

11. Benneyan JC, Lloyd RC, Plsek PE: Statistical process control as a tool for research and healthcare improvement. *Qual Saf Health Care* 2003, 12:458–464

12. Iezzoni, LI (ed.): *Risk Adjustment for Measuring Health Care Outcomes, 3rd edition.* Chicago, IL, Health Administration Press, 2003.

13. Alberta Health and Wellness: *Calculating Small Area Analysis: Definition of Sub-regional Geographic Units in Alberta.* Alberta, 2003, accessed on-line at http://www.health.alberta.ca/documents/Geo-Calculating-Small-Area-2003.pdf

14. Statistics Canada: *Canadian Community Health Survey (CCHS).* accessed on-line at http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SurvId=3226&SurvVer=0&InstaId=15282&InstaVer=4&SDDS=3226&lang=en&db=IMDB&dbg=f&adm=8&dis=2

15. Korn EL, Graubard BI. Analysis of Health Surveys. New York, NY, Wiley, 1999

16. Rogerson P, Yamada I. Statistical Detection and Surveillance of Geographic Clusters. Boca Raton, FL, Chapman and Hall, 2008

17. Gail M, Benichou J, editors. Encyclopedia of Epidemiologic Methods. Chichester, England, Wiley, 2000, p.872

18. Barss P, Smith GS, Baker SP, Mohan D: *Injury prevention : an international perspective epidemiology, surveillance, and policy.* New York, Oxford University Press, 1998

19. Ministry of Transportation, Government of Canada: *Vision 2010 – Making Canada's Roads the Safest in the World*. 2002, accessed on-line at http://www.ccmta.ca/english/pdf/rsv_report_02_e.pdf

20. Birkmeyer JD: Primer on Geographic Variation in Health Care. *Effective Clinical Practice* 2001, 4(5):232-233

21. Terrin, N, Schmid, CJ, Lau J: In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology* 2005, 58: 894–901

22 Spiegelhalter, DJ: Handling over-dispersion of performance indicators. *Qual Saf Health Care* 2005, 14:347-351

23 Kim H, Kriebel D: Regression models for public health surveillance data: a simulation study. *Occup Environ Med* 2009, 66:733–739

24 Ohlssen DI, Sharples LD, Spiegelhalter DJ: A hierarchical modelling framework for identifying unusual performance in health care providers. *J. R. Statist. Soc. A* 2007, 170:865-890

25. Berk R, MacDonald J: Overdispersion and Poisson Regression. *Journal of Quantitative Criminology* 2008, 24: 269–284

26. Pires AN, Amado C: Interval estimators for a binomial proportion: comparison of twenty methods. *Revstat* 2008, 6(2):165-197

27. Rothman KJ, Greenland S, Lash TL: *Modern Epidemiology (3rd ed.)*. Philadelphia, PA Lippincott Williams & Wilkins, 2008

Figure 2-1. Funnel plot of motor vehicle traffic mortality rates

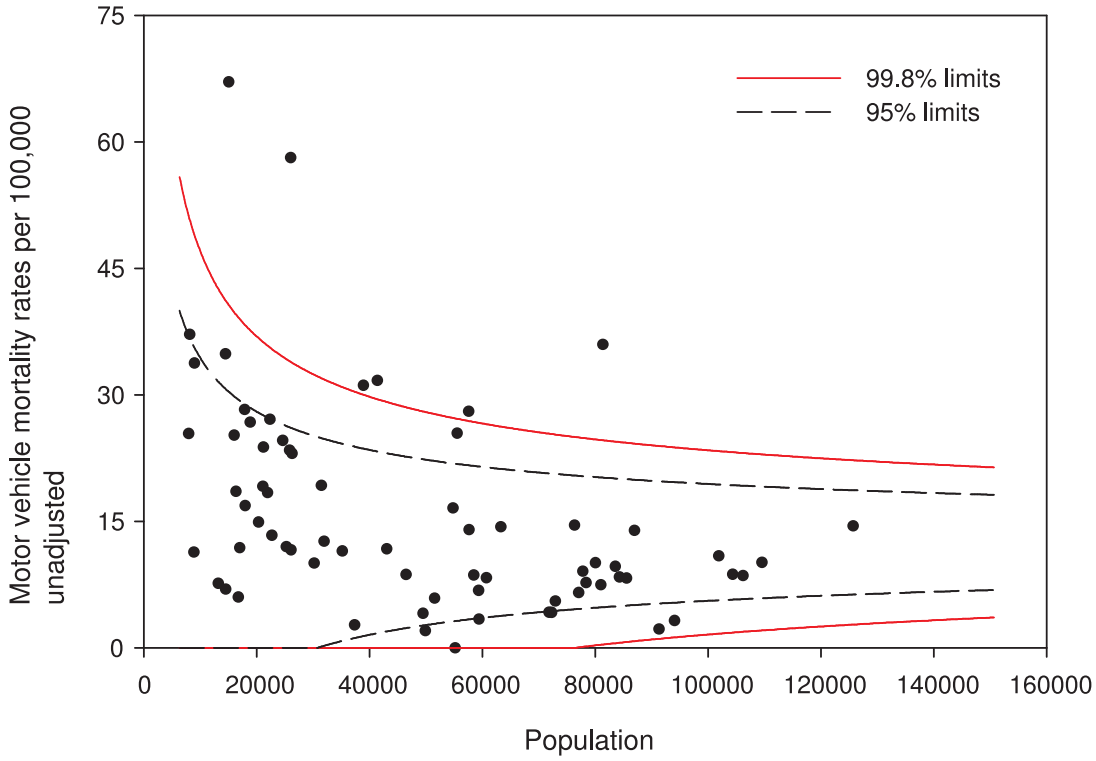
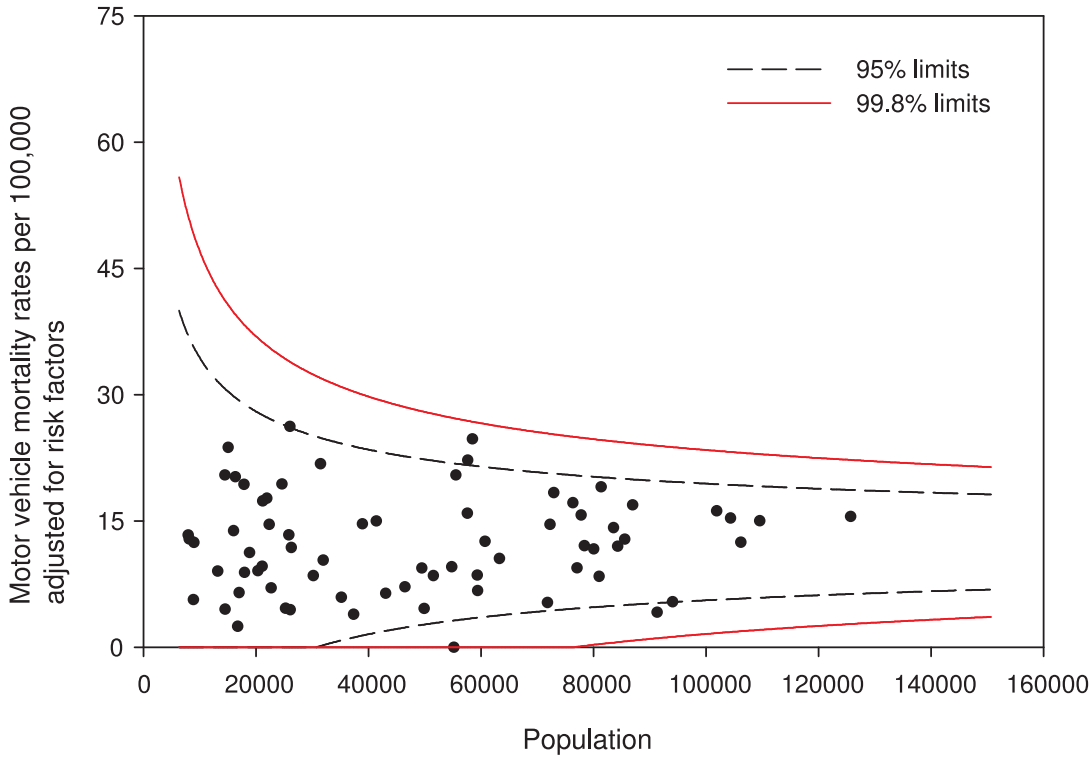


Figure 2-2. Funnel plot of adjusted motor vehicle traffic mortality rates



Motor vehicle mortality rates are adjusted for age, sex, seat belt use, and road type and utilization.

Figure 2-3. Alberta map of unadjusted z-scores, motor vehicle traffic mortality rates

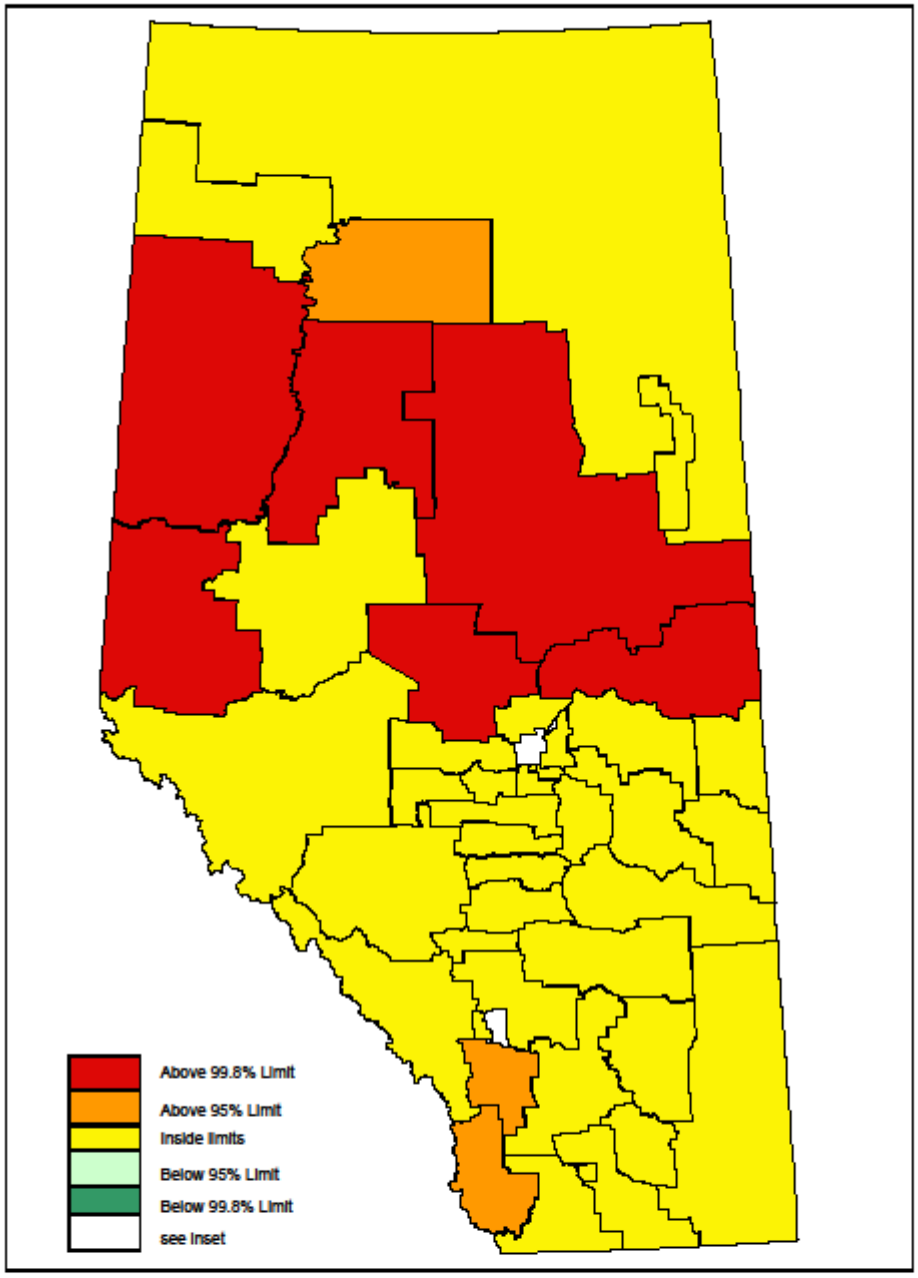


Figure 2-4. Edmonton map of unadjusted z-scores, motor vehicle traffic mortality rates

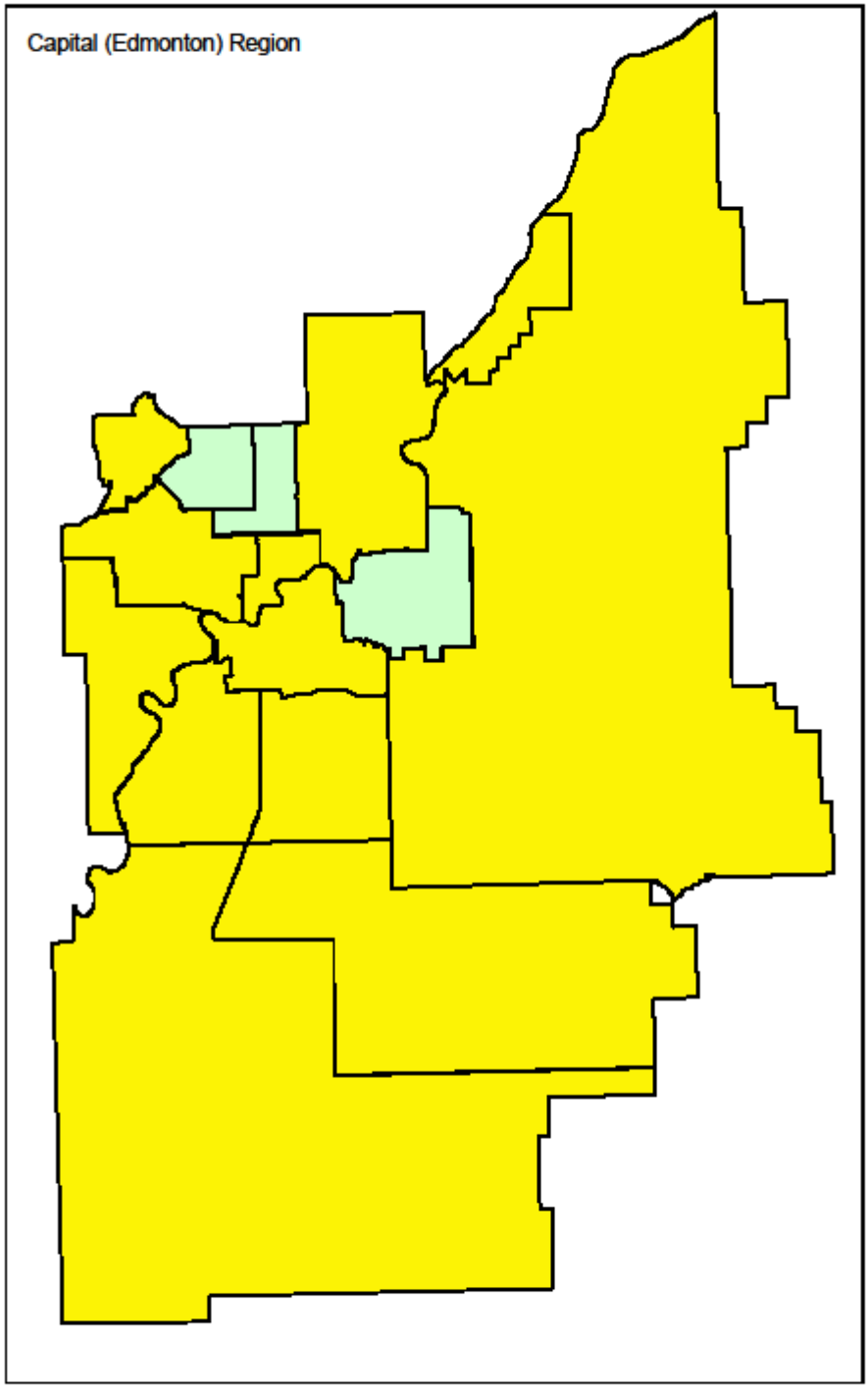


Figure 2-5. Calgary map of unadjusted z-scores, motor vehicle traffic mortality rates

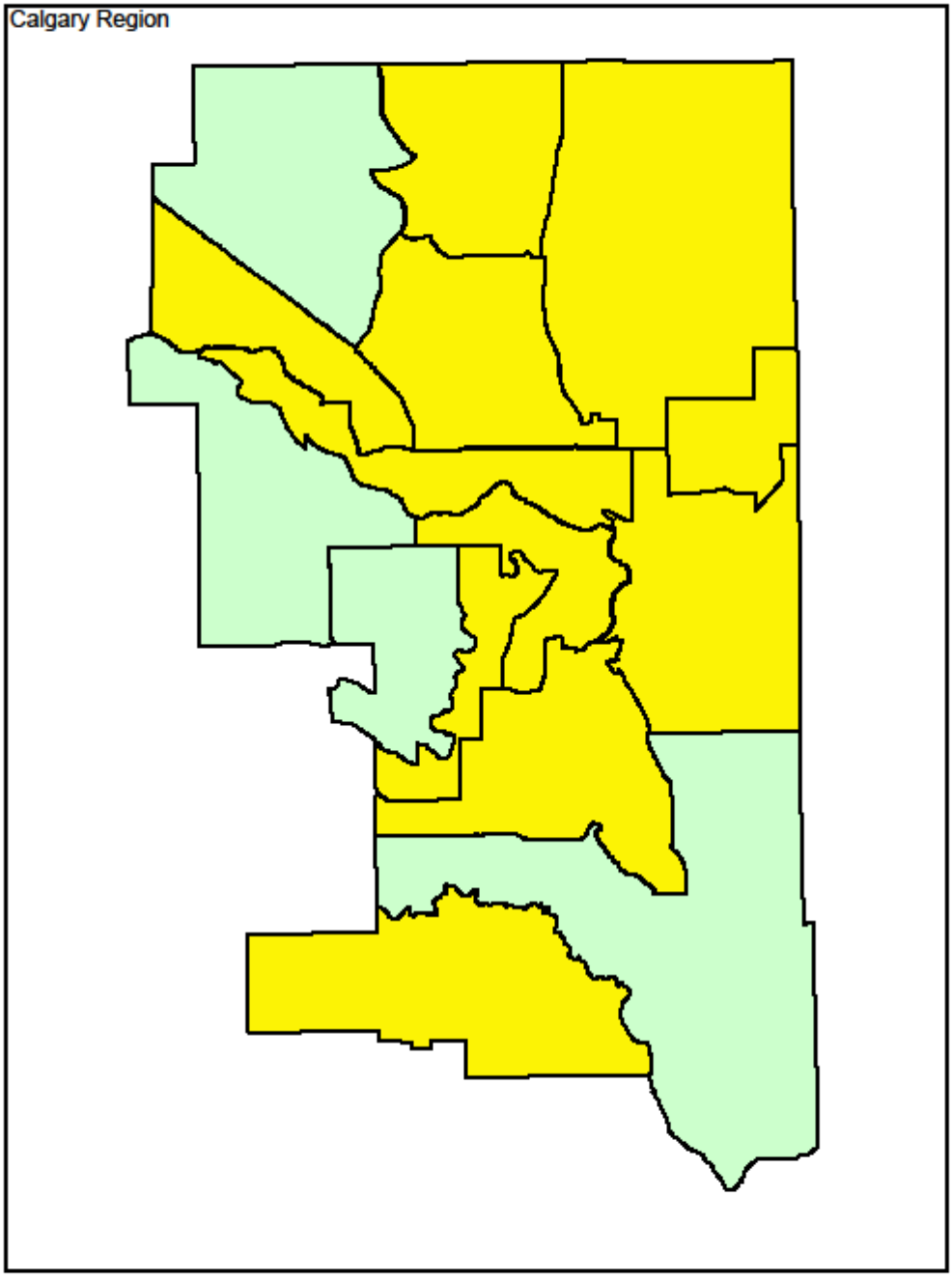


Figure 2-6. Alberta map of adjusted z-scores, motor vehicle traffic mortality rates

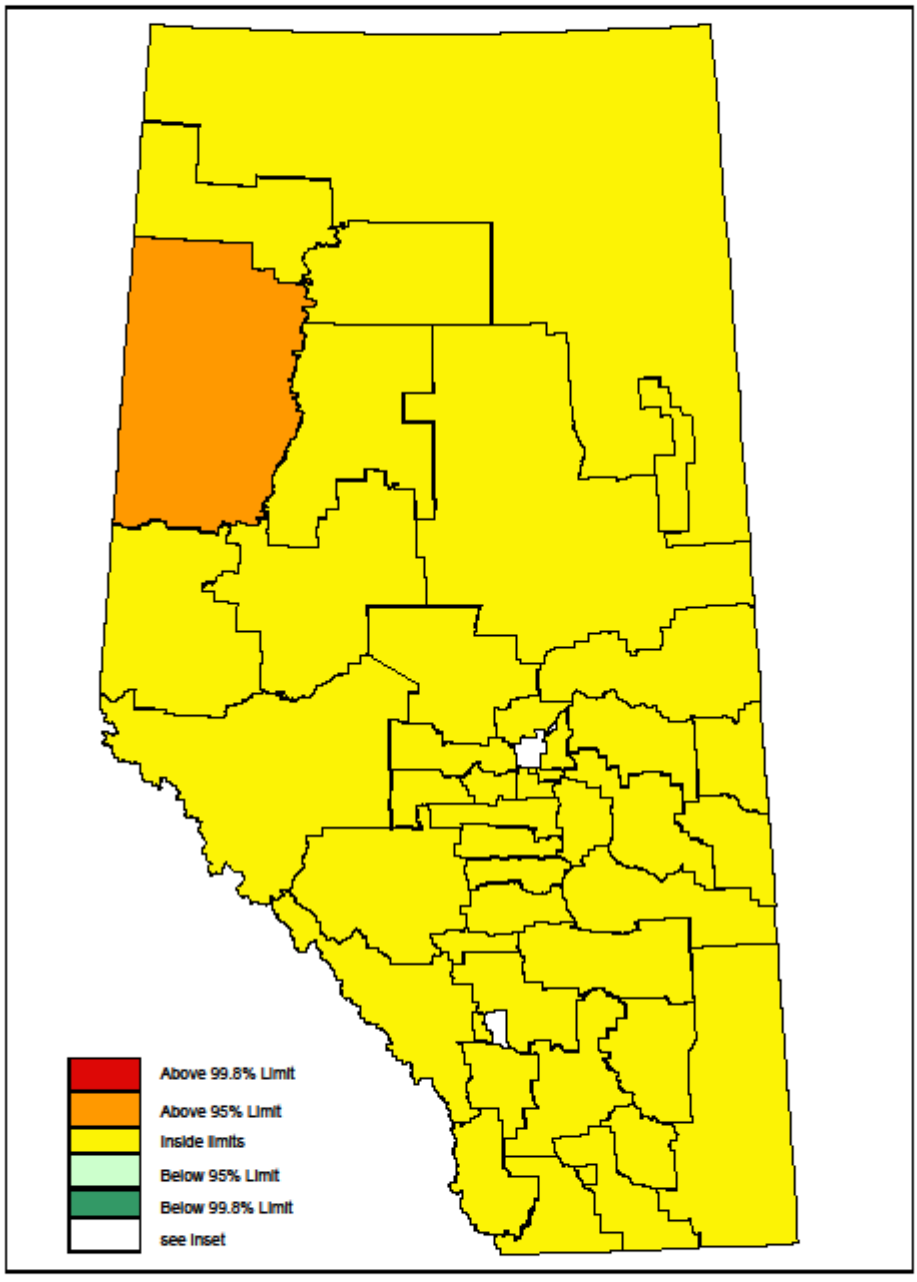


Figure 2-7. Edmonton map of adjusted z-scores, motor vehicle traffic mortality rates

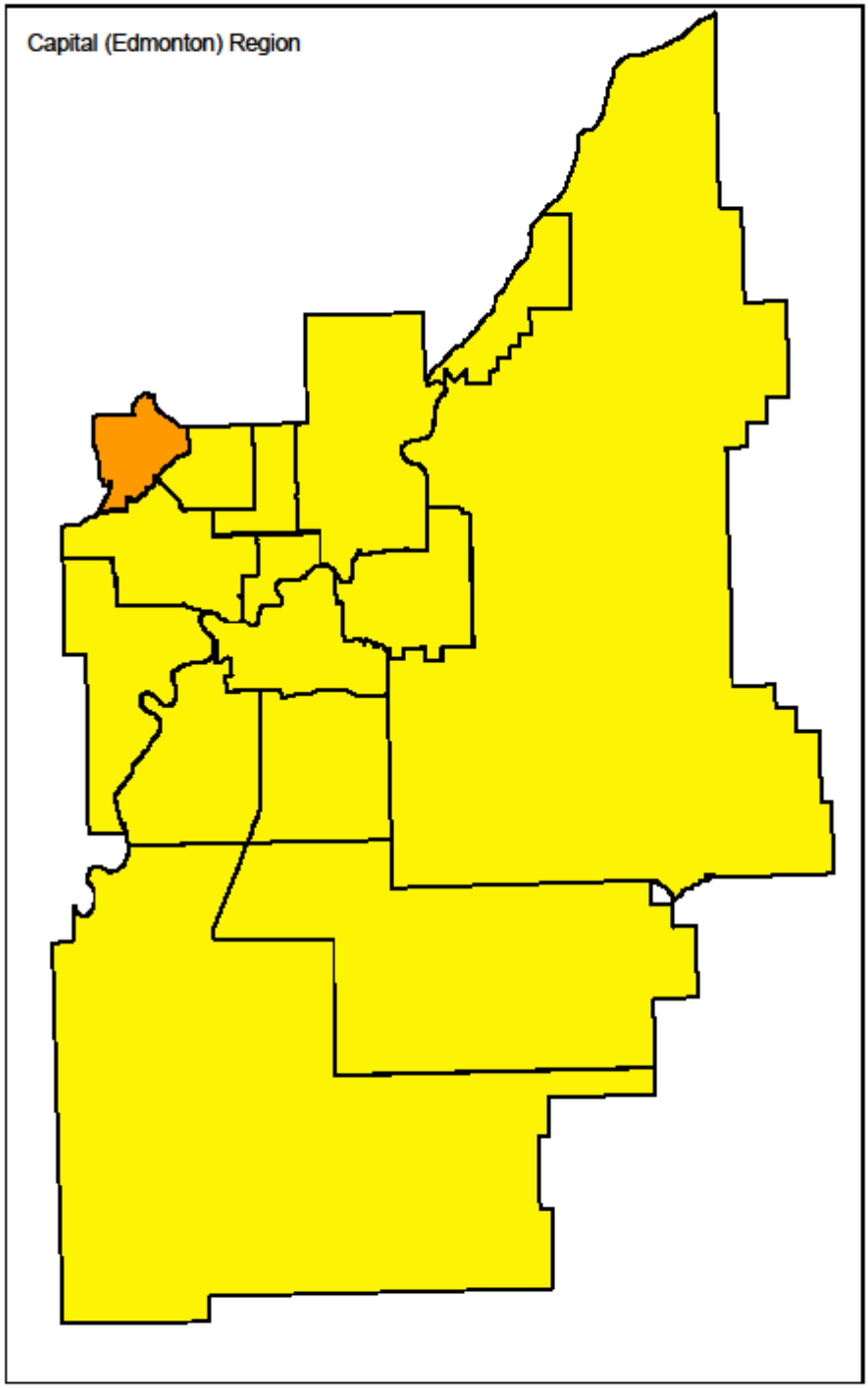


Figure 2-8. Calgary map of adjusted z-scores, motor vehicle traffic mortality rates

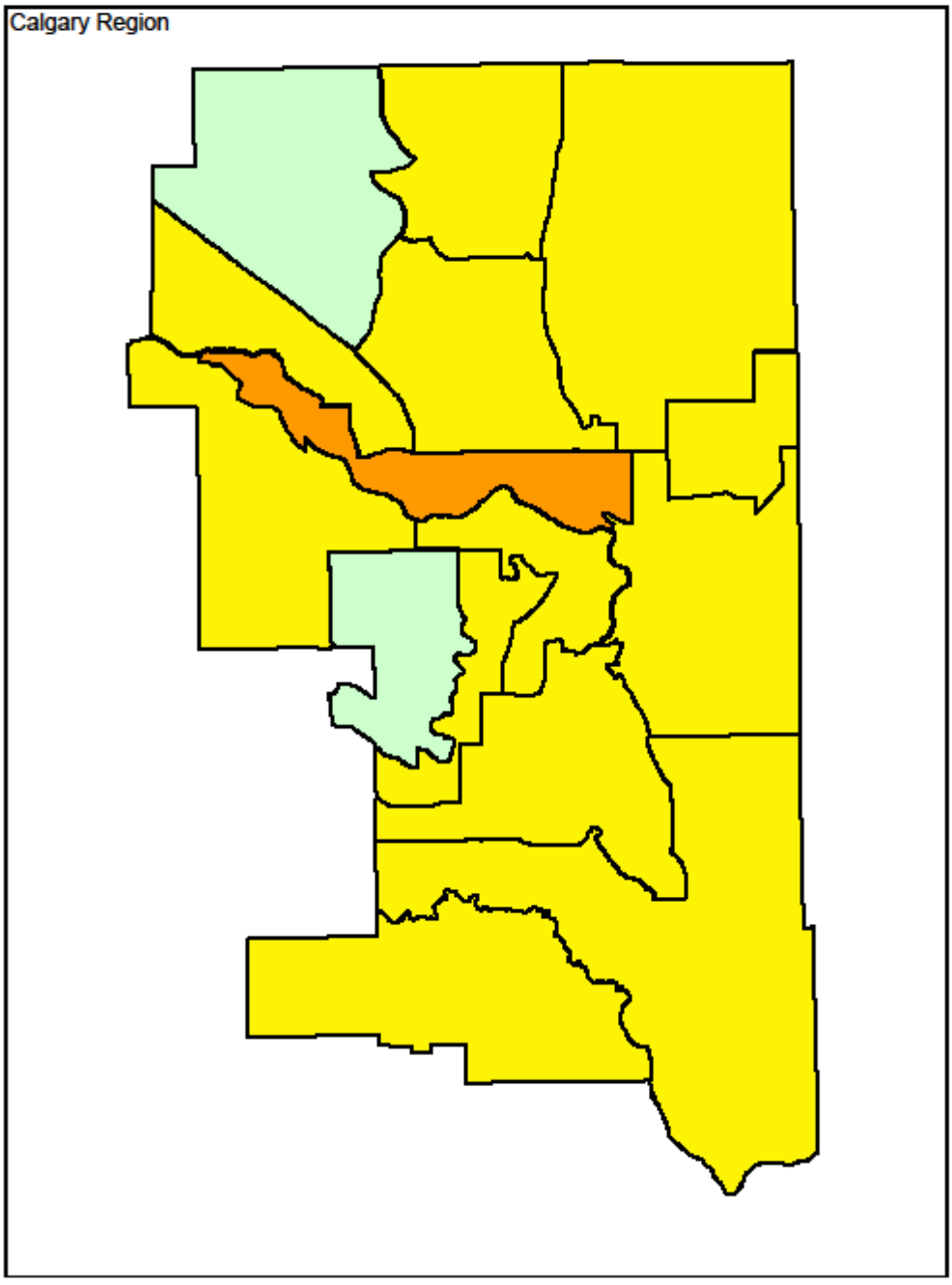
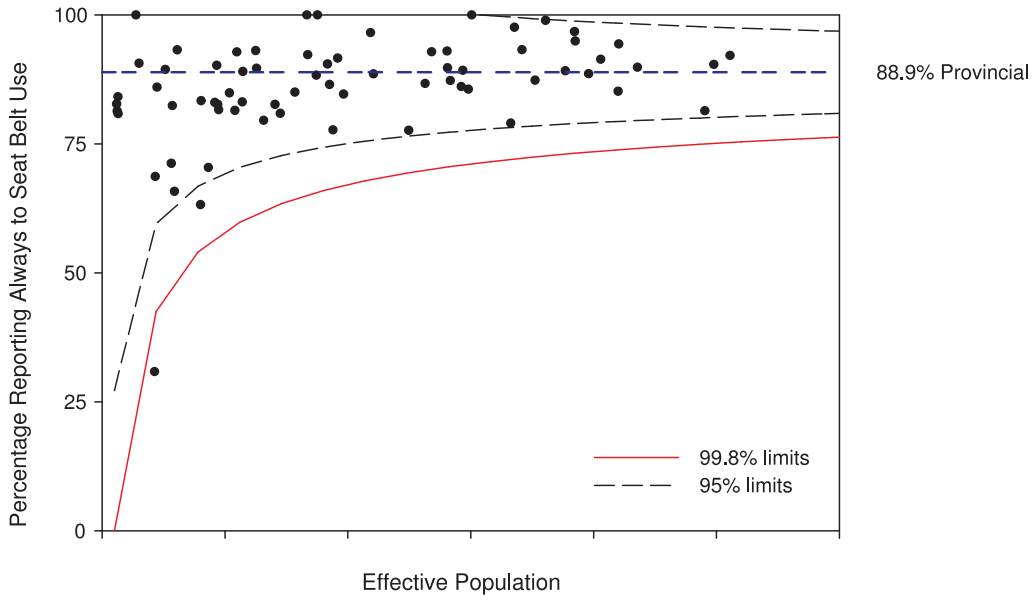
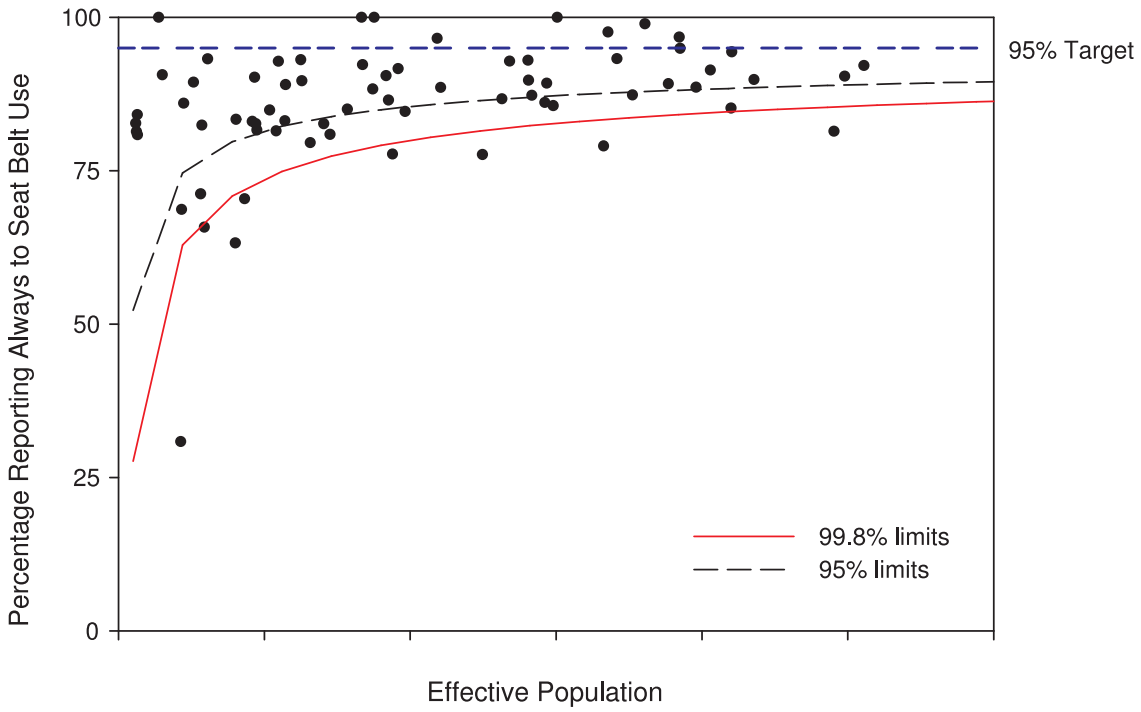


Figure 2-9. Funnel plot of seat belt use rates



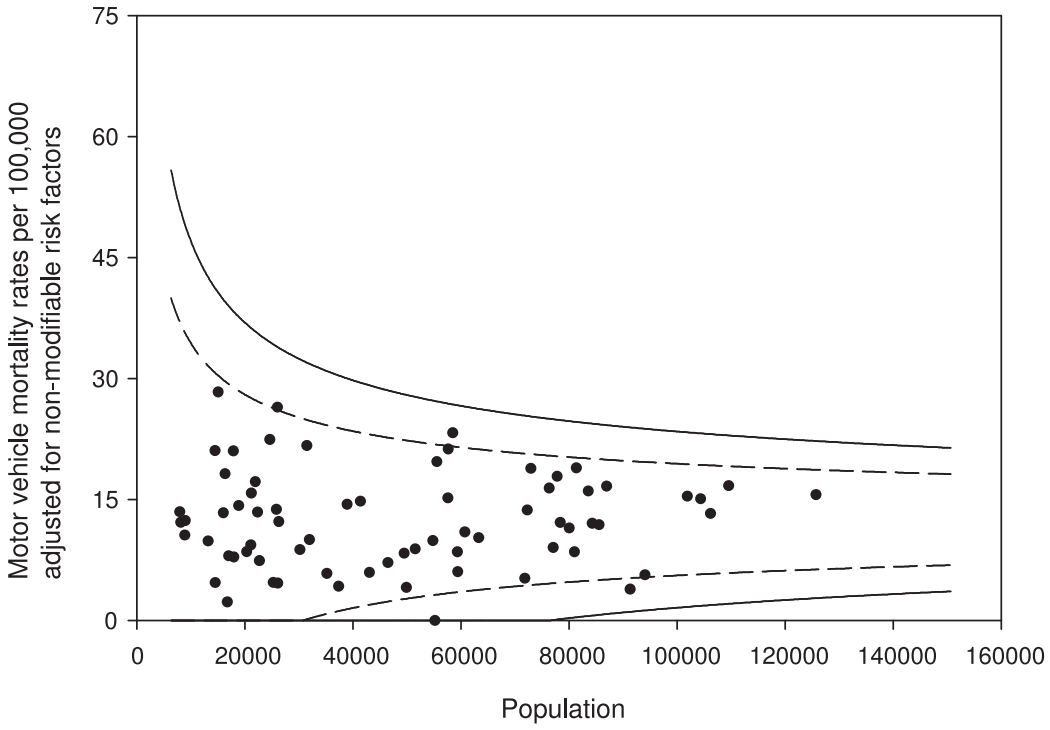
The x-axis values have been suppressed to maintain confidentiality.

Figure 2-10. Funnel plot of seat belt use rates centered at target 95%



The x-axis values have been suppressed to maintain confidentiality.

Figure 2-11. Funnel plot adjusted for nonmodifiable risk factors



Motor vehicle mortality rates are adjusted for age, sex, and road type and utilization.

Table 2-1. Modeling between subregion variation in motor vehicle traffic mortality rates

		Scaled Pearson goodness of fit*	Outside 95% limits (#)	Outside 99.8% limits (#)
Unadjusted		3.78	16	6
Adjusting for:				
	Age, sex	1.55	9	2
	Age, sex, drinking and driving	1.55	8	1
	Age, sex, seat belt use	1.52	7	2
	Age, sex, seat belt use, road type and utilization	1.34	5	0

*the scaled Pearson goodness of fit measures is scaled by the degrees of freedom so that the expected value of 1 represents good model fit.

CHAPTER 3: SPATIAL SCAN STATISTICS

An edited version of the following paper has been submitted for publication as:

Multiple, Irregular, Spatial-Temporal Cluster Identification in Public Health Surveillance

Douglas C. Dover

School of Public Health, University of Alberta, Edmonton, Alberta

Alberta Ministry of Health, Edmonton, Alberta

This paper develops and evaluates the MultScan spatial scan statistic. First, the funnel plot motivation is discussed. The MultScan algorithm for multiple clusters is described, along with MultScan-single, a single cluster detection variation. MultScan is evaluated through a number of simulations examining different cluster shapes in both space and time, differing numbers of clusters, and different underlying disease rates inside and outside the cluster. Finally, MultScan is briefly demonstrated on immunization data in Alberta. The last section of this Chapter presents a technical description of spatial scan statistics.

3.1 Abstract

Spatial scan statistics are important tools in the public health surveillance methodologists' arsenal. The ability to accurately identify clusters of health events and inform action is a foundational role in public health surveillance. The most common spatial scan methods today use fixed shapes in the spatial or temporal dimensions and suffer from low accuracy. A novel spatial scan is proposed based upon insights from health performance monitoring. The method can simultaneously detect multiple clusters that are irregularly shaped in space and time. Performance is evaluated through simulations of public health relevant clustering scenarios and through the application to immunization coverage in Alberta, Canada. The new method was able to increase accuracy as measured by positive predictive value while maintaining computation speed and sensitivity.

3.2 Introduction

Cluster detection is widely used in public health surveillance. After reviewing current cluster detection techniques, a cluster detection method motivated by statistical process control concepts that is capable of rapidly identifying multiple clusters having irregular spatial-temporal shapes is proposed. Its performance characteristics are examined through simulation and a real data example.

Identifying geographic clusters of elevated health risk leads to further understanding disease aetiologies (Waller and Gotway, 2004), addressing community concerns (Center for Disease Control (CDC), 1990; Government of Alberta, 2011), and better targets public health interventions (Lieu *et al.*, 2015; Groseclose and Buckeridge, 2017). Spatial scan statistics allow public health analysts to identify the location of a geographic cluster of excess risk. Accuracy is a key characteristic in any spatial scan method since these results help allocate public health resources. If methods erroneously identify a cluster larger than the true cluster or identify one when none exist, resources may be diverted from real public health issues to address artefactual problems. When methods fail to

identify or identify only a portion of the true cluster, insufficient resource allocation may lead to the continuance of the health issue.

The most widely used spatial scan is Kulldorff's (Kulldorff, 1997) circular scan (KSS), a method able to simultaneously assess the statistical significance, excess rate, and location of a geographic cluster. KSS performs quite well when the underlying cluster is circular, but accuracy drops off dramatically when the underlying cluster is no longer compact (Goujon-Bellec *et al.*, 2011). Clusters of public health importance occurring in space and time will almost certainly be irregularly shaped and temporally dynamic. As the underlying exposure changes over time, so will the placement of any excess health events. For example, a communicable disease can see localized rates sharply increase and then decrease, while geographic clustering of behavioural risk factors will change with demographic shifts, changes in social norms, and in response to public health actions. While a variety of other spatial shapes have been investigated [ellipses (Kulldorff *et al.*, 2006), squares (Neill and Moore, 2004)], Tango (Tango and Takahashi, 2012) directly addressed the issue of irregularly shaped clusters with an exhaustive search approach and many other computationally feasible approaches have since been proposed (Duczmal and Assuncao, 2004; Patil and Taillie, 2004; Assunção *et al.*, 2006; Duczmal *et al.*, 2007). These spatial shapes have been extended temporally, for example, extending a circle to a cylinder in time. Irregular shapes have been analogously extended (Tango and Takahashi, 2005). A dynamic shape placed in time using rectangular pyramids has also been proposed (Iyengar, 2005). This leaves largely unexplored the area of fully irregularly shaped spatial-temporal spatial scan methods.

A common feature to all of these likelihood-based methods is low positive predictive value resulting from capturing many geographies that are not part of the true cluster. They regularly include neighbouring geographies with only slightly elevated rates in the cluster or combine geographies to obtain clusters that are "unrealistically large" and can have "peculiar shapes" (Tango, 2008). Methods to deal with these features using arbitrary tuning parameters, for example by limiting the search depth (Yiannakoulias, Rosychuk and Hodgson, 2007; Tango, 2008) or penalizing the irregularity (Duczmal, Kulldorff and Huang, 2006), have only been moderately successful.

As the volume of electronically available data continues to increase in public health surveillance repositories worldwide, it is becoming common to search for multiple clusters to guide public health interventions and programmes. The most common approach to assessing more than one cluster is to begin with the most likely cluster and then order all other searched clusters by their test statistic. Every subsequent cluster that is statistically significant (using the single cluster test values) and has no overlap with previously identified significant clusters is considered significant (Zhang, Assuncao and Kulldorff, 2010). This is a statistically conservative approach. A sequential approach to controlling the significance levels limited to identifying up to three clusters has been developed (Li *et al.*, 2011). A Bayesian approach (Wakefield and Kim, 2013) placing a prior distribution on the number of clusters correctly addresses the issue of significance level for any number of simultaneous clusters but is quite sensitive to the choice of prior. Strategies to ensure detection at the correct significance level are still required.

A spatial-temporal scan method inspired by applications of statistical process control in public health surveillance is proposed. It allows the detection of possibly multiple, irregularly shaped, spatial or spatial-temporal clusters while using a data-based approach to limit the computational burden and increase the positive predictive power. Performance is examined through a series of simulations and comparisons to Kulldorff's spatial scan. The method is applied to a real scenario examining childhood immunization rates.

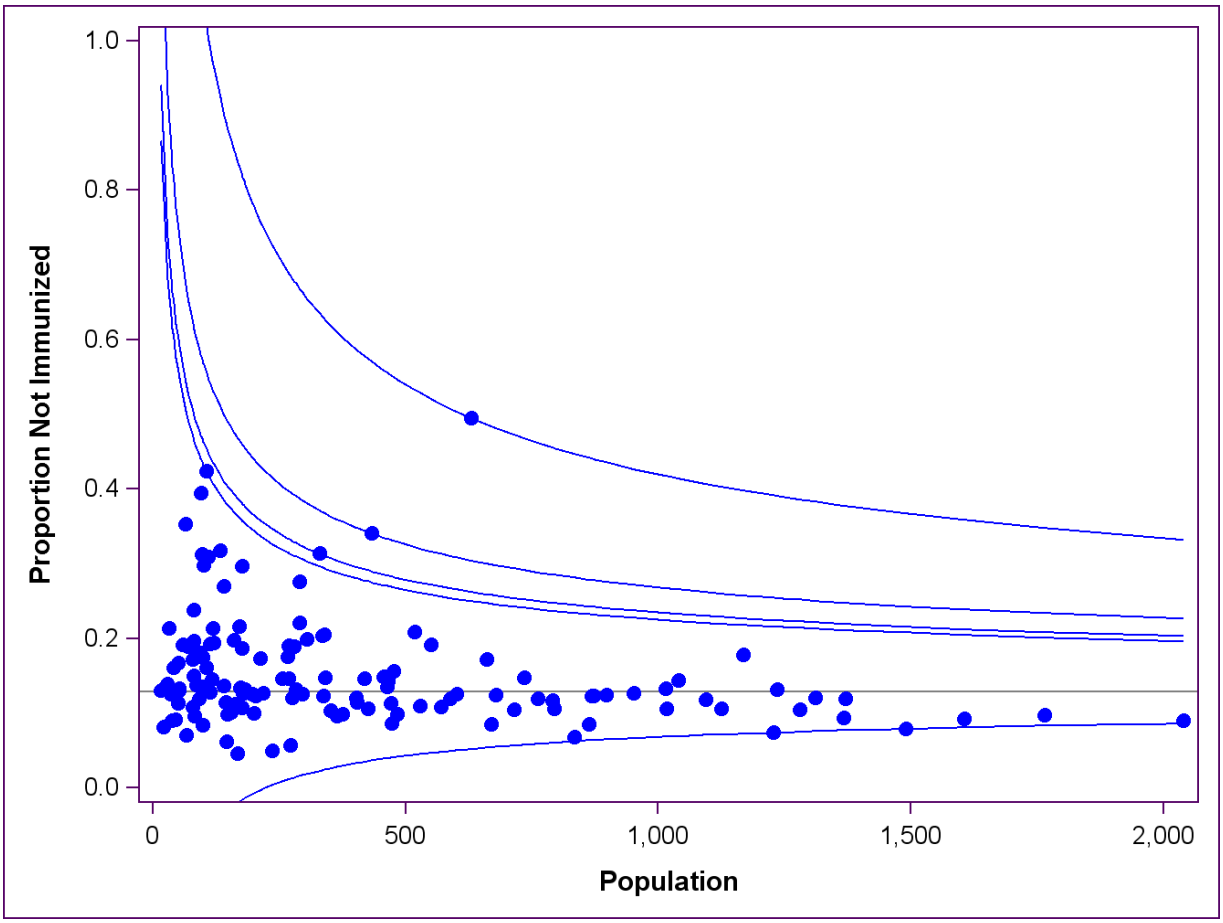
3.3 METHODS

3.3.1 MULTIPLE CLUSTER, MULTIPLE DIRECTION SPATIAL SCAN (MULTSCAN) ALGORITHM

MultScan is a novel cluster detection algorithm motivated by statistical process control concepts. These concepts are briefly reviewed and then the proposed algorithm and testing methods are described.

Statistical process control methods presume that an underlying stable process with some statistical variability generates the observed data. This process is then monitored for deviations. The funnel plot is a statistical process control tool that is commonly used in health system performance monitoring (Spiegelhalter, 2005) and health surveillance (Woodall, 2006; Dover and Schopflocher, 2011) to identify anomalous performance or health events in pre-defined units, such as hospitals or geographic areas, of varying sizes. Funnel plots examine a scatter plot of rates by population, overlaying control limits. These control limits are created under the assumption that the observed rates in all geographies are generated from a common distribution, analogous to Score tests in classical statistics (Poirier, 1995). They represent the range in which observations are expected to fall, for different population sizes, assuming that all observations are generated from the same underlying process. Using binomial proportions as an example, the 95% control limits would be $r_0 \pm \sqrt{\frac{r_0(1-r_0)}{n}}$, where r_0 is an overall proportion (either observed or a target) and n varies over a relevant range of population sizes (Dover and Schopflocher, 2011). Geographies falling outside of the control limits warrant further investigation. This provides a natural ordering of the data – those points lying on the furthest percentile funnel plot limit are of most interest (the top most control limit in Figure 3-1). Viewed another way, standardized scores can be computed as the observed proportion minus the expected proportion divided by a standard error. Score z-scores use standard errors computed assuming the null hypothesis that proportions are generated from the overall proportion and are computed as $z = \frac{r_i - r_0}{\sqrt{\frac{r_0(1-r_0)}{n_i}}}$, for each geography i. Areas with the largest Score z-scores are of most interest.

Figure 3-1. Funnel Plot of 2015 Proportions not immunized with multiple contours

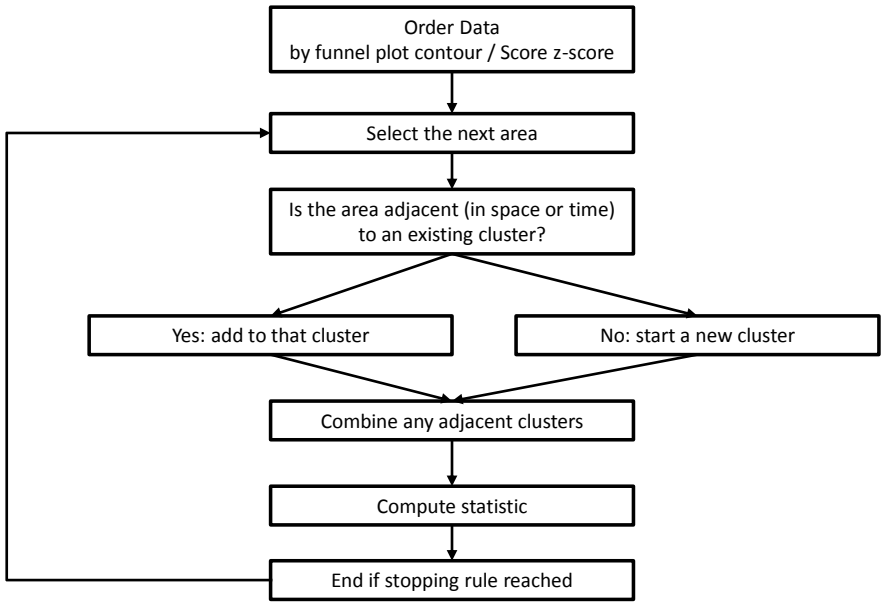


Immunization coverage proportions for measles-mumps-rubella at age 2 for Alberta local geographic areas in 2015 are shown. Control limits for the four largest and one smallest Score z-score values are displayed.

To build clusters, MultScan first sorts the data by Score z-score (refer to Figure 3-1 for a visualization of the z-score ordering as funnel plot control limits). By ordering the data, the MultScan algorithm is computationally efficient, needing to pass through the data only once. Once the data is sorted, clusters are then created and tested. The first geography (i.e. that with the largest Score z-score value) defines the first cluster. The next geography in order is selected; if it is adjacent to a pre-existing cluster, it is added to that cluster and disjoint clusters are combined if they are adjacent; otherwise, it begins a

new cluster. At each step there are one or more simultaneously identified clusters. The number of clusters at each step can grow or shrink, depending on the adjacency of the geographies in that step. The algorithm terminates when a pre-defined stopping rule, such as half the population or half the geographies having been scanned, is met. The algorithm flow is shown in Figure 3-2. The set of geographies identified jointly is considered – for example, if there are three active clusters at a step, the set of three active clusters is the most likely set of clusters. The likelihood ratio for each set of clusters is jointly evaluated by testing the hypothesis that the rates in the active clusters are different from the rate in the remaining geographies. This is a straightforward extension of the original likelihood statistic (Jung, 2009; Li *et al.*, 2011).

Figure 3-2. MultScan Algorithm



Simultaneous multiple cluster detection requires a modification to the testing framework used in spatial scan statistics. Having multiple clusters is equivalent to a model where each cluster has its own rate parameter (Jung, 2009). In the general case, an overall test of significance would scale the likelihood by the degrees of freedom, in this case the number of clusters. This allows direct comparisons of the (scaled) log likelihood between steps with different numbers of clusters (parameters). The standard approach to significance testing, using Monte Carlo simulation since there is no closed form distribution of the likelihood ratio statistics generated under the multiple searches carried out in a spatial scan, in conjunction with the extreme value distribution, is performed on the scaled likelihood ratio statistics.

MultScan generalizes the detection of irregular spatial shapes to irregular spatial-temporal phenomena by defining adjacency in terms of the usual spatial adjacency plus adjacency in time. Adjacency is defined in terms of both space and time, so that geographies are adjacent if (a) they are in the same time period and geographically adjacent, or (b) the same geography is adjacent in time. The MultScan algorithm sorts all geography-time data points by their Score z-scores and proceeds in the same fashion as the spatial-only case. Clusters can thus be simultaneously detected in different spatial locations with different begin and end times and with different patterns of growth.

To allow a direct comparison to methods that detect only single clusters, a variation of MultScan is implemented that, at each step when the set of potential clusters is identified, evaluates the likelihood for each potential cluster individually. The single most likely cluster is then the single cluster across all potential clusters, across all steps, with the largest likelihood ratio. This variant is referred to as MultScan-single.

3.3.2 Implementation Details

MultScan was set to stop searching once half of the geographies had been considered. Clusters were considered statistically significance at the 5% level. Cluster detection, simulation, and analyses were all carried out in SAS version 9. Kulldorf's spatial scan and MultScan were both programmed by the author in SAS.

3.3.3 Simulation Scenarios

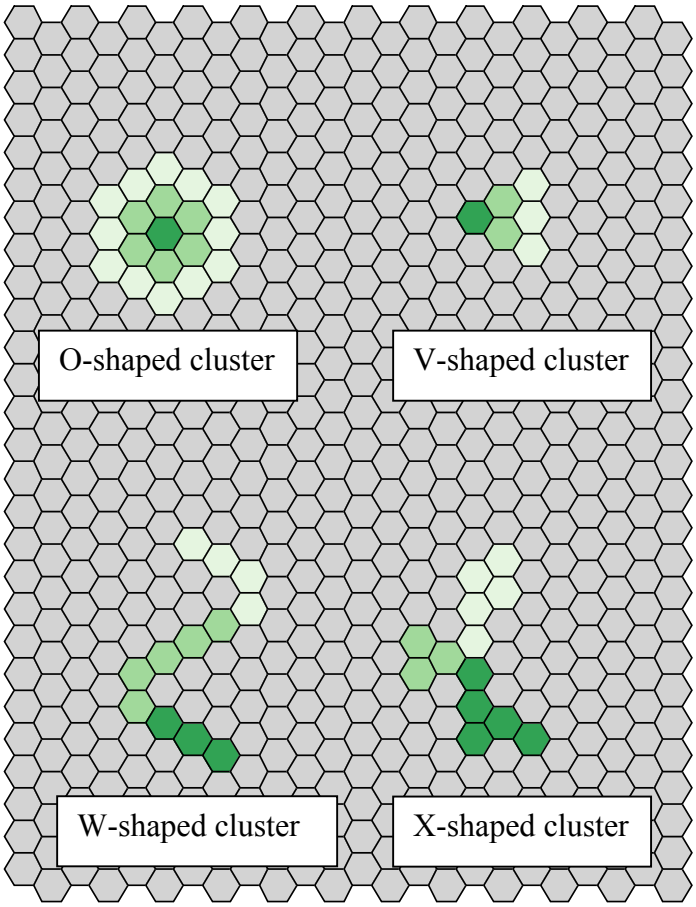
Simulations were performed on a fixed 24 by 27 hexagonal grid containing 648 geographies with varying populations. The population was distributed in geographic areas cyclically taking on one of seven values (1,500 to 10,500 in increments of 1,500) for a total population of 3,879,000. This structure mimics the magnitude of population variation and number of administratively defined geographies likely to be observed in public health surveillance practice in Alberta, Canada. Loosely speaking, this is similar to census tracts in Alberta. Simulations were performed by varying three dimensions: relative risk within the cluster compared to the remainder of the areas, overall disease counts, and spatial-temporal cluster configuration.

In the first dimension, a range of relative risks expected to be seen in health surveillance were examined: 1.5, 2.0, and 3.0.

In the second dimension, annual disease counts of 600, 2,000, 6,000, and 12,000 in the total population were examined. These correspond to a range of epidemiologically plausible rare disease counts to be examined using a spatial scan, from very rare conditions such as multiple sclerosis incidence to relatively more common conditions such as all cancer incidence. Counts were simulated using a Poisson model.

In the third dimension, a variety of cluster configurations were considered. Cluster shapes correspond to one of four epidemiologically plausible shapes. Specifically, four shapes were used that could correspond to the following mechanisms: (1) O: an O-shaped circular, compact cluster – a localized outbreak or point source, (2) V: a compact V-shaped cluster – an aerially dispersed pathogen, (3) W: a W-shaped cluster – an exposure along a geographic feature like a river, and (4) X: a spider-shaped cluster – spread from a central person in a social network. Figure 3-3 shows the four shapes.

Figure 3-3. Spatial-Temporal Cluster Shapes Simulated



The first configuration was spatial only, placing each of the four shapes in space. This was designed to examine the ability to detect different shapes of clusters.

The second configuration was spatial-temporal, allowing cluster to grow through time. For example, the circular cluster begins as only the dark green region (Figure 3-3, top-left). In the next year, it grew to include the next concentric layer, so the cluster includes the dark and medium green areas. In the final year, the cluster included all of the green shaded areas.

The third configuration was spatial only, placing pairs of clusters (O with V, W with X) spatially to examine the ability to detect multiple clusters.

The fourth configuration was spatial-temporal, representing a cluster propagating through time. For example, the first year of a compact cluster begins at a central location (Figure 3-3, dark green), the second year was the area surrounding but not including the first location (medium green only), and the next year was the area surrounding but not including the areas from the previous year (light green only). As these propagations are not adjacent in space or time, they are in fact three distinct space-time clusters.

Four years of data were used for the simulations of temporal clusters, with clusters spanning years two through four. One hundred simulations were carried out for each scenario.

3.3.4 Performance Metrics

Performance of both MultScan and MultScan-single was examined from the perspective of a public health surveillance user by means of a number of metrics. First, usual power and detection power were examined. Usual power is the proportion of simulations in which the algorithm identifies a statistically significant cluster using p-values < .05 cluster and detection power is the proportion of simulations identifying a statistically significant cluster that contains at least one geography from the true cluster. High power represents the ability of the method to detect a cluster when one exists, without considering the accuracy. Positive predictive value and sensitivity were then examined. Positive predictive value is the proportion of the identified geographies that are also part of the true cluster. Low positive predictive value implies that the identified cluster has a substantial amount of noise because it contains many non-cluster geographies, whereas high positive predictive value implies that the identified cluster contains mostly the true cluster (although not necessarily all of the true cluster). Sensitivity is the fraction of the true cluster that is identified in the detected cluster. High sensitivity means that the algorithm is detecting a substantial portion of the true cluster (but possibly also much more). The ability to exactly detect clusters was then examined. Finally, to understand how MultScan utilizes an unrestricted number of clusters, the number of clusters identified was examined.

The performance of MultScan was compared to Kulldorff's spatial scan. Briefly, KSS searches for circular clusters and is the de facto gold standard in public health (Goujon-Bellec *et al.*, 2011). KSS can detect multiple circular clusters by considering non-overlapping secondary clusters that are significant as the same level as the first cluster. When only the original single cluster version of Kulldorff's method is considered, it is referred to as KSS-single.

Performance characteristics were compared between MultScan-single and KSS-single and between MultScan and KSS for the single static cluster and the single growth cluster scenarios. MultScan and KSS are compared for the static multiple cluster and the propagating cluster configurations.

Regression analysis on the simulation results was used to compare between methods and across simulation dimensions. Regressions were run with each of the performance characteristics – usual power, detection power, positive predictive value, sensitivity – as outcomes. For each outcome, separate regression models were run for single clusters, for multiple clusters, and for propagating clusters. This yields a total of 12 models (4 characteristics x 3 cluster types). Each regression model used a combined "information" dimension to capture the joint effect of the relative risk and disease count on performance (see Table 3-1 footnote for formula). Specifically, each regression models included an indicator for the type of statistic (MultScan, MultScan-single, KSS, KSS-single), a continuous measure of information, an indicator of information greater than three to allow for non-linear effects, an indicator of the shape of the cluster(s), an indicator of the growth pattern (static, growth, or propagating). All possible interactions between the information, the cluster shape, and the cluster growth pattern were considered. Results statistically significant at the 5% level and with performance differences of at least 5 percentage points (for power and sensitivity only) are discussed.

Table 3-1. High and Low Information Scenarios

Disease Count	Cluster Relative Risk 1.5	2.0	3.0
600	0.0	1.4	2.8
2,000	1.2	2.6	4.0
6,000	2.3	3.7	5.1
12,000	3.0	4.4	5.8

Dark green corresponds to high information scenarios and light green to low information scenarios. The values in the table are the information values computed as:

Information is defined as $\log\left(count \cdot (RR-1)^2\right) - 6$. *High information is defined as values 3.5 or higher. This heuristic is based upon the Wald statistic which has more power with increasing distance in the estimate, the (RR-1) term, and disease counts for a fixed population size. The product term is then linearized and approximately centred by taking the log and subtracting six.*

3.4 RESULTS

The relative risk and disease count dimensions were combined into a single "information" dimension. A cut-point of 3.5 was selected after a visual examination of the performance characteristics by information. Values below this are referred to as "low information" scenarios and include all scenarios with relative risk 1.5, scenarios with relative risk of 2.0 and disease counts of up to 2,000, and the scenario with relative risk of 3.0 and disease count of 600. Values above the cut-point are referred to as "high information" scenarios and include scenarios with relative risk of 2.0 and disease counts of 6,000 or more and scenarios with relative risk of 3.0 and disease counts of 2,000 or more (see Table 3-1). Linear line segments allowing for one step increase were used in

the modelling. Tables 3-2 and 3-3 summarize the results for usual power, positive predictive value, and sensitivity.

Table 3-2. Simulation Results

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
1	Static Single Cluster	single cluster	O	0.01	0.00	0.01	0.11	0.11	0.13	0.18
				1.21	0.19	0.22	0.35	0.40	0.37	0.45
				1.40	0.21	0.25	0.39	0.45	0.40	0.49
				2.31	0.36	0.41	0.57	0.67	0.58	0.70
				2.60	0.40	0.46	0.63	0.74	0.64	0.76
				2.78	0.43	0.49	0.66	0.78	0.67	0.80
				3.01	0.47	0.53	0.71	0.83	0.72	0.85
				3.70	1.00	1.00	0.81	0.96	0.81	0.93
				3.99	1.00	1.00	0.84	0.97	0.83	0.94
				4.39	1.00	1.00	0.88	0.98	0.86	0.96
				5.09	1.00	1.00	0.95	0.99	0.92	0.98
				5.78	1.00	1.00	1.00	1.00	0.98	1.00
			V	0.01	0.00	0.00	0.02	0.05	0.10	0.14
				1.21	0.18	0.19	0.27	0.28	0.36	0.38
				1.40	0.21	0.22	0.31	0.32	0.40	0.42
				2.31	0.38	0.38	0.50	0.49	0.60	0.60
				2.60	0.43	0.43	0.56	0.54	0.66	0.66
				2.78	0.46	0.46	0.60	0.58	0.70	0.70
				3.01	0.50	0.50	0.65	0.62	0.75	0.74
				3.70	0.96	0.96	0.75	0.70	0.85	0.81
				3.99	0.97	0.96	0.78	0.71	0.87	0.81
				4.39	0.98	0.97	0.83	0.73	0.90	0.81
				5.09	1.00	0.99	0.92	0.75	0.95	0.81
				5.78	1.00	1.00	1.00	0.77	1.00	0.81

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	**Growth Pattern**	**Detection Method**	**Shape**	**Information**						
1	**Static Single Cluster**	**single cluster**	**X**	**0.01**	0.00	0.00	0.14	0.14	0.08	0.15
				1.21	0.22	0.24	0.38	0.29	0.27	0.35
				1.40	0.27	0.29	0.41	0.31	0.30	0.38
				2.31	0.51	0.52	0.60	0.42	0.44	0.53
				2.60	0.58	0.60	0.65	0.46	0.49	0.58
				2.78	0.63	0.64	0.69	0.48	0.51	0.61
				3.01	0.68	0.70	0.73	0.51	0.55	0.64
				3.70	1.00	1.00	0.76	0.48	0.71	0.70
				3.99	1.00	1.00	0.79	0.47	0.75	0.71
				4.39	1.00	1.00	0.84	0.47	0.82	0.73
				5.09	1.00	1.00	0.91	0.46	0.92	0.76
				5.78	1.00	0.99	0.99	0.45	1.00	0.79
			W	**0.01**	0.00	0.00	0.12	0.15	0.04	0.03
				1.21	0.24	0.24	0.34	0.25	0.22	0.23
				1.40	0.28	0.29	0.37	0.27	0.25	0.26
				2.31	0.52	0.52	0.54	0.35	0.38	0.41
				2.60	0.59	0.59	0.59	0.38	0.43	0.46
				2.78	0.64	0.63	0.62	0.39	0.45	0.49
				3.01	0.70	0.69	0.66	0.41	0.49	0.52
				3.70	1.00	1.00	0.74	0.43	0.69	0.57
				3.99	1.00	1.00	0.78	0.43	0.74	0.59
				4.39	1.00	1.00	0.82	0.44	0.80	0.62
				5.09	1.00	1.00	0.89	0.44	0.91	0.67
				5.78	1.00	0.99	0.96	0.44	1.00	0.71

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
1	Static Single Cluster	multiple cluster	O	**0.01**	0.00	0.01	0.26	0.07	0.04	0.18
				1.21	0.13	0.22	0.46	0.31	0.12	0.45
				1.40	0.14	0.25	0.49	0.35	0.13	0.49
				2.31	0.24	0.41	0.65	0.53	0.20	0.70
				2.60	0.27	0.46	0.69	0.59	0.22	0.76
				2.78	0.28	0.49	0.72	0.62	0.23	0.80
				3.01	0.31	0.53	0.76	0.67	0.24	0.85
				3.70	0.91	1.00	0.94	0.75	0.34	0.93
				3.99	0.92	1.00	0.95	0.77	0.42	0.94
				4.39	0.94	1.00	0.96	0.79	0.53	0.96
				5.09	0.96	1.00	0.99	0.84	0.72	0.98
				5.78	0.99	1.00	1.00	0.88	0.91	1.00
			V	**0.01**	0.00	0.00	0.05	0.14	0.00	0.13
				1.21	0.12	0.19	0.35	0.28	0.09	0.38
				1.40	0.14	0.22	0.40	0.30	0.11	0.42
				2.31	0.24	0.38	0.63	0.40	0.21	0.61
				2.60	0.27	0.43	0.70	0.43	0.24	0.67
				2.78	0.29	0.46	0.75	0.45	0.26	0.70
				3.01	0.31	0.50	0.80	0.48	0.28	0.75
				3.70	0.88	0.96	0.93	0.57	0.38	0.81
				3.99	0.90	0.96	0.94	0.58	0.46	0.82
				4.39	0.93	0.97	0.96	0.59	0.57	0.82
				5.09	0.97	0.99	0.99	0.61	0.77	0.84
				5.78	1.00	1.00	1.00	0.62	0.97	0.85

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
1	Static Single Cluster	multiple cluster	X	**0.01**	0.00	0.00	0.14	0.13	0.00	0.16
				1.21	0.16	0.24	0.46	0.25	0.05	0.35
				1.40	0.19	0.29	0.51	0.27	0.06	0.38
				2.31	0.33	0.52	0.76	0.36	0.11	0.53
				2.60	0.38	0.60	0.83	0.39	0.12	0.58
				2.78	0.41	0.64	0.88	0.41	0.13	0.61
				3.01	0.44	0.70	0.94	0.43	0.14	0.65
				3.70	0.97	1.00	0.95	0.40	0.24	0.71
				3.99	0.98	1.00	0.96	0.41	0.35	0.74
				4.39	0.99	1.00	0.96	0.41	0.49	0.77
				5.09	1.00	1.00	0.97	0.42	0.74	0.82
				5.78	1.00	0.99	0.99	0.42	0.98	0.87
			W	**0.01**	0.00	0.00	0.24	0.15	0.01	0.17
				1.21	0.15	0.24	0.50	0.23	0.06	0.32
				1.40	0.18	0.29	0.54	0.25	0.07	0.34
				2.31	0.31	0.52	0.75	0.31	0.10	0.46
				2.60	0.35	0.59	0.81	0.33	0.12	0.50
				2.78	0.38	0.63	0.85	0.34	0.12	0.52
				3.01	0.41	0.69	0.90	0.36	0.13	0.55
				3.70	0.97	1.00	0.96	0.37	0.27	0.64
				3.99	0.98	1.00	0.96	0.37	0.37	0.66
				4.39	0.99	1.00	0.97	0.38	0.51	0.70
				5.09	1.00	1.00	0.98	0.38	0.75	0.76
				5.78	1.00	0.99	0.98	0.39	0.99	0.83

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
1	Single Growth Cluster	single cluster	O	**0.01**	0.00	0.02	0.63	0.46	0.22	0.70
				1.21	0.35	0.46	0.76	0.63	0.37	0.78
				1.40	0.41	0.52	0.78	0.66	0.40	0.80
				2.31	0.70	0.85	0.88	0.78	0.51	0.86
				2.60	0.79	0.95	0.91	0.82	0.55	0.88
				2.78	0.85	1.00	0.93	0.85	0.57	0.89
				3.01	0.92	1.00	0.95	0.88	0.60	0.91
				3.70	0.97	1.00	0.94	0.94	0.88	0.90
				3.99	0.98	1.00	0.95	0.95	0.90	0.89
				4.39	0.98	1.00	0.96	0.96	0.92	0.87
				5.09	0.98	1.00	0.97	0.98	0.96	0.85
				5.78	0.99	1.00	0.98	1.00	1.00	0.82
			V	**0.01**	0.00	0.00	0.00	0.02	0.00	0.04
				1.21	0.17	0.16	0.23	0.23	0.16	0.31
				1.40	0.20	0.18	0.30	0.26	0.21	0.35
				2.31	0.35	0.32	0.63	0.42	0.44	0.56
				2.60	0.40	0.36	0.73	0.46	0.51	0.62
				2.78	0.43	0.39	0.80	0.50	0.55	0.66
				3.01	0.47	0.42	0.88	0.53	0.61	0.71
				3.70	0.95	0.96	0.90	0.63	0.75	0.73
				3.99	0.96	0.97	0.91	0.64	0.78	0.74
				4.39	0.98	0.98	0.92	0.66	0.83	0.74
				5.09	1.00	1.00	0.94	0.69	0.91	0.75
				5.78	1.00	1.00	0.97	0.72	1.00	0.75

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
1	Single Growth Cluster	single cluster	X	**0.01**	0.00	0.00	0.33	0.26	0.00	0.22
				1.21	0.30	0.36	0.58	0.36	0.08	0.38
				1.40	0.36	0.42	0.61	0.37	0.10	0.40
				2.31	0.66	0.75	0.79	0.45	0.21	0.53
				2.60	0.75	0.85	0.85	0.47	0.24	0.57
				2.78	0.81	0.92	0.89	0.48	0.26	0.59
				3.01	0.88	1.00	0.93	0.50	0.29	0.62
				3.70	1.00	1.00	0.91	0.49	0.57	0.63
				3.99	1.00	1.00	0.92	0.49	0.64	0.63
				4.39	1.00	1.00	0.93	0.49	0.73	0.64
				5.09	0.99	0.99	0.95	0.49	0.88	0.65
				5.78	0.99	0.99	0.98	0.50	1.00	0.66
			W	**0.01**	0.00	0.00	0.26	0.20	0.03	0.26
				1.21	0.29	0.35	0.52	0.29	0.14	0.39
				1.40	0.34	0.41	0.56	0.30	0.16	0.41
				2.31	0.63	0.72	0.75	0.37	0.24	0.51
				2.60	0.71	0.82	0.81	0.40	0.27	0.54
				2.78	0.77	0.88	0.85	0.41	0.28	0.56
				3.01	0.84	0.96	0.90	0.43	0.30	0.58
				3.70	1.00	1.00	0.89	0.43	0.55	0.60
				3.99	1.00	1.00	0.91	0.44	0.61	0.59
				4.39	1.00	1.00	0.93	0.46	0.70	0.59
				5.09	0.99	0.99	0.96	0.49	0.86	0.57
				5.78	0.99	0.99	0.99	0.51	1.00	0.56

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
1	Static Single Cluster	multiple cluster	O	**0.01**	0.03	0.02	0.00	0.28	0.00	0.70
				1.21	0.22	0.46	0.36	0.36	0.07	0.79
				1.40	0.25	0.52	0.41	0.37	0.08	0.81
				2.31	0.40	0.85	0.71	0.43	0.16	0.87
				2.60	0.44	0.95	0.80	0.45	0.19	0.90
				2.78	0.47	1.00	0.86	0.46	0.21	0.91
				3.01	0.51	1.00	0.93	0.47	0.23	0.93
				3.70	0.93	1.00	0.94	0.43	0.53	0.97
				3.99	0.94	1.00	0.95	0.45	0.59	0.97
				4.39	0.96	1.00	0.96	0.47	0.68	0.98
				5.09	0.99	1.00	0.98	0.51	0.84	0.98
				5.78	1.00	1.00	1.00	0.55	1.00	0.98
			V	**0.01**	0.00	0.00	0.00	0.04	0.00	0.04
				1.21	0.09	0.16	0.19	0.14	0.04	0.31
				1.40	0.11	0.18	0.25	0.15	0.06	0.35
				2.31	0.19	0.32	0.53	0.22	0.13	0.56
				2.60	0.21	0.36	0.62	0.24	0.15	0.62
				2.78	0.23	0.39	0.68	0.26	0.17	0.66
				3.01	0.24	0.42	0.75	0.27	0.19	0.71
				3.70	0.89	0.96	0.93	0.31	0.33	0.73
				3.99	0.91	0.97	0.94	0.31	0.41	0.75
				4.39	0.94	0.98	0.95	0.32	0.52	0.79
				5.09	0.98	1.00	0.98	0.33	0.72	0.84
				5.78	1.00	1.00	1.00	0.34	0.92	0.90

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
1	Static Single Cluster	multiple cluster	X	**0.01**	0.00	0.00	0.25	0.18	0.00	0.19
				1.21	0.20	0.36	0.51	0.26	0.03	0.37
				1.40	0.23	0.42	0.55	0.28	0.03	0.40
				2.31	0.40	0.75	0.75	0.34	0.06	0.54
				2.60	0.46	0.85	0.81	0.36	0.07	0.58
				2.78	0.49	0.92	0.85	0.38	0.08	0.61
				3.01	0.53	1.00	0.89	0.39	0.09	0.65
				3.70	0.97	1.00	0.95	0.36	0.20	0.71
				3.99	0.98	1.00	0.96	0.36	0.31	0.73
				4.39	0.99	1.00	0.97	0.37	0.46	0.76
				5.09	1.00	0.99	0.98	0.37	0.72	0.81
				5.78	1.00	0.99	1.00	0.38	0.98	0.86
			W	**0.01**	0.00	0.00	0.06	0.18	0.00	0.37
				1.21	0.17	0.35	0.40	0.22	0.03	0.48
				1.40	0.20	0.41	0.45	0.23	0.04	0.50
				2.31	0.34	0.72	0.71	0.26	0.08	0.59
				2.60	0.39	0.82	0.79	0.27	0.09	0.62
				2.78	0.42	0.88	0.85	0.28	0.10	0.63
				3.01	0.45	0.96	0.91	0.29	0.10	0.65
				3.70	0.96	1.00	0.94	0.31	0.26	0.70
				3.99	0.97	1.00	0.95	0.32	0.36	0.72
				4.39	0.98	1.00	0.96	0.33	0.50	0.75
				5.09	1.00	0.99	0.99	0.35	0.75	0.79
				5.78	1.00	0.99	1.00	0.37	0.99	0.84

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
2	Static Two Clusters	multiple cluster	O V	0.01	0.00	0.00	0.25	0.26	0.00	0.06
				1.21	0.15	0.26	0.53	0.42	0.04	0.27
				1.40	0.18	0.31	0.57	0.44	0.05	0.30
				2.31	0.33	0.54	0.78	0.56	0.11	0.45
				2.60	0.37	0.61	0.84	0.60	0.13	0.50
				2.78	0.40	0.65	0.89	0.62	0.14	0.53
				3.01	0.44	0.71	0.94	0.65	0.16	0.57
				3.70	0.99	1.00	0.95	0.73	0.29	0.84
				3.99	0.99	1.00	0.96	0.73	0.37	0.85
				4.39	0.99	1.00	0.97	0.73	0.50	0.87
				5.09	0.99	1.00	0.99	0.74	0.71	0.91
				5.78	0.99	1.00	1.00	0.74	0.92	0.94
			W X	0.01	0.00	0.00	0.25	0.24	0.04	0.06
				1.21	0.21	0.32	0.55	0.31	0.06	0.22
				1.40	0.25	0.38	0.59	0.32	0.06	0.24
				2.31	0.44	0.64	0.82	0.38	0.07	0.36
				2.60	0.50	0.73	0.89	0.40	0.07	0.40
				2.78	0.54	0.78	0.94	0.41	0.07	0.42
				3.01	0.58	0.85	0.99	0.42	0.08	0.45
				3.70	0.99	1.00	0.98	0.40	0.22	0.68
				3.99	0.99	1.00	0.98	0.41	0.32	0.72
				4.39	0.99	1.00	0.98	0.42	0.48	0.77
				5.09	0.99	1.00	0.98	0.44	0.74	0.85
				5.78	0.99	1.00	0.98	0.46	1.00	0.93

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
2	Two Growth Clusters	multiple cluster	O V	**0.01**	0.00	0.01	0.27	0.19	0.18	0.58
				1.21	0.21	0.44	0.54	0.34	0.15	0.66
				1.40	0.25	0.51	0.59	0.37	0.14	0.67
				2.31	0.45	0.84	0.80	0.49	0.12	0.72
				2.60	0.51	0.95	0.86	0.52	0.11	0.74
				2.78	0.55	1.00	0.90	0.55	0.11	0.75
				3.01	0.60	1.00	0.95	0.58	0.10	0.76
				3.70	0.97	1.00	0.94	0.51	0.38	0.89
				3.99	0.97	1.00	0.95	0.55	0.44	0.87
				4.39	0.97	1.00	0.97	0.60	0.54	0.85
				5.09	0.97	1.00	0.99	0.69	0.70	0.80
				5.78	0.97	1.00	1.00	0.77	0.86	0.76
			W X	**0.01**	0.00	0.00	0.21	0.21	0.00	0.17
				1.21	0.25	0.41	0.51	0.28	0.00	0.28
				1.40	0.29	0.48	0.56	0.29	0.00	0.29
				2.31	0.50	0.82	0.78	0.35	0.03	0.38
				2.60	0.56	0.93	0.86	0.37	0.05	0.40
				2.78	0.60	0.99	0.90	0.38	0.06	0.42
				3.01	0.66	1.00	0.96	0.39	0.07	0.44
				3.70	1.00	1.00	0.96	0.37	0.24	0.62
				3.99	1.00	1.00	0.96	0.38	0.34	0.65
				4.39	1.00	1.00	0.97	0.39	0.49	0.68
				5.09	1.00	1.00	0.98	0.42	0.75	0.73
				5.78	1.00	1.00	0.99	0.44	1.00	0.78

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
3	Skip Cluster	multiple cluster	O	**0.01**	0.00	0.00	0.00	0.00	0.00	0.43
				1.21	0.17	0.34	0.31	0.10	0.07	0.61
				1.40	0.19	0.40	0.37	0.12	0.08	0.64
				2.31	0.34	0.71	0.66	0.21	0.15	0.79
				2.60	0.38	0.81	0.75	0.24	0.17	0.83
				2.78	0.41	0.87	0.81	0.26	0.18	0.86
				3.01	0.44	0.95	0.88	0.28	0.20	0.89
				3.70	0.97	0.99	0.99	0.31	0.39	0.95
				3.99	0.98	0.99	0.99	0.32	0.47	0.96
				4.39	0.98	0.99	1.00	0.32	0.58	0.96
				5.09	1.00	0.99	1.00	0.33	0.78	0.97
				5.78	1.00	0.99	1.00	0.34	0.97	0.98
			V	**0.01**	0.07	0.04	0.00	0.00	0.00	0.00
				1.21	0.13	0.11	0.26	0.00	0.06	0.16
				1.40	0.13	0.12	0.31	0.01	0.07	0.19
				2.31	0.18	0.18	0.60	0.10	0.16	0.34
				2.60	0.19	0.20	0.69	0.13	0.19	0.39
				2.78	0.20	0.21	0.75	0.15	0.21	0.42
				3.01	0.21	0.22	0.82	0.17	0.23	0.46
				3.70	0.79	0.82	0.93	0.20	0.35	0.64
				3.99	0.82	0.85	0.94	0.20	0.42	0.67
				4.39	0.88	0.90	0.94	0.21	0.51	0.73
				5.09	0.97	0.97	0.96	0.21	0.68	0.82
				5.78	1.00	1.00	0.98	0.22	0.84	0.91

Table 3-2. – continued –

					Modelled Performance Metrics					
					Usual Power		Positive Predictive Value		Sensitivity	
					MultScan	KSS	MultScan	KSS	MultScan	KSS
Number of Clusters	Growth Pattern	Detection Method	Shape	Information						
3	Skip Cluster	multiple cluster	X	**0.01**	0.01	0.00	0.00	0.08	0.00	0.02
				1.21	0.13	0.15	0.28	0.20	0.02	0.15
				1.40	0.14	0.17	0.33	0.22	0.03	0.17
				2.31	0.23	0.28	0.62	0.31	0.07	0.26
				2.60	0.26	0.32	0.71	0.34	0.08	0.29
				2.78	0.28	0.34	0.77	0.36	0.08	0.31
				3.01	0.30	0.37	0.84	0.38	0.09	0.33
				3.70	0.91	0.96	0.95	0.41	0.11	0.47
				3.99	0.93	0.97	0.96	0.42	0.21	0.52
				4.39	0.95	0.98	0.97	0.42	0.36	0.58
				5.09	0.98	1.00	0.98	0.43	0.62	0.69
				5.78	1.00	1.00	1.00	0.44	0.87	0.80
			W	**0.01**	0.02	0.00	0.00	0.00	0.00	0.17
				1.21	0.13	0.14	0.28	0.12	0.05	0.26
				1.40	0.14	0.16	0.34	0.14	0.06	0.27
				2.31	0.22	0.28	0.63	0.23	0.10	0.34
				2.60	0.25	0.32	0.72	0.26	0.11	0.36
				2.78	0.27	0.34	0.78	0.28	0.12	0.37
				3.01	0.29	0.37	0.85	0.30	0.13	0.39
				3.70	0.91	0.95	0.96	0.33	0.14	0.44
				3.99	0.93	0.96	0.96	0.33	0.24	0.48
				4.39	0.95	0.97	0.97	0.34	0.38	0.54
				5.09	0.99	0.99	0.99	0.35	0.61	0.64
				5.78	1.00	1.00	1.00	0.35	0.84	0.74

Table 3-3. Simulation Results Summary

Cluster Configuration		Method Type	Relative Performance*		
			Usual Power	**Positive Predictive Value**	**Sensitivity**
Single Cluster (static or growth)	Circular	Single cluster	KSS superior only in low information	KSS superior only in mid-information static circular cluster	MultScan superior in high information; KSS superior in low information
	Other shapes		Similar	**MultScan** superior	MultScan superior in high information; KSS superior in low information
	Circular	Multiple cluster	KSS superior only in low information	**MultScan** superior	**KSS** superior for static circular cluster; KSS superior in low and middle information growth cluster
	Other shapes		KSS superior only in low information	**MultScan** superior	MultScan superior at highest information; KSS superior in low and middle information

*bolding of the method name indicates scenarios where that method is broadly superior

Table 3-3. – continued --

Cluster Configuration		Method Type	Relative Performance*		
			Usual Power	**Positive Predictive Value**	**Sensitivity**
Two Clusters (static or growth)		Multiple cluster	KSS superior only in low information	**MultScan** superior	MultScan superior at highest information; KSS superior in low and middle information
Propagating Clusters (three clusters)		Multiple cluster	Similar	**MultScan** superior	MultScan superior in highest information irregular clusters; **KSS** superior in remaining scenarios

*bolding of the method name indicates scenarios where that method is broadly superior

3.4.1 Usual Power and Detection Power

For configurations of single spatial clusters, MultScan-single and KSS-single performed comparably in both static and growth clusters in the high information range. For example, Figure 3-4(a) shows usual power for single circular cluster simulations. The solid line represents the modelled performance of MultScan-single and the dashed line the modelled performance of KSS-single. Both methods showed the expected increasing power to detect clusters with increasing information – higher cluster relative risk and/or higher disease counts. MultScan-single and KSS-single performed similarly with 100% power in the high information scenarios. MultScan methods displayed lower power than KSS methods in the low information scenarios. This pattern held for the single static cluster configurations and the single growth cluster configurations (for both single and multiple cluster methods), as well as the two cluster and propagating cluster configurations.

The spatial-temporal shapes of the clusters appeared to affect power, although it was difficult to discern a consistent pattern. Irregular shapes (W, X) had higher power than compact shapes for static clusters but spatial-temporal clusters with growth showed higher power for KSS-single. MultScan-single has similar power for most growth shapes.

The results for detection power are virtually identical to the usual power results (not shown).

In summary, the methods showed comparable power in all high information scenarios. In low information scenarios, KSS often displayed higher power.

Figure 3-4(a). Usual power for a circular static cluster

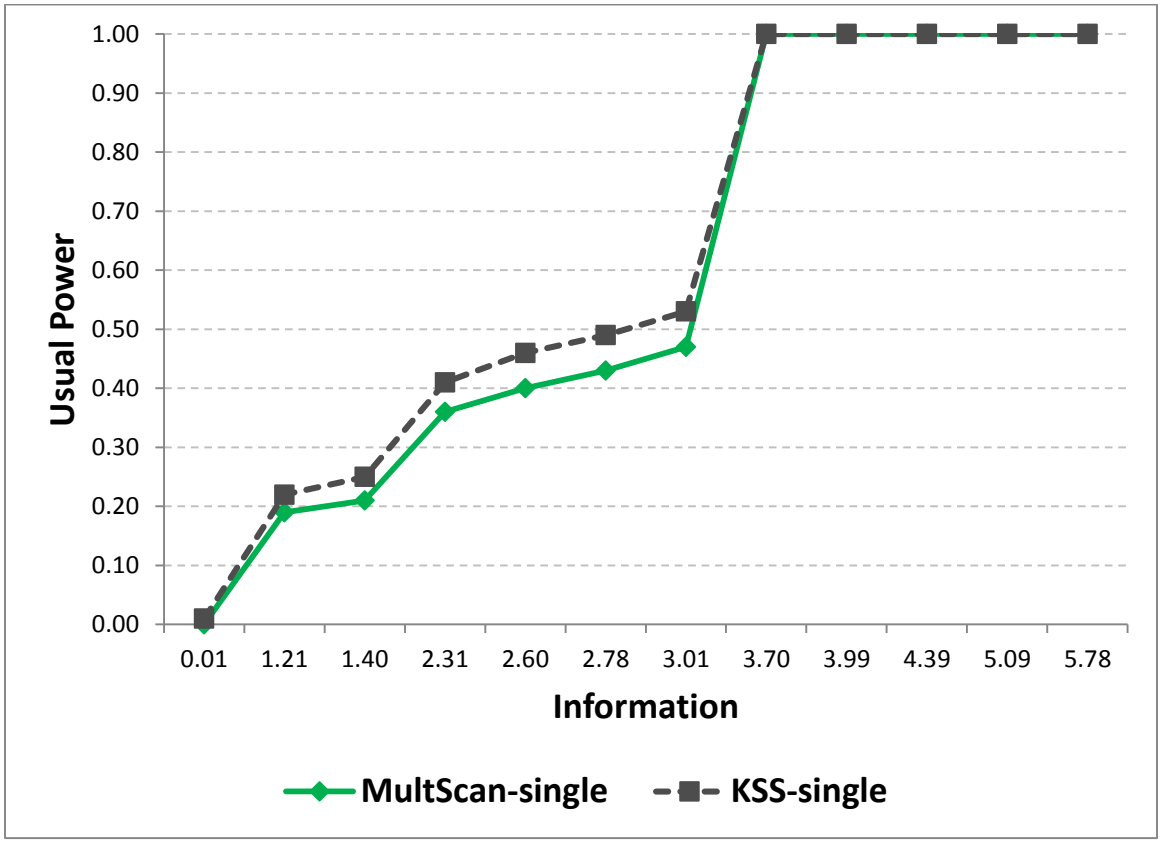


Figure 3-4(b). PPV for MultScan-single static clusters

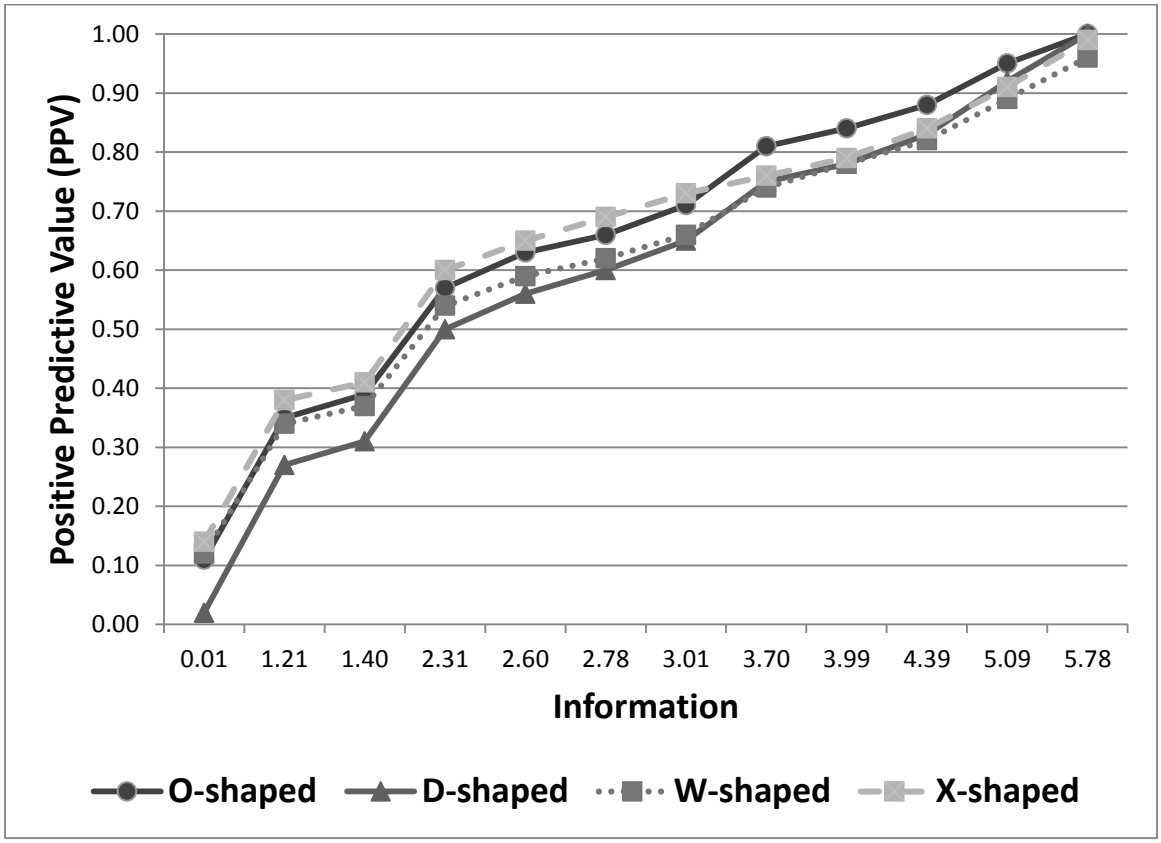


Figure 3-4(c). Sensitivity for an X-shaped static cluster

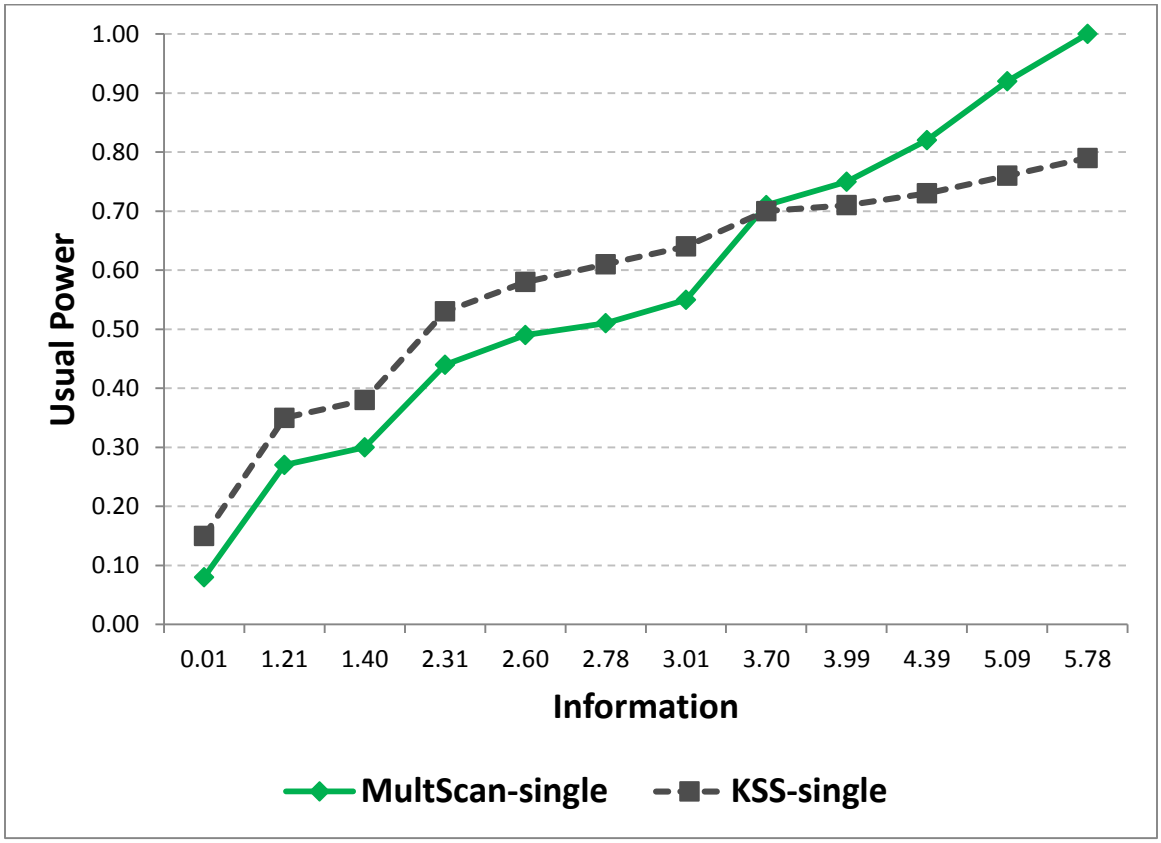


Figure 3-4(d). Number of clusters detected by MultScan

3.4.2 Positive Predictive Value

Shape did not affect the ability of MultScan to accurately detect any type or number of clusters. Figure 3-4(b) shows as an example the positive predictive value curves for MultScan-single for each single static cluster shape. MultScan displayed good positive predictive value in low information scenarios and very high positive predictive value in all high information scenarios. A similar pattern is seen with the growth clusters, pairs of clusters, and propagating clusters.

For KSS, shape significantly impacted positive predictive value. KSS had highest positive predictive value for circular shapes, then declining for V-shapes and then irregular shapes. This was an expected result as KSS uses only circles (or cylinders) when scanning.

Across virtually every scenario, MultScan had much higher positive predictive value than KSS. The only exception was in the middle information scenarios with the static circular cluster where KSS able to outperform MultScan.

3.4.3 SENSITIVITY

Across all spatial-temporal shapes and for both MultScan and KSS, compact clusters had higher sensitivity than irregularly shaped clusters. Higher sensitivity was seen in the compact O and V clusters compared to the irregular X and W clusters. MultScan and KSS displayed a similar ordering across shapes.

With single static clusters and growth clusters, a cross-over in the relative performance of MultScan and KSS was observed. In the low information scenarios, KSS had higher sensitivity. This flipped and MultScan had higher sensitivity by the high information scenarios. Figure 3-4(c) illustrates this effect for the X-shaped spatial cluster.

A similar cross-over effect was seen in the two cluster and propagating scenarios. MultScan overtakes or matches KSS in terms of sensitivity again, but only in the highest information scenarios.

3.4.4 EXACT DETECTION

The paucity of exact detection lead to not modelling this outcome. For a static circular cluster, the only type that KSS could detect exactly, KSS-single exactly detected an average of 30.4% of low and 83.0% of high information simulations; KSS had corresponding exact detection of 22.4% and 64.5%. MultScan methods never detected more than 1% of clusters exactly in low information scenarios. In a static circular cluster, MultScan and MultScan-single had exact detection of 11.6% and 34.4%. For non-circular static clusters, MultScan and MultScan-single had exact detection of 12.1% and 23.6%. This declined to 2.8% and 10.4% respectively for growth clusters. MultScan exactly detected 4.6% of multiple clusters, <1% of multiple growth clusters, 4.3% of propagating clusters.

3.4.5 MULTSCAN BEHAVIOUR

The performance of the MultScan method exhibited an interesting feature in terms of the number of clusters detected. In the low information range, MultScan usually detected only one cluster. The number of detected clusters then increased beyond the number of true clusters, and finally converged back down toward the true number of clusters [Figure 3-4(d)]. This behaviour occurred regardless of the true number of clusters.

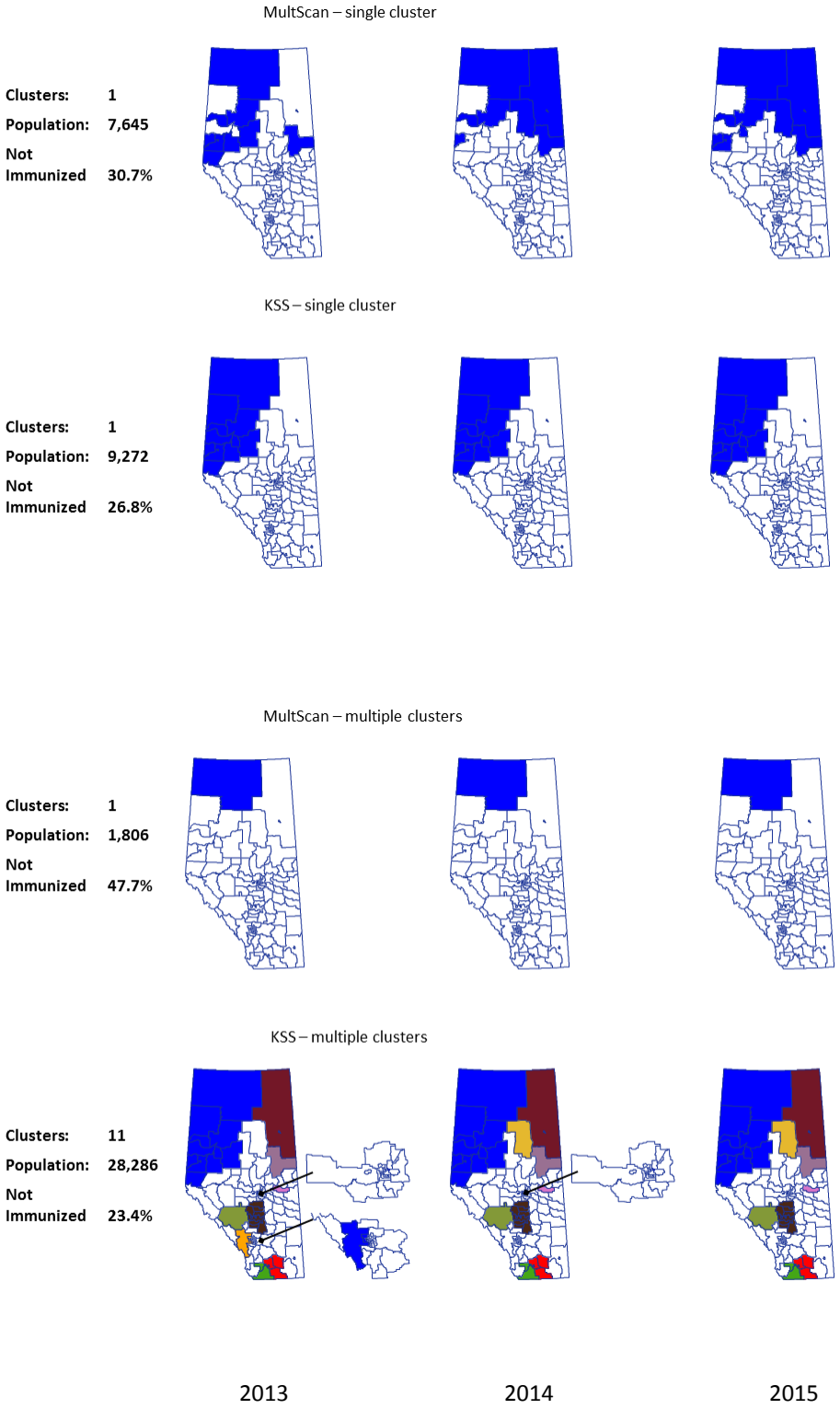
3.4.6 REAL DATA – CHILDHOOD IMMUNIZATION

To test MultScan and MultScan-single with real data, childhood immunization coverage proportions across small geographies in Alberta were examined. Alberta is a province in Canada with a population of approximately 4.2 million and 56,000 births in 2015 (Alberta Ministry of Health, 2017a). Coverage for measles-mumps-rubella at age 2 for 132 local planning areas (geographies designed for health planning and surveillance) were publicly available from the Alberta Ministry of Health's online tool (Alberta Ministry of Health, 2017b). Data were from 2013 to 2015.

Spatial-temporal scans were carried out looking for areas where the unimmunized proportions were highest. MultScan, when scanning for multiple clusters, identified a single cluster covering a population of 1,806 children with an unimmunized proportion of 47.7% (see Figure 3-5). The cluster represents a single rural geography. In contrast, KSS identified 10 distinct clusters covering a total population of two year olds of 28,286 with an unimmunized proportion of 23.4%. The clusters are geographically dispersed around the province in rural north, central and south Alberta. Most clusters spanned all three years.

MultScan-single identified a cluster covering a population of two year olds of 7,645 with an unimmunized proportion of 30.7% in the northern portion of the province. The cluster beginning in 2013 has two disconnected components, joining geographically in 2014 and expanding slightly in 2015. KSS-single identified a larger cluster in the north–west portion of the province. It had a population of 9,272 with 26.8% unimmunized spanning all three years.

Figure 3-5. Clusters of Proportions Not Immunizing for Measles, Alberta, 2013 to 2015



3.5 DISCUSSION

As recent outbreaks in previously rare diseases have shown, contemporary health outbreaks can quickly spread across multiple geographic areas. To date, analytic efforts to identify the location, shape, and changes in time in these clusters have been limited, which in turn has prevented public health officials from acting with the best information. This study demonstrates that MultScan can simultaneously identify multiple clusters that change and grow irregularly through time, which means it best mimics the changes that naturally occur as clusters shift, expand, and recede.

MultScan utilizes insights from statistical process control theory in its approach by ordering the data by Score z-scores (funnel plot contours). This data driven approach likely contributes to the high levels of positive predictive value observed. Other spatial scan methods using rankings have met with mixed results. Using a rate based ranking algorithm and search, Patil and Taillie performed similarly to other likelihood based methods that attempt a modified exhaustive search (Patil and Taillie, 2004). Using an algorithm similar to MultScan but relying on Wald z-scores, Que and Tsui showed increased positive predictive values but at the cost of much lower sensitivity (Que and Tsui, 2011). Both methods maximize the likelihood, but do so relying on ordered data without attempting to approximate an exhaustive search. MultScan appears to compensate for the loss of sensitivity by virtue of the Score z-scores. Score z-scores create a more stable ordering of areas by virtue of the assumption that the variance estimate comes from the entire population whereas the variance estimate in Wald z-scores depends on the more variable observed rates in the small area data. This reinforces the suggestion (Tango and Takahashi, 2005; Assunção *et al.*, 2006) that the unchecked maximization of the likelihood is the underlying cause of poor positive predictive value in many spatial scan methods. While the Score z-score ordering is one of the greatest strengths of the MultScan algorithm, it also represents a limitation. The data based ordering makes it easy to exclude geographically or temporally adjacent areas because

they are ranked further down the list. This likely explains the only adequate levels of sensitivity observed.

Regardless of the shape or number of clusters in space or time, MultScan displayed high levels of positive predictive value, virtually always outperforming KSS. As indicated by the positive predictive values of 80% to 100% in high information scenarios, MultScan identified clusters with very little noise. The high positive predictive values observed are likely driven by the interesting way that MultScan arrives at the true number of clusters. The simulations found that the algorithm initially overestimates the number of clusters by ignoring connecting area with only slightly elevated rates; as it collects more information though, these bridge pieces are added into the cluster, and the algorithm detects the exact number of clusters. This maintains the utility of the method in public health practice. In situations when limited information is available and there are multiple clusters in close (but not adjacent) geographic proximity, health professionals will be able to combine the MultScan results with epidemiologic and local knowledge to decide how to deploy resources in the identified clusters and in the "bridging" areas between them.

MultScan maintained adequate levels of sensitivity in the simulations. In general, MultScan had lower sensitivity in low information scenarios, but its sensitivity rivalled and then exceeded KSS as information increased. MultScan displayed levels of usual power and detection power comparable to KSS in all high information scenarios. Across the low information scenarios, KSS displayed greater power. These low information scenarios, comprised of small relative risks and particularly low disease rates, plague all cluster detection methods, leading many studies to focus on higher information scenarios (Goujon-Bellec *et al.*, 2011). While this may be understandable from a statistical perspective, it is important for public health practitioners to note that no method can be relied upon in low information scenarios to accurately detect spatial-temporal clusters.

The types of simulations represent a limitation of this study. The spatial boundaries are regular, equally sized, polygons and the population structure follows a fixed pattern. Results on regular data may not generalize to real world conditions where, in Alberta for example, there are very large rural geographic areas with small populations and large urban populations within small geographic boundaries. The simulations have captured

variability in population size, as well as a variety of rare health event rates and a set of plausible cluster shapes.

Exactly identifying clusters provides public health professionals with quality information to help plan and execute public health interventions. To date, this has been one of the greatest limitations of spatial scan statistics; a recent review of five scan statistics determined that "none of the methods was ever able to detect the underlying cluster *exactly*, irrespective of the population size, shape or relative risk of the cluster" (Goujon-Bellec *et al.*, 2011). In contrast to these methods, MultScan and MultScan-single showed the ability to exactly detect clusters in high information scenarios. The ability of MultScan to exactly identify two changing clusters is a substantial improvement over existing methods, although the low exact detection rates are an indication that additional research is needed in this area.

The case study consistently had MultScan based methods identifying clusters with higher proportions not immunizing than did the corresponding KSS methods. This finding was consistent with the simulation results demonstrating that MultScan had higher positive predictive value than KSS. This suggests that the clusters identified by MultScan are more likely to contain only the true cluster while the clusters identified by KSS are more likely to include areas not actually a part of the true cluster. Public health interventions focused on increasing immunization uptake are costly endeavours. Targeting them more widely than necessary (e.g. by relying on KSS results) would waste limited public health resources. This implies that focusing on areas identified by MultScan would be the most effective use of resources.

3.6 CONCLUSION

As the spatial-temporal shapes and number of clusters will never be known in advance in real applications, it is vital to be able to detect multiple, irregularly shaped clusters in space and time. The high positive predictive power observed in MultScan allows effective allocations of public health resources, although this comes with lower sensitivity in some cases. The ability to exactly identify irregularly shaped clusters in

both space and time is a characteristic not observed in other methods. MultScan represents a useful tool in public health surveillance and applications requiring the identification of spatial-temporal clusters.

3.7 Appendix – Input Data Files

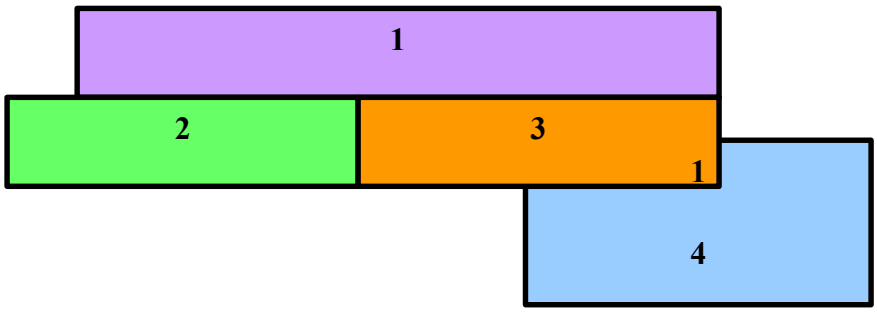
MultScan requires two input datasets, an adjacency matrix and the count data.

The adjacency matrix is used to identify which geographic areas are adjacent for the purposes of grouping areas into geographic clusters. The dataset contains two columns of geographic identifiers. The pair of geographies identifies a pair of geographic areas that are adjacent to each other. An example of the data used to create the adjacency matrix is illustrated in Figure 3-6. In this example, depicted in Figure 3-7, geography 1 is adjacent to geographies 2 and 3, but not 4; geography 2 is adjacent to geographies 1 and 3, but not 4; geography 3 is adjacent to geographies 1, 2, and 3; geography 4 is adjacent to only geography 3. Adjacency is defined by sharing a portion of their boundaries.

Figure 3-6. MultScan Input Dataset: Adjacency

Geographic Identifier 1	Geographic Identifier 2
1	2
1	3
2	1
2	3
3	1
3	2
3	4
4	3
...	...

Figure 3-7. Adjacency Example



The count dataset used in MultScan includes a geographic identifier, an optional year variable, the count of cases, and the population at risk. The annual population rate and

subsequent annual Score z-scores are created from this data. An example of the data used is illustrated in Figure 3-8.

Figure 3-8. MultScan Input Dataset: Count Data

Geographic Identifier	Year*	Count	Population
1	2003	98	1,043
1	2004	117	1,218
1	2005	132	1,287
2	2003	1,217	11,957
2	2004	1,246	12,850
2	2005	1,329	13,244
3	2003	620	5,421
3	2004	618	5,344
3	2005	661	5,498
4	2003	314	3,300
4	2004	351	3,428
4	2005	401	3,910
…	…	…	…

Kulldorff's spatial scan uses the same count dataset as MultScan for input, but rather than using an adjacency matrix relies on distances from centroids of the geographic areas. Centroids are given in latitude and longitude. An example of the centroid dataset is shown in Figure 3-9.

Figure 3-9. KSS Input Dataset: Geographic Centroids

Geographic Identifier	Latitude	Longitude
1	54.63	-114.61
2	53.46	-114.87
3	53.45	-113.48
4	52.23	-112.02
…	…	…

3.8 Appendix: The Statistical Basis for Spatial Scan Statistics

Spatial scans determine if a cluster exists, where it is spatially located, and determine its statistical significance. This is accomplished in a four step process:

1. Apply a search algorithm
2. Compute a test statistic for each set of geographies in the algorithm
3. Note the most likely cluster, the set of geographies with the largest value of the test statistic
4. Evaluate the statistical significance of the most likely cluster

I begin by reviewing these steps for Kulldorff's circular spatial scan (Kulldorff and Nagarwalla, 1995), and then show how these are modified for the MultScan algorithm.

3.8.1 Search Algorithm

The search algorithm for Kulldorff's circular spatial scan is an exhaustive search of all "circular" clusters of geographies. There is, in the spatial case, a set of n geographies. The algorithm starts at the centroid if each geographic region, hence starting with a single geography to evaluate the test statistic for. Then, ever larger circles are created until another centroid is included, and the test statistic is computed for these geographies. This continues until a logical limit is reached, such as half of the population being included in the potential cluster.

It should be noted, that although the algorithm proceeds using circles and centroids of geographic regions, this does not ensure that the resulting cluster identified is circular or even connected. Since the underlying geographic regions are not circular, an agglomeration of them likely will not be as well. In cases with very non-compact geographic regions, regions with long spidery shapes for example, it is possible to include two geographies that have centroids close together, but are separated by a third geographic region that happens to have a thin arm separating the original two geographies.

The MultScan search algorithms I propose are single pass searches that pre-order the geographic regions and then sequentially build potential clusters. The pre-ordering is accomplished by sorting the geographic regions by their Score z-scores. In sequence, each region is then added to an existing cluster if it is adjacent to it, or the region begins a new cluster. In the case where the geography connects to two or more potential clusters, the now connected clusters merge to become a single cluster. The MultScan algorithm stops here at each step, with a set of potential clusters (i.e. there may be more than one cluster identified at each step in the algorithm). In the case of the MultScan single algorithm (designed to detect only a single cluster), there is an additional sub-search where each of the identified potential clusters is examined.

3.8.2 TEST STATISTICS

Both Kulldorff's circular scan and MultScan single use similar likelihood ratio statistics as the basis for evaluating potential clusters. The derivation of the likelihood used in MultScan in the Poisson case follows.

Let k $= 1..K$ represent K clusters.

Let □ ent the set of geogrphies that comprise cluster k

Let $i = 1..N$ represent N geographies.

Let α_0 represent the overall rate (for all geographies).

In the case of clusters, let α represent the underlying rate in the non-cluster geographies and let RRα represent the rate in the cluster geographies with relative risk RR.

Recall that the Poisson distribution for independent observations with rate α_0, populations p_i, and counts c_i is given by

c_i □ $(\alpha_0 \cdot p_i) = \dfrac{(\alpha_0 \cdot p_i)^{c_i} e^{-\alpha_0 \cdot p_i}}{c_i!}$

The Poisson likelihood (L_{H_0}) for N observations under H_0, the null hypothesis of a common rate α_0 (no clustering), is given by

$$L_{H_0} = \prod_{i=1}^{N} \frac{(\alpha_0 \cdot p_i)^{c_i} e^{-\alpha_0 \cdot p_i}}{c_i!}$$

Under H_1, the hypothesis of K clusters, the geographies within the clusters share a rate of RRα and the geographies outside the K clusters share a rate α. The likelihood reflects this, with the first term being the likelihood for geographies in the clusters and the second term being the likelihood for the non-cluster geographies.

$$L_{H_1} = \prod_{k=1}^{K} \prod_{i \in \square} \frac{(RR\alpha \cdot p_i)^{c_i} e^{-RR\alpha \cdot p_i}}{c_i!} \bullet \prod_{\square} \frac{(\alpha \cdot p_i)^{c_i} e^{-\alpha \cdot p_i}}{c_i!}$$

The likelihood ratio (LR) is given by the ratio of the likelihood under H_1 to the likelihood under H_0 and is given by

$$LR = \frac{L_{H_1}}{L_{H_0}}$$

$$= \frac{\prod_{k=1}^{K} \prod_{i \in \square} \frac{(RR\alpha \cdot p_i)^{c_i} e^{-RR\alpha \cdot p_i}}{c_i!} \bullet \prod_{\square} \frac{(\alpha \cdot p_i)^{c_i} e^{-\alpha \cdot p_i}}{c_i!}}{\prod_{i=1}^{N} \frac{(\alpha_0 \cdot p_i)^{c_i} e^{-\alpha_0 \cdot p_i}}{c_i!}}$$

Simplifying and taking the natural log, the loglikelihood ratio (LLR) is given by

$$LLR = \sum_{k=1}^{K} \sum_{i \in \square} \left(c_i \log(RR\alpha \cdot p_i) - RR\alpha \cdot p_i\right) + \sum_{\square} \left(c_i \log(\alpha \cdot p_i) - \alpha \cdot p_i\right) - \sum_{i=1}^{N} \left(c_i \log(\alpha_0 \cdot p_i) - \alpha_0 \cdot p_i\right)$$

The MultScan likelihood derived above is a generalization of Kulldorff's formulation (where only one cluster is searched for, i.e. K=1). The MultScan single algorithm that searches for only a single cluster also has K=1.

The test statistic used by Kulldorff is simply the log likelihood ratio. For MultScan, a change is required to the test statistic to account for the fact that there are (potentially) multiple clusters being simultaneously evaluated. To do this, I use the fact that minus two times the log likelihood ratios is asymptotically has a chi-squared distribution, and scale the log likelihood ratio by the degrees of freedom. In this case, the degrees of freedom is the number of clusters. So at each step, the test statistic is LLR/K , where K is the number of clusters.

3.8.3 Evaluating Statistical Significance of the Most Likely Cluster(s)

The most likely cluster has the largest value of the test statistic. The next step is to determine the statistical significance of the most likely cluster. This is a challenging task, because the area being tested is under examination specifically because it had the largest value of the test statistic, out of a huge number of test statistics being generated. This is a multiple testing issue, similar to "data dredging" in epidemiology. There is no closed form for a general spatial scan with varying populations (Glaz, Pozdnyakov and Wallenstein, 2009). To deal with these issues, Kulldorff proposed Monte Carlo testing. This process involves repeatedly creating a new simulated set of outcomes through simulation under the assumption of no clusters and recording the maximum value of the test statistic. This creates an empirical distribution for the test statistic and the test statistic from the observed data has a significance level equivalent to the percentile the observed test statistics fits in. For example, if the observed test statistic is 36, and this is the 3rd largest value out of a combined 999 simulations plus the observed, the p-value would be 3/1,000 or p=.003.

The Monte Carlo approach to significance testing is computationally intensive. Each simulation is costly in terms of its runtime, and the number of simulations must be large

to obtain a reliable distribution of the test statistic. Implementations of Kulldorff's spatial scan usually default to 999 simulations.

A more computationally feasible approach has been proposed by Abrams et al (Abrams, Kleinman and Kulldorff, 2010). This approach uses a parametric distribution, specifically the Gumbel distribution. This distribution is an intuitively sensible choice as the Gumbel distribution is a distribution of maxima. Other distributions have been considered (Jung, 2009; Abrams, Kleinman and Kulldorff, 2010), but the Gumbel outperforms all others for all forms of spatial scan statistics examined to date. The Gumbel distribution is a two parameter distribution, and can be parameterized as a function of the sample mean and standard deviation. Thus Monte Carlo simulations under the null are still required, but only enough to estimate a mean and standard deviation.

3.8.4 Extending the Spatial Scan to include Covariates

The same arguments that suggest adjustment for covariates in funnel plots apply to adjustment in spatial scanning. Fundamentally, it is a matter of interpretation. In spatial scans, the interest is in the identification of geographic regions that are unusual, and require further investigation. To be interesting and unusual implies that the cause of the unusualness is unknown. In other words, that known factors that affect the outcome of interest are already accounted for.

There are two approaches to accounting for covariates. The first naturally extends the distributional assumption in the spatial scans to include covariates. The second approach disentangles the modelling step from the spatial scan step. Jung (Jung, 2009) extended the Poisson circular spatial scan to Poisson with covariates, and then generalized the process to the entire class of generalized linear models. Any familiar regression model – logistic, Poisson, normal, exponential – can directly be seen to be able to incorporate the cluster concept by including as a covariate an indicator for membership in the potential cluster. In fact, Kulldorff's Poisson and binomial spatial scan statistics are special cases of Jung's framework. In this way, the spatial scan procedure can be seen to compute a regression based test statistic at each step of the algorithm. In principle, this process could

be extended to any regression framework (not only generalized linear models) including models such as the Cox proportional hazards model used in survival analysis. The second approach distinguishes between the modelling step and the cluster evaluation step. It is the same approach to risk adjustment examined in Chapter 2 on funnel plots and is frequently used in spatial scan statistics (Kulldorff *et al.*, 1997). The modelling is carried out first, and an adjusted outcome is computed. This adjusted outcome is then the input to the spatial scan algorithm. Both the regression method and the two step method for including covariates can be used.

3.9 REFERENCES

Abrams, A. M., Kleinman, K. and Kulldorff, M. (2010) 'Gumbel based p-value approximations for spatial scan statistics', *International Journal of Health Geographics*, 9(1), p. 61.

Alberta Ministry of Health (2017a) *Interactive Health Data Application (IHDA) - Demographics - Population Estimates - Adjusted*. Available at http://www.ahw.gov.ab.ca/IHDA_Retrieval/.

Alberta Ministry of Health (2017b) *Interactive Health Data Application (IHDA) - Immunization - Childhood Coverage Rates*. Available at http://www.ahw.gov.ab.ca/IHDA_Retrieval/.

Assunção, R. *et al.* (2006) 'Fast detection of arbitrarily shaped disease clusters.', *Statistics in medicine*, 25, pp. 723–742.

Center for Disease Control (CDC) (1990) 'Guidelines for investigating clusters of health events.', *MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports TA* -, 39(RR-11), pp. 1–23.

Dover, D. C. and Schopflocher, D. P. (2011) 'Using funnel plots in public health surveillance', *Population Health Metrics*, 9(1), p. 58.

Duczmal, L. *et al.* (2007) 'A genetic algorithm for irregularly shaped spatial scan statistics', *Computational Statistics and Data Analysis*, 52, pp. 43–52.

Duczmal, L. and Assuncao, R. (2004) 'A simulated annealing strategy for the detection of arbitrarily shaped spatial clusters', *Computational Statistics and Data Analysis*, 45, pp. 269–286.

Duczmal, L., Kulldorff, M. and Huang, L. (2006) 'Evaluation of Spatial Scan Statistics for Irregularly Shaped Clusters', *Journal of Computational and Graphical Statistics*, pp. 428–442.

Glaz, J., Pozdnyakov, V. and Wallenstein, S. (2009) *Scan statistics: methods and applications*. 1st edn. Birkhäuser Basel.

Goujon-Bellec, S. *et al.* (2011) 'Detection of clusters of a rare disease over a large territory: performance of cluster detection methods', *International Journal of Health Geographics*, p. 53.

Government of Alberta (2011) *Guidelines for the Investigation of Clusters of Non-Communicable Health Events*. Edmonton, AB. Available at http://www.health.alberta.ca/documents/Investigation-Clusters-Guidelines-2011.pdf.

Groseclose, S. L. and Buckeridge, D. L. (2017) 'Public Health Surveillance Systems: Recent Advances in Their Use and Evaluation', *Annual Review of Public Health*. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, California 94303-0139, USA, 38(0), pp. 57–79.

Iyengar, V. S. (2005) 'Space-time clusters with flexible shapes', *MMWR Suppl*, 54, pp. 71–76.

Jung, I. (2009) 'A generalized linear models approach to spatial scan statistics for covariate adjustment', *Statistics in Medicine*. John Wiley & Sons, Ltd., 28(7), pp. 1131–1143.

Kulldorff, M. (1997) 'A spatial scan statistic', *Communications in Statistics - Theory and Methods*, 26(6), pp. 1481–1496.

Kulldorff, M. *et al.* (1997) 'Breast cancer clusters in the northeast United States: a geographic analysis', *American journal of epidemiology*. Oxford Univ Press, 146(2), pp. 161–170.

Kulldorff, M. *et al.* (2006) 'An elliptic spatial scan statistic.', *Statistics in medicine*, 25, pp. 3929–3943.

Kulldorff, M. and Nagarwalla, N. (1995) 'Spatial disease clusters: detection and inference.', *Statistics in medicine*, 14, pp. 799–810.

Li, X. Z. *et al.* (2011) 'A spatial scan statistic for multiple clusters', *Mathematical Biosciences*, 233(2), pp. 135–142.

Lieu, T. *et al.* (2015) 'Geographic clusters in underimmunization and vaccine refusal.', *Pediatrics TA* -, 135(2), pp. 280–289.

Neill, D. B. and Moore, A. W. (2004) 'Rapid detection of significant spatial clusters', *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*, p. 256.

Patil, G. and Taillie, C. (2004) 'Upper level set scan statistic for detecting arbitrarily shaped hotspots', *Environmental and Ecological statistics*, 11, pp. 189–197.

Poirier, D. J. (1995) *Intermediate Statistics and Econometrics: A Comparative Approach*. The MIT Press.

Que, J. and Tsui, F.-C. (2011) 'Rank-based spatial clustering: an algorithm for rapid outbreak detection.', *Journal of the American Medical Informatics Association : JAMIA*, 18, pp. 218–224.

Spiegelhalter, D. J. (2005) 'Funnel plots for comparing institutional performance', *Statistics in Medicine*, 24(8), pp. 1185–1202.

Tango, T. (2008) 'A spatial scan statistic with a restricted likelihood ratio', *Japanese Journal of Biometrics*, 29(2), pp. 75–95.

Tango, T. and Takahashi, K. (2005) 'A flexibly shaped spatial scan statistic for detecting clusters', *International Journal of Health Geographics*, 4(1), p. 11.

Tango, T. and Takahashi, K. (2012) 'A flexible spatial scan statistic with a restricted likelihood ratio for detecting disease clusters.', *Statistics in medicine*, pp. 4207–18.

Wakefield, J. and Kim, A. (2013) 'A Bayesian model for cluster detection.', *Biostatistics (Oxford, England)*, 14, pp. 752–65.

Waller, L. A. and Gotway, C. A. (2004) *Applied Spatial Statistics for Public Health Data*, *Environmental Health*.

Woodall, D. H. (2006) ‘The Use of Control Charts in Health-Care and Public-Health Surveillance’, *J Qual Technol*, 38(2), pp. 89–104.

Yiannakoulias, N., Rosychuk, R. J. and Hodgson, J. (2007) ‘Adaptations for finding irregularly shaped disease clusters.’, *International journal of health geographics*, 6, p. 28.

Zhang, Z., Assuncao, R. and Kulldorff, M. (2010) ‘Spatial scan statistics adjusted for multiple clusters’, *Journal of Probability and Statistics*, 2010, pp. 1–11.

CHAPTER 4: CHILDHOOD IMMUNIZATION IN ALBERTA

An edited version of the following paper will be submitted as:

> Use of funnel plots and spatial scans to guide immunization surveillance and intervention
>
> Douglas C. Dover
>
> Shannon MacDonald

The paper integrates the funnel plot visualizations and adjustment methodology with the MultScan spatial scan statistic. These methods are applied to childhood immunization in Alberta accounting for a wide variety of factors related to immunization uptake.

4.1 Abstract

4.1.1 Background

Monitoring of childhood immunization coverage is a key tool for public health to control vaccine preventable diseases. Understanding which determinants of immunization are influential in which geographic areas helps in targeting public health interventions.

4.1.2 Data

The routine childhood immunization schedule for DTaP-IPV-Hib is 2, 4, 6, and 18 months. Coverage for these doses in an Alberta cohort born in 2010 is measured at age two. Determinants of immunization measured include families with 3 or more children, Vitamin K uptake, recently moved, having a regular family physician, being a single parent, income quintile, and immunization provider. All data were publicly available at the local planning area geographies in Alberta, Canada from a Government of Alberta Open Data website.

4.1.3 Methods

Immunization coverage in small geographic areas was visualized using funnel plots. Adjustment was carried out using log-binomial regressions accounting for coverage in the previous dose. Overdispersion was measured using Pearson's goodness of fit statistic. Spatial scans were used to identify any geographic clustering of under- or over-immunization.

4.1.4 Results

Substantial overdispersion in coverage was observed in all doses before adjustment. After adjustment, the first dose still showed moderate overdispersion. Remarkably, the second and third doses showed underdispersion. The fourth dose showed moderate overdispersion. After adjustment, there was no geographic clustering of immunization coverage in any dose.

4.1.5 CONCLUSION

Childhood immunization coverage, accounting for determinants of immunization, was analyzed using funnel plots, regression, and spatial scans. In Alberta, the determinants of immunization measured in this study accounted for a substantial portion of the observed overdispersion. No geographic clustering was observed after accounting for the determinants. This approach to analyzing immunization coverage is cost-effective, sustainable, and provides a rich description for public health professionals.

4.2 INTRODUCTION

One of the top ten achievements of public health in the 20th century, immunization has been extremely successful in reducing the burden of communicable disease morbidity and mortality (Center for Disease Control (CDC), 2011). Vaccine preventable diseases with severe clinical outcomes are now rare due to immunization being the norm. Despite, or perhaps due to, this success, public health is facing a new challenge in the area of immunization – the unexpected decline in vaccine uptake by parents for routine childhood immunizations. "Vaccine hesitancy refers to delay in acceptance or refusal of vaccines despite availability of vaccine services. Vaccine hesitancy is complex and context specific, varying across time, place and vaccines. It is influenced by factors such as complacency, convenience and confidence." (World Health Organization (WHO), 2014). Public health officials are responsible for protecting the health of the population, this mandate clearly extending to protecting the population from vaccine preventable diseases. To effectively do this, public health officials must monitor the variation in and determinants of immunization uptake to determine what interventions are required where and in what sub-populations. Immunization surveillance systems can, when well designed, provide this information.

For immunization surveillance systems to be widely used by public health officials, they need to provide accurate information while still being adaptable to new situations. This can be achieved by incorporating five factors into surveillance system design (Lee *et al.*, 2010). The first is that the surveillance system needs to have a strong theoretical framework, which ultimately ensures accurate conclusions and effective public health actions. The second is the ability to accurately detect clusters of immunization behavior or changes in behavior over time; this sensitivity is critical for responding to changes and directing resources where they are most needed. Third, the surveillance system needs to be timely in both its data collection and analysis to allow interventions to occur when they are still possible and most likely to be effective. Fourth, flexibility in the system to incorporate new or local knowledge about immunization behavior ensures that the system remains current. Finally, the system needs to be simple to implement. Since surveillance activities are resource-limited, any system needs to use methods and processes that are

sound and cost effective. Surveillance systems that combine these five factors will be used and trusted by public health officials, which in turn creates trust in the communities where they are acting.

To guide public health action, the surveillance system must explain immunization coverage in the population across pre-existing geographic administrative areas using knowledge of factors influencing immunization behavior. Not only must the system be able to identify areas where the population is at unusually high risk, it must also simultaneously identify those most amenable to intervention. While examples of complex analyses of immunization coverage exist in the research literature (Omer *et al.*, 2008; Lieu *et al.*, 2015), a robust methodology meeting the population-based, ongoing, systematic surveillance needs of public health officials has yet to be developed.

We propose an immunization surveillance system based upon geographically aggregated data and methods. We examine the system's ability to explain geographic variation in immunization coverage at age two, to identify areas of unusually high or low coverage, and to inform public health actions.

4.3 METHODS

4.3.1 THEORETICAL FRAMEWORK

We used the Behavioural Model of Health Services Use (Aday and Andersen, 1974) to characterize factors affecting immunization uptake into Predisposing Characteristics, Enabling Resources, and Health Care System Characteristics. Predisposing Characteristics that have been identified in the literature include religious beliefs (Mills *et al.*, 2005; van der Wal *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016), birth order (Pearce *et al.*, 2008; MacDonald *et al.*, 2014; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016), and immunization-specific knowledge/beliefs including anticipated pain from the immunization event (Mills *et al.*, 2005; MacDonald *et al.*, 2014; Larson *et al.*, 2014) and the protective effects of the vaccine (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; MacDonald *et al.*, 2014; Larson *et al.*, 2014;

Thomson, Robinson and Vallée-Tourangeau, 2016), direct benefits via protection from disease (Larson *et al.*, 2014), indirect benefits from helping others through herd immunity (Mills *et al.*, 2005; MacDonald *et al.*, 2014; Larson *et al.*, 2014), and potential adverse reactions (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; MacDonald *et al.*, 2014; Larson *et al.*, 2014; Thomson, Robinson and Vallée-Tourangeau, 2016). Enabling Resources such as having a regular medical provider (Mills *et al.*, 2005; Falagas and Zarkadoulia, 2008; Dubé *et al.*, 2013; MacDonald *et al*., 2014; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016), socio-economic resources (Pearce *et al.*, 2008; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016), and being a two-parent family (Pearce *et al.*, 2008; Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016) have all been proposed as factors that support childhood immunization. Health Care System Characteristics related to immunization include having recently moved (Falagas and Zarkadoulia, 2008; MacDonald *et al.*, 2014), which may represent an ability to integrate into the local health care system, and characteristics of the immunization provider (Larson *et al.*, 2014; de Cantuária Tauil, Sato and Waldman, 2016). Many of these factors can vary in prevalence geographically and are thus expected to explain some portion of the geographic variation in coverage.

4.3.2 Setting and Population

Alberta, Canada is a province with a population of 4 million (Alberta Ministry of Health, 2017a) with a universal, publicly funded health care system. Childhood immunization programmes are a part of this system and are delivered through public health clinics throughout Alberta. We will examine DTaP-IPV-Hib immunization protecting against diphtheria, tetanus, pertussis, polio and *Haemophilus influenzae* type B. The routine childhood schedule in Alberta is four doses administered at 2, 4, 6, and 18 months of age (also referred to as doses 1 to 4) with coverage at age two being routinely reported.

We examined immunization coverage at age two for the 2010 birth cohort (50,630 births, (Alberta Ministry of Health, 2017a) at a small geographic level. The local geographic areas used in this study were the 132 geographies used by the province for the planning, delivery, and evaluation of health and health care delivery.

4.3.3 Data Sources and Variables

Data were obtained from the Alberta Ministry of Health Interactive Health Data Application (IHDA) (Alberta Ministry of Health, 2017a). Immunization coverage estimates in the IHDA were based on data from Imm/ARI (Alberta's provincial Immunization and Adverse Reactions following Immunization repository). Imm/ARI is a population-based repository and is considered complete for publicly funded childhood immunizations administered since 2006. Population based coverage at age 2 in 2012 for each of four doses of DTaP-IPV-Hib were sequentially analyzed. Provincially, the coverage estimates were 92.1, 90.3, 87.9 and 74.6% for 2, 4, 6 and 18 month doses respectively. Coverage estimates are based upon time-to-event methodology (Alberta Ministry of Health, 2017b) using a birth cohort and estimating the probability of being immunized at age 2 (Alberta Ministry of Health, 2017b).

Data on the factors related to immunization coverage were based on aggregate, publicly available measures from the 2012 Canadian Community Health Survey (CCHS), the 2011 National Household Survey (NHS), and 2006-2012 Alberta Notice of Birth forms. The Canadian Community Health Survey is a national health survey designed to measure health and health determinants at the national, provincial and sub-provincial level (Béland, 2002). The National Household Survey (NHS) is the optional survey attached to the 2011 Canadian census (Hamel and Laniel, 2014). The Alberta Notice of Birth data contains information collected on all births in the province (Sahni, Lai and MacDonald, 2014). All data was publicly available on the Alberta Ministry of Health's IHDA with accompanying documentation.

The predisposing factor of the percentage of families with three or more children was available from the NHS. Refusal of Vitamin K at birth to prevent vitamin K deficiency bleeding has been related to vaccine hesitancy (Sahni, Lai and MacDonald, 2014), presumably representing similar beliefs to immunization. We used the proportion of live births accepting Vitamin K as a proxy measure of vaccine confidence.

The enabling resource of having a regular medical provider was available from CCHS. Socio-economic resources measured by income quintile and single parent families were both available from the NHS.

Health care system integration was captured from the NHS as the percentage of families having moved at least once within the last five years. Characteristics specific to each immunization provider are captured using an indicator of the former 2003 Regional Health Authority, the structure under which immunization programmes were delivered at the time in Alberta.

4.3.4 Analytic Methods

The proportion of children immunized for doses at 2, 4, 6, and 18 months in each of the 132 local areas was examined visually using funnel plots (as described in (Dover and Schopflocher, 2011). Funnel plots are scatter plots of the population size by immunization proportion with control limits overlaid. The control limits, similar to confidence intervals, capture the expected variation in observations assuming the only difference between areas is due to population sizes. Points beyond the 95% control limits are worthy of further investigation and points beyond the 99.8% control limits are considered alerts (Dover and Schopflocher, 2011). In the case of immunization proportions, binomial control limits were used. Analytically, the Pearson goodness-of-fit statistics was used to quantify any amount of overdispersion. Overdispersion represents an excess of variability in the observed data over what would be expected from the statistical model used. A value of 1 suggests no overdispersion; values less than 1 suggest under-dispersion; and values greater than 1 suggest overdispersion. In public health, the most likely cause of this overdispersion is unmeasured covariates (Birkmeyer, 2001; Dover and Schopflocher, 2011). Funnel plots were examined for all four doses of DTaP-IPV-Hib for both the crude immunization proportions as well as the proportions adjusted for the factors affecting immunization.

Adjustment was carried out with ecological log-binomial regressions (Wacholder, 1986) on coverage proportions for each dose, sequentially adding the Health Care System

Characteristics, Enabling Resources, and Predisposing Characteristics to identify overdispersion at each step. The models for doses 2 through 4 included as a covariate the previous doses' coverage proportions, replacing the usual intercept term. In this way, the interpretation of the model coefficients is the ***change*** in effect between subsequent doses. This approach has the advantage of removing the effect of any common unmeasured covariates in the analyses of subsequent doses. For example, if there is an unmeasured covariate that has the same effect on coverage for each dose, the model for dose 2 would have accounted for it through the inclusion of the dose 1 observed proportion in the model. If the effect of an unmeasured covariate were to increase for dose 2 (vis-à-vis dose 1), this would be observed as increased overdispersion in the dose 2 adjusted model. To aid in the interpretation of the relative risks for continuous factors, the relative risks are given for a percentage point range depending on the observed range of each factor. Model fit was assessed by examining influential observations and the linearity of effects assumption by testing quadratic terms for all continuous factors.

Geographic clusters of over or under-immunization were evaluated using MultScan (Chapter 3), a spatial scan statistic capable of simultaneously detecting multiple irregularly shaped clusters in small area data. MultScan identifies clusters by passing through the data (which is ordered by funnel plot contour values) and creating clusters out of adjacent geographic areas. The binomial likelihood was computed for each set of clusters and the set of clusters with the maximal value of the likelihood was identified as the most likely set of clusters. A Monte Carlo method utilizing the extreme value distribution was then used to obtain p-values for the set of clusters, adjusted for the number of clusters identified. Spatial scans were carried out for unadjusted and fully adjusted coverage proportions for each dose.

Data management and analyses were carried out in SAS 9.3. Statistical significance was evaluated at the $p < 0.05$ level. Ethical approval was obtained from the Health Research Ethics Board at the University of Alberta.

4.4 RESULTS

Table 4-1 shows the distribution of the measured factors for the 132 local areas, as well as the distribution of dose 1 through 4 coverage proportions.

The results of the log-binomial regression are given in Table 4-2. No observations had undue influence in any of the models. All factors were found to be significant in the dose 1 regression model. Dose 2 coverage was 97.3% (96.6-98.1) of dose 1 coverage (i.e. it was reduced by 2.7%), dose 3 coverage was similar at 98.2% (97.4-99.0) of dose 2, and dose 4 coverage dropped to 90.2% (89.0-91.4) of dose 3 coverage.

Beginning with Health Care System Characteristics, the indicator of the delivery system (former regional health authorities) shows significant heterogeneity in effects between former regional health authorities, the relative risk between the region with the highest dose 1 rate and the region with the lowest rate being 1.075 (95% confidence interval 1.056, 1.095) in dose 1; these remained unchanged (the factor was not significant) until dose 4 when all authorities had varying degrees of decline in immunization coverage and the spread between the highest and lowest increased to a ratio of 1.206 (95% CI 1.176, 1.237). In the proportion of families recently moving, the relative risk across the 60 percentage point range was 1.036 (95% CI 1.018-1.087) and there was no change in impact on immunization coverage after the first dose. Turning to Enabling Characteristics, the relative risk for the 20 percentage point range in the proportion of individuals with regular family doctors was 1.075 (95% CI 1.058-1.094) in dose 1, an additional 1.025 (95% CI 1.005-1.045) increase at the second dose and not changing thereafter. The proportion of single parent families displayed a non-linear response in the first dose, so a quadratic term was added to the first dose model. Relative risks were less than one for single parent family percentages less than 20% or greater than 35%. After the first dose, each subsequent dose had an additional (linear) effect with relative risks across a 30 percentage point range of 0.979 (95% CI 0.959-0.999), 0.992 (95% CI 0.970-1.000), and 0.997 (95% CI 0.996-0.998) for doses 2 through 4, respectively. The dose 1 model captured a U-shaped effect of income quintile, with the richest (IQ5) and poorest (IQ1) quintiles showing high coverage. The largest effect was between IQ1 and IQ4 with risk of 1.020 (95% CI 1.011-1.029); there were no changes in these effects after dose 1. Finally, when examining Predisposing Characteristics, we found that the percentage of

families with three or more children had an incremental effect over every subsequent dose, lowering the chance of immunization for each subsequent dose with relative risks of 0.844 (95% CI 0.813-0.876), 0.944(95% CI 0.906-0.9987), 0.930 (95% CI 0.887-0.976), and 0.858 (95% CI 0.798-0.922) for doses 1 to 4, respectively, across a 50 percentage point range. The relative risk of a percentage point increase in Vitamin K uptake was 1.072 (95% CI 1.057-1.087) in dose 1. It showed no change in effect until dose 4, where it was associated with an additional relative risk of 1.027 (95% CI 1.003-1.052).

Funnel plots of the unadjusted coverage proportions and the adjusted coverage proportions are shown in Figure 4-1. The large number of points very far outside the unadjusted funnels compared to their corresponding adjusted funnel plots graphically illustrates the substantial explanatory power of the regression model. The unadjusted dose 1 model had an overdispersion measure of 12.50, showing approximately 12 times the expected variation between areas (Table 4-3). After accounting for the Health Care System factors, the overdispersion measure reduced to 7.12 and after accounting for Enabling factors and Predisposing Characteristics reduced further to 4.50. The model for dose 2 including only the dose 1 coverage proportion showed no evidence of overdispersion, with a value of 1.08. Adding of all factors reduced the dose 2 value to 0.71, indicating a slight amount of under-dispersion; the same pattern is observed in dose 3 where the overdispersion measure started at 1.01 and ended at 0.54. Dose 4 exhibited a different pattern of overdispersion with a value of 7.47 after accounting for only the dose 3 coverage proportion. Adding Health Care System factors reduced the measure to 2.96 and adding the remaining factors reduced it to 2.09.

As the Health Care System Characteristics are added to the model, substantial reductions in overdispersion can be seen, particularly in the dose 1 and dose 4 models where the overdispersion is reduced by almost half. The Enabling Resources further reduced the overdispersion by approximately 20% in doses 2 through 4. The Predisposing Characteristics further reduced overdispersion by nearly 20% in dose 1 and 10% in doses 2 through 4. The dose 1 model has substantial overdispersion present even after

adjustment. Remarkably, after adjustment the dose 2 and 3 models show under dispersion. The Dose 4 model shows slight overdispersion after adjustment.

The spatial scan of the unadjusted dose 1 proportions did not find any significant clusters, although the most likely cluster was a single local area in the northern part of the province. Unadjusted dose 2 and 3 spatial scans each identified a single significant cluster, the same northern local area seen in dose 1. The dose 4 result was uninformative, identifying a huge connected region covering nearly half of the province. The spatial scan results for the adjusted proportions were markedly different: no significant clusters were identified for any dose.

4.5 Discussion

Aggregate models of immunization coverage can account for substantial portions of geographic variability in coverage proportions as illustrated with Alberta data. Our models confirm the impact of Predisposing Characteristics, Enabling Resources and Health Care System Characteristics on coverage across all four childhood doses of DTaP-IPV-Hib in Alberta.

The results in this ecologic model generally reflect what is expected from individual level models in the literature. For Predisposing Characteristics, Vitamin K acceptance was associated with increased vaccine coverage (Sahni, Lai and MacDonald, 2014); a higher proportion of families with three or more children was associated with decreased coverage (MacDonald *et al.*, 2014). For Enabling Resources, the fraction with family doctors was associated with increased coverage (MacDonald *et al.*, 2014) and the effect of income quintile was U-shaped (Lemstra *et al.*, 2007; MacDonald *et al.*, 2014). The effect of single parent families was unusual, with the inverted U-shaped response curve in dose 1: an increasing and then decreasing effect on coverage. Subsequent doses saw stronger effects in the expected direction, with lower coverage as the proportion of single parent families increased. The reason for this is not clear from this study as we would have expected decreasing coverage or no change with higher proportions of single parent families (Lemstra *et al.*, 2007; MacDonald *et al.*, 2014). The Health Care System proxy

of former regional health authority displayed heterogeneity as expected. The percentage of families having moved recently showed the opposite effect than expected (MacDonald *et al.*, 2014; Bell, Simmonds and MacDonald, 2015), with increasing mobility associated with increasing coverage. It may be that during the relatively prosperous and high growth times in the province during the study period, the percentage moving is capturing an effect more related to affluence than disconnection with the health care system. This result is puzzling, suggesting an avenue for further research.

Although all factors included in the dose 1 model were significantly related to immunization, the funnel plot and overdispersion measure both show that substantial overdispersion remained after adjustment. This indicates that there are important aspects of dose 1 immunization behavior not captured by the model. This is not surprising as our model did not include data on indicators related to perceived benefits (protecting oneself or others from disease) or risks (fear of needles, pain, risk of adverse events, inconvenience). Previous research has found these factors to have substantial effects on immunization coverage proportions. These missing factors could go a long way to accounting for the remaining geographic variability in coverage and identifying appropriate indicators represents an area for future research.

The models for doses 2 and 3 show the noteworthy phenomenon of under-dispersion: less variability than expected from the statistical model. This suggests that once children have been immunized for their first dose, they are more likely than chance would imply to return for the second dose and third doses. This means that public health is performing well in terms of retention after children enter the immunization system at two months until 6 months. This is an extraordinary "win" for public health programmes and understanding the reasons for this phenomenon could lead to future interventions with improved programme performance.

The dose 4 model then shows substantial overdispersion even after adjusting for the dose 3 coverage proportion and accounting for the additional effects of factors acting only on the forth dose. Again, this indicates an absence of important factors in the dose 4 model. We hypothesize that access and convenience issues play a larger role than in previous doses. The DTaP-IPV-Hib schedule in Alberta has the first three doses early after birth

(months 2, 4, and 6) and the fourth dose at 18 months. By the time a child is 18 months old, routine visits to public health clinics have ended and many parents have likely returned to work, creating health care scheduling difficulties(Bell, Simmonds and MacDonald, 2015). Since the fourth dose is also the final dose, parents may perceive more of the inconveniences of getting the immunization relative to its benefits.

The spatial scan of immunization coverage identified clusters in the unadjusted coverage proportions. However, after adjusting for Predisposing Characteristics, Enabling Resources and Health Care System Characteristics, no clusters were detected. This implies that the apparent clustering is actually due to clustering of the risk factors rather than a geographic anomaly.

The results of the funnel plots, modelling, and spatial scans all have direct implications for public health practice. The substantial reductions in variability between areas when Health Care System Characteristics were introduced into the model suggest that understanding these characteristics across the province could lead to learnings which, if applicable elsewhere, could reduce the disparities in coverage. This is particularly apparent in dose 4. The difference between the best and worst performing former regional health authorities was an increased relative risk of immunization of 20%. This factor accounted for the majority of the overdispersion in both dose 1 and dose 4, substantially reducing the overdispersion in each case. The Enabling Resources of having a family doctor accounted for a large portion of the overdispersion as did the Predisposing Characteristic of higher vitamin K update, and a moderate amount of overdispersion was captured by the Predisposing Characteristic of belonging to a family with three or more children. The public health implication of each of these observations differs, depending upon the type of characteristic. Since the Enabling Resource of having a family doctor accounts for a substantial amount of variability, interventions targeting communities with low family doctor proportions may be indicated. Similarly, targeting programmes providing logistical supports in communities with high proportions of multi-child families may be indicated. The Predisposing Characteristic of low vitamin K uptake, and possible corresponding concerns over vaccine safety or efficacy, suggest areas where public health affecting change is liable to be a long term goal. Such knowledge should be

taken into account when assessing the performance of local immunization programmes. Interestingly, all factors included in the model accounted for some between region variability, as indicated by their statistical significance in the dose 1 model. Had a factor affecting individual immunization behaviour had a similar effect across all areas, and hence not been significant in explaining between area variability, this factor would suggest provincial level programming options. The spatial scan did not identify any anomalous clustering of over or under immunization after adjustment. Any identified clusters would have warranted further investigation to more fully understand what unmeasured effects would be at play. Taken together, these results point to a very complex array of factors with unequal distributions across areas in the province calling for geographically tailored immunization programmes.

4.6 LIMITATIONS

While the combination of modelling and visualization of aggregate immunization data performs well, this study is not without limitations. Some factors known to contribute to vaccination behaviours are not captured in existing data sources. For example, beliefs around pain and fear of needles, (see for example, (MacDonald *et al.*, 2014) for an Alberta study) are not regularly captured. Even when measures of factors are available, they may not be specific. Recent moves from the NHS, as an example, captures moves within five years for all households within the geography. A more specific measure would capture the number of moves within the last two years for parents of two year olds since this is the population of interest and the effect changes across the number of moves (Bell, Simmonds and MacDonald, 2015).

As this study relies upon ecologic analyses, caution around ecologic fallacy must be addressed. The factors in the model have been selected based upon existing studies carried out with individual level data, removing the possibility of incorrect identification of a causal factor. However, the effect estimates themselves should not be interpreted as individual but as area level effects. Also, there is the possibility to misinterpret non-significant risk factors. Non-significant factors in the ecologic model may mean (a) the

factors are not related to immunization coverage, or (b) that the factor does not vary significantly across the geographic areas studied. The former is unlikely as the factor was *a priori* selected because of a relationship, whereas the latter is quite possible to have occurred.

4.7 Conclusion

Surveillance of childhood immunization coverage has historically relied upon measures of coverage at a point in time, such as coverage at age two, national targets, or days delayed. There have been calls for small area surveillance of immunization coverage (Smith and Singleton, 2008), ecologic analyses of immunization data (Omer *et al.*, 2008; Lieu *et al.*, 2015), as well as spatial scans applied to immunization data (Omer *et al.*, 2008; Lieu *et al.*, 2015). These methods provide effective, simple and affordable tools in providing evidence to identify which of a very wide array of public health interventions is most appropriate in the circumstances. We have proposed a system integrating the visualizations provided by funnel plots to understand the variability between regions, with modelling based upon readily available data on factors related to immunization, and with spatial scans to identify anomalous areas of over or under immunization. This system provides public health professionals a sustainable surveillance system with a rich description of the immunization landscape directing effective interventions and programmes.

Table 4-1. Characteristics of the Geographic Sample

Variable	n*	Mean	Std Dev**	Minimum	Maximum
Immunization Coverage					
Dose 1	132	91.0%	5.3%	65.5%	99.1%
Dose 2	132	88.9%	6.3%	59.0%	99.3%
Dose 3	132	86.3%	7.1%	51.5%	98.9%
Dose 4	132	71.0%	11.1%	34.4%	92.2%
System Characteristics					
% moved within 5 years	132	43.8%	12.5%	14.7%	86.8%
Former Health Authority					
1	10				
2	5				
3	29				
4	19				
5	10				
6	27				
7	17				
8	12				
9	3				
Enabling Factors					
% of single parent families	132	22.3%	8.1%	0.0%	44.2%
% with regular family doctor	132	80.3%	4.5%	57.6%	89.9%
Income Quintile					
1 (lowest)	32				
2	32				
3	26				
4	26				
5 (highest)	16				
Predisposing Characteristics					
% with 3+ children	132	19.9%	7.4%	0.0%	50.0%
% Vitamin K uptake	132	99.6%	0.3%	98.1%	100.0%

*n refers to the number of local geographic areas in the study

**standard deviation

Table 4-2. Summary of Regression Model Building reporting Relative Risks (RR) for Ranges

Dose 1				
Parameter	**RR**	*95% confidence Interval*		p-value
Intercept*	0.944	*0.937*	*0.951*	<.0001
% with 3+ children	0.844	*0.813*	*0.876*	<.0001
% Vitamin K uptake	1.072	*1.057*	*1.087*	<.0001
% moved within 5 years	1.036	*1.018*	*1.055*	<.0001
% with regular family doctor	1.075	*1.058*	*1.094*	<.0001
% of single parent families	1.001	*1.001*	*1.002*	<.0001
Income Quintile**	1.020	*1.011*	*1.029*	0.0068
Former Health Authority **	1.075	*1.056*	*1.095*	<.0001

Dose 2				
Parameter	**RR**	*95% confidence Interval*		p-value
Intercept*	0.973	*0.966*	*0.981*	<.0001
% with 3+ children	0.944	*0.906*	*0.985*	0.0073
% Vitamin K uptake	0.996	*0.983*	*1.009*	0.5538
% moved within 5 years	0.991	*0.974*	*1.01*	0.3827
% with regular family doctor	1.025	*1.005*	*1.045*	0.0152
% of single parent families	0.979	*0.959*	*0.999*	0.0125
Income Quintile**	1.005	*1.003*	*1.014*	0.2597
Former Health Authority **	1.024	*1.007*	*1.041*	0.5651

Table 4-2. – continued –

Dose 3				
Parameter	**RR**	*95% confidence Interval*		p-value
Intercept*	0.982	*0.974*	*0.990*	<.0001
% with 3+ children	0.930	*0.887*	*0.976*	0.0030
% Vitamin K uptake	1.005	*0.990*	*1.002*	0.5270
% moved within 5 years	0.991	*0.970*	*1.012*	0.2902
% with regular family doctor	1.014	*0.991*	*1.037*	0.1703
% of single parent families	0.992	*0.970*	*1.015*	0.0037
Income Quintile**	1.009	*0.998*	*1.020*	0.4022
Former Health Authority **	1.030	*1.006*	*1.054*	0.1291

Dose 4				
Parameter	**RR**	*95% confidence Interval*		p-value
Intercept*	0.902	*0.890*	*0.914*	<.0001
% with 3+ children	0.858	*0.798*	*0.922*	<.0001
% Vitamin K uptake	1.027	*1.003*	*1.052*	0.0282
% moved within 5 years	0.999	*0.966*	*1.032*	0.9040
% with regular family doctor	1.029	*0.993*	*1.066*	0.0931
% of single parent families	0.914	*0.885*	*0.944*	<.0001
Income Quintile**	1.024	*1.007*	*1.041*	0.7740
Former Health Authority **	1.206	*1.176*	*1.237*	<.0001

*Intercept term in Dose 2-4 models is the previous doses' value

**relative risks are between the highest and lowest categories

The relative risk (RR) measures the change in risk over a percentage-point range, given by:

Variable:	Percentage-point range
% with 3+ children	50
% Vitamin K uptake	1
% moved within 5 years	60
% with regular family doctor	20
% of single parent families	30

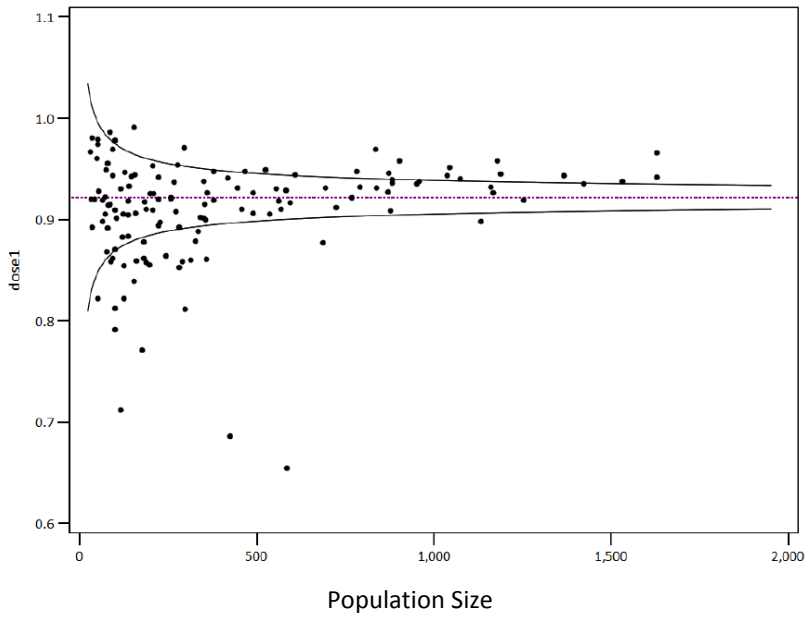
Table 4-3. Overdispersion

	Factor Added to Model		Overdispersion Measure			
			Dose 1	Dose 2	Dose 3	Dose 4
NULL Model*			12.50	1.08	1.01	7.47
System Characteristics	Former Health Authority		7.76	0.95	0.76	2.96
	Moved within last 5 years		7.12	0.95	0.77	2.95
Enabling Factors	Single Parent		6.96	0.79	0.62	2.26
	Income Quintile		6.96	0.80	0.61	2.23
	Regular family doctor		5.89	0.76	0.61	2.24
Predisposing Characteristics	3+ children		5.51	0.70	0.54	2.12
	Vitamin K uptake		4.50	0.71	0.54	2.09
Final overdispersion			4.50	0.71	0.54	2.09

**Null Model for dose 1 contains only an intercept term; null models for doses 2-4 contain the previous doses' observed value.*

Figure 4-1. Funnel Plots of unadjusted and adjusted immunization coverage

Dose 1 – Unadjusted



Dose 1 – Adjusted

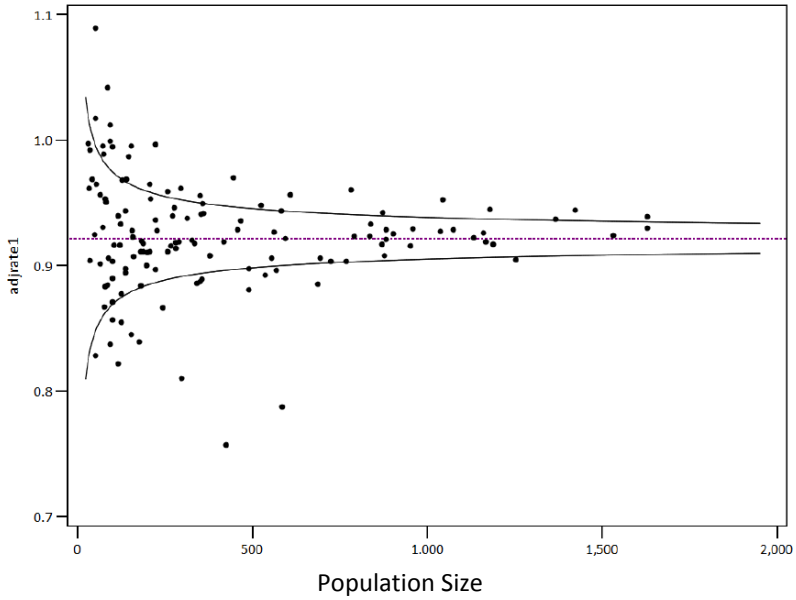
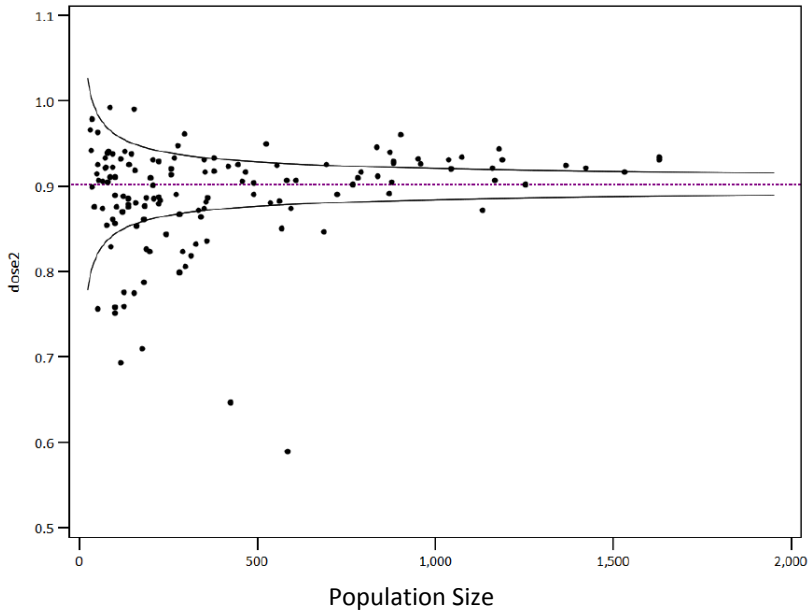


Figure 4-1. – continued –

Dose 2 – Unadjusted



Dose 2 – Adjusted

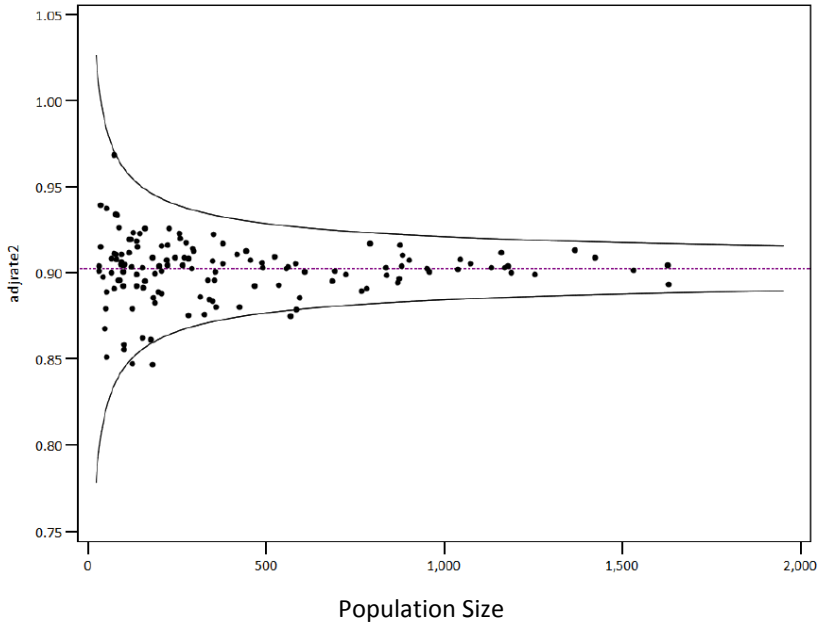
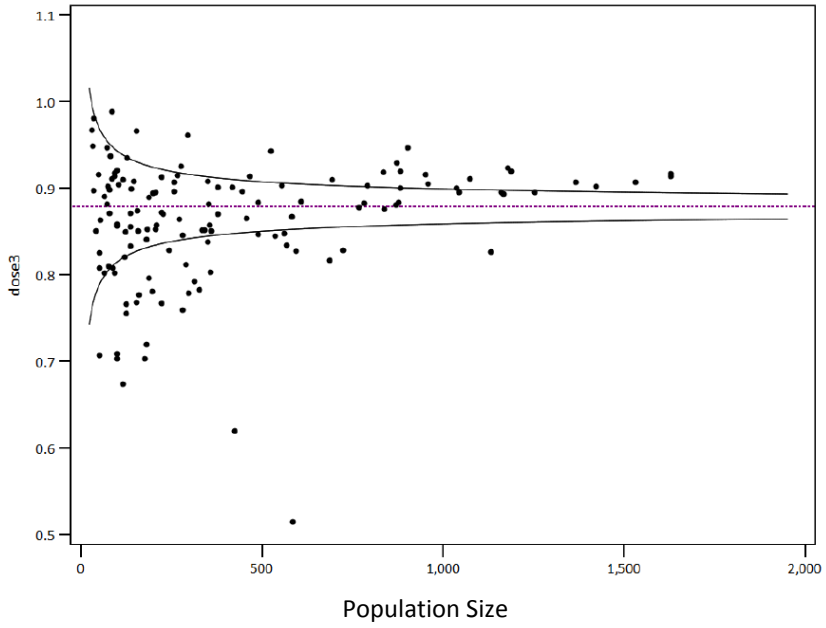


Figure 4-1. – continued –

Dose 3 – Unadjusted



Dose 3 – Adjusted

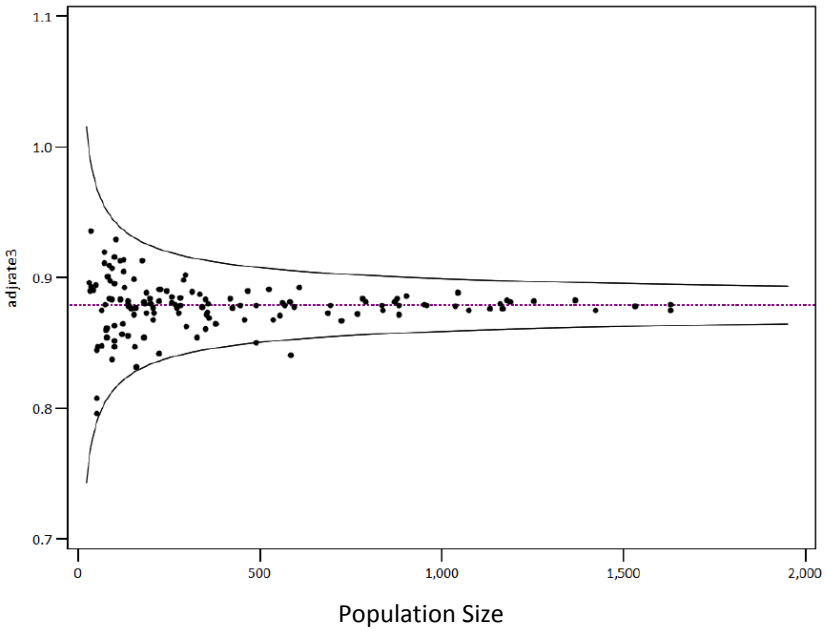
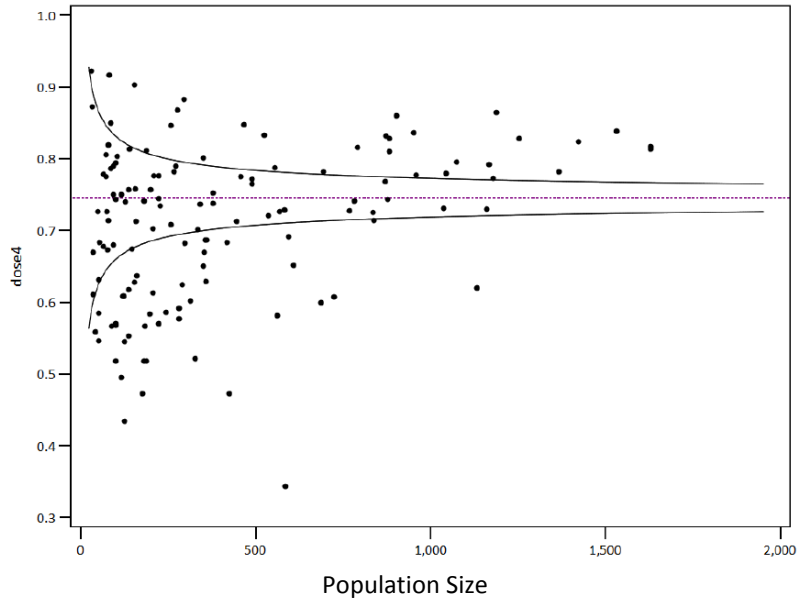
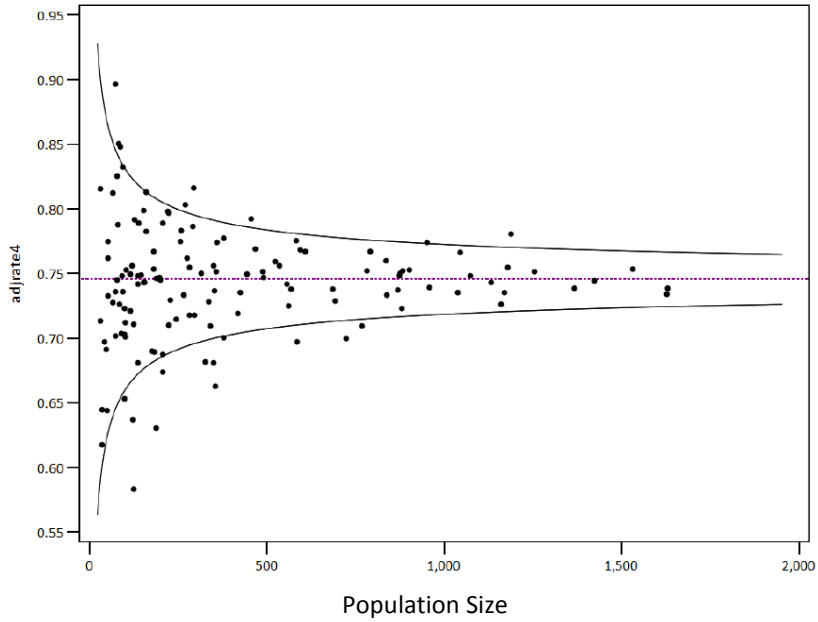


Figure 4-1. – continued –

Dose 4 – Unadjusted



Dose 4 – Adjusted



4.8 Appendix: Regression Results

The relative risk results presented in Table 4-2 are readily interpretable. They are based on transformations of the log-binomial regression results, but are made interpretable by computing the relative risk over relevant data ranges for each of the covariates. For completeness, the full model results and correlations are shown in Table 4-4 and Table 4-5.

Table 4-4. Log-Binomial Regression Results

Parameter		Dose 1			Dose 2			Dose 3			Dose 4		
		Estimate	*Standard Error*	*p-value*	*Estimate*	*Standard Error*	*p-value*	*Estimate*	*Standard Error*	*p-value*	*Estimate*	*Standard Error*	*p-value*
*Intercept**		-0.0577	0.0035	*<.0001*	-0.0270	0.0038	*<.0001*	-0.0182	0.0044	*<.0001*	-0.1031	0.0068	*<.0001*
% with 3+ children		-0.0034	0.0004	*<.0001*	-0.0011	0.0004	*0.0073*	-0.0014	0.0005	*0.0030*	-0.0030	0.0007	*<.0001*
% Vitamin K uptake		0.0696	0.0070	*<.0001*	-0.0040	0.0068	*0.5538*	0.0049	0.0078	*0.5270*	0.0268	0.0122	*0.0282*
% moved within 5 years		0.0006	0.0001	*<.0001*	-0.0001	0.0002	*0.3827*	-0.0002	0.0002	*0.2902*	0.0000	0.0003	*0.9040*
% with regular family doctor		0.0036	0.0004	*<.0001*	0.0012	0.0005	*0.0152*	0.0008	0.0006	*0.1703*	0.0015	0.0009	*0.0931*
% of single parent families		0.0012	0.0003	*<.0001*	-0.0006	0.0002	*0.0125*	-0.0008	0.0003	*0.0037*	-0.0032	0.0004	*<.0001*
% of single parent families squared		-0.0016	0.0004	*<.0001*									
Income Quintile	2	-0.0106	0.0039	*0.0068*	-0.0048	0.0043	*0.2597*	-0.0041	0.0049	*0.4022*	-0.0023	0.0079	*0.7740*
(reference=1)	3	-0.0110	0.0039	*0.0052*	-0.0043	0.0044	*0.3355*	0.0011	0.0049	*0.8227*	0.0002	0.0076	*0.9751*
	4	-0.0198	0.0043	*<.0001*	-0.0024	0.0045	*0.5922*	-0.0027	0.0052	*0.5987*	0.0001	0.0081	*0.9932*
	5	-0.0067	0.0039	*0.0894*	-0.0013	0.0044	*0.7715*	-0.0088	0.0050	*0.0786*	-0.0241	0.0078	*0.0020*
Former	1	-0.0458	0.0081	*<.0001*	0.0046	0.0080	*0.5651*	-0.0136	0.0090	*0.1291*	-0.1345	0.0156	*<.0001*
Health Authority	2	0.0082	0.0060	*0.1772*	-0.0020	0.0073	*0.7803*	-0.0136	0.0088	*0.1248*	-0.1431	0.0172	*<.0001*
(reference=3)	4	-0.0069	0.0055	*0.2059*	-0.0021	0.0064	*0.7473*	-0.0097	0.0071	*0.1738*	-0.1872	0.0128	*<.0001*
	5	0.0070	0.0074	*0.3463*	0.0219	0.0069	*0.0014*	0.0004	0.0081	*0.9560*	-0.0279	0.0147	*0.0582*
	6	0.0006	0.0032	*0.8614*	0.0150	0.0034	*<.0001*	-0.0070	0.0038	*0.0676*	-0.0388	0.0060	*<.0001*
	7	0.0266	0.0061	*<.0001*	0.0111	0.0072	*0.1209*	0.0020	0.0080	*0.8084*	-0.0348	0.0129	*0.0072*
	8	-0.0141	0.0085	*0.0985*	0.0027	0.0095	*0.7792*	-0.0297	0.0112	*0.0079*	-0.1699	0.0183	*<.0001*
	9	-0.0081	0.0096	*0.3995*	0.0108	0.0090	*0.2277*	-0.0016	0.0105	*0.8785*	-0.0753	0.0185	*<.0001*

Table 4-5. Correlation Matrix

	Dose 1	*Dose 2*	*Dose 3*	*Dose 4*	*% with regular family doctor*	*% with 3+ children*	*% of single parent families*	*% moved within 5 years*	*% Vitamin K uptake*
Dose 1	1.00	0.95	0.89	0.68	0.17	-0.38	0.17	0.16	0.24
Dose 2	0.95	1.00	0.95	0.76	0.21	-0.37	0.07	0.14	0.20
Dose 3	0.89	0.95	1.00	0.82	0.28	-0.40	0.00	0.15	0.18
Dose 4	0.68	0.76	0.82	1.00	0.22	-0.38	-0.10	0.16	0.09
% with regular family doctor	0.17	0.21	0.28	0.22	1.00	0.06	-0.30	-0.05	-0.06
% with 3+ children	-0.38	-0.37	-0.40	-0.38	0.06	1.00	-0.34	-0.42	0.12
% of single parent families	0.17	0.07	0.00	-0.10	-0.30	-0.34	1.00	0.26	0.11
% moved within 5 years	0.16	0.14	0.15	0.16	-0.05	-0.42	0.26	1.00	-0.04
% Vitamin K uptake	0.24	0.20	0.18	0.09	-0.06	0.12	0.11	-0.04	1.00

4.9 REFERENCES

Aday, L. A. and Andersen, R. (1974) 'A Framework for the Study of Access to Medical Care', *Health services research*, 9(3), pp. 208–220.

Alberta Ministry of Health (2017a) *Interactive Health Data Application (IHDA) - Demographics - Population Estimates - Adjusted*. Available at http://www.ahw.gov.ab.ca/IHDA_Retrieval/.

Alberta Ministry of Health (2017b) *Interactive Health Data Application (IHDA) - Immunization - Childhood Coverage Rates*. Available at http://www.ahw.gov.ab.ca/IHDA_Retrieval/.

Béland, Y. (2002) 'Canadian community health survey--methodological overview', *Health reports*. Statistics Canada, 13(3), p. 9.

Bell, C. A., Simmonds, K. A. and MacDonald, S. E. (2015) 'Exploring the heterogeneity among partially vaccinated children in a population-based cohort', *Vaccine*. Elsevier, 33(36), pp. 4572–4578.

Birkmeyer, J. D. (2001) 'Primer on Geographic Variation in Health Care', *Effective Clinical Practice*, 4(5), pp. 232–233.

de Cantuária Tauil, M., Sato, A. P. S. and Waldman, E. A. (2016) 'Factors associated with incomplete or delayed vaccination across countries: a systematic review', *Vaccine*. Elsevier, 34(24), pp. 2635–2643.

Center for Disease Control (CDC) (2011) 'Ten Great Public Health Achievements—United States , 2001-2010', *JAMA*, 306(1), pp. 36–38.

Dover, D. and Schopflocher, D. (2011) 'Using funnel plots in public health surveillance', *Population Health Metrics*, 9(1), p. 58.

Dubé, E. *et al.* (2013) 'Vaccine hesitancy: an overview', *Human vaccines & immunotherapeutics*. Taylor & Francis, 9(8), pp. 1763–1773.

Falagas, M. E. and Zarkadoulia, E. (2008) ‘Factors associated with suboptimal compliance to vaccinations in children in developed countries: a systematic review’, *Current medical research and opinion*. Taylor & Francis, 24(6), pp. 1719–1741.

Hamel, M. and Laniel, N. (2014) ‘Producing official statistics via voluntary surveys--the National Household Survey in Canada’, *Statistical Journal of the IAOS*. IOS Press, 30(3), pp. 237–242.

Larson, H. J. *et al.* (2014) ‘Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007--2012’, *Vaccine*. Elsevier, 32(19), pp. 2150–2159.

Lee, L. *et al.* (2010) *Principles and Practice of Public Health Surveillance*. 3rd edn. Oxford University Press, USA.

Lemstra, M. *et al.* (2007) ‘Disparity in childhood immunizations.’, *Paediatrics & child health*, 12(10), pp. 847–852.

Lieu, T. *et al.* (2015) ‘Geographic clusters in underimmunization and vaccine refusal.’, *Pediatrics TA* -, 135(2), pp. 280–289.

MacDonald, S. E. *et al.* (2014) 'Parental concern about vaccine safety in Canadian children partially immunized at age 2: A multivariable model including system level factors', *Human Vaccines & Immunotherapeutics,* 10(9), pp.2603–2611.

Mills, E. *et al.* (2005) ‘Systematic review of qualitative studies exploring parental beliefs and attitudes toward childhood vaccination identifies common barriers to vaccination’, *Journal of clinical epidemiology*. Elsevier, 58(11), pp. 1081–1088.

Omer, S. B. *et al.* (2008) ‘Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis’, *American Journal of Epidemiology*. Oxford Univ Press, 168(12), pp. 1389–1396.

Pearce, A. *et al.* (2008) ‘Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study’, *bmj*. British Medical Journal Publishing Group, 336(7647), pp.

754–757.

Sahni, V., Lai, F. Y. and MacDonald, S. E. (2014) 'Neonatal vitamin K refusal and nonimmunization', *Pediatrics*. Am Acad Pediatrics, pp. 497-503.

Smith, P. J. and Singleton, J. A. (2008) 'Vaccination Coverage Estimates for Selected Counties: Achievement of Healthy People 2010 Goals and Association with Indices of Access to Care, Economic Conditions, and Demographic Composition', *Public health reports*. Association of Schools of Public Health, 123(2), pp. 155–172.

Thomson, A., Robinson, K. and Vallée-Tourangeau, G. (2016) 'The 5As: A practical taxonomy for the determinants of vaccine uptake', *Vaccine*. Elsevier, 34(8), pp. 1018–1024.

Wacholder, S. (1986) 'Binomial regression in GLIM: estimating risk ratios and risk differences.', *American journal of epidemiology*, 123(1), pp. 174–84.

van der Wal, M. F. *et al.* (2005) 'Vaccination rates in a multicultural population.', *Archives of disease in childhood*, 90(1), pp. 36–40.

World Health Organization (WHO) (2014) *Report of the SAGE working group on vaccine hesitancy*. Available at http://www.who.int/immunization/sage/meetings/2014/october/1_Report_WORKING_GROUP_vaccine_hesitancy_final.pdf.

CHAPTER 5: DISCUSSION

5.1 CONTEXT AND AIM

The overarching aim of this set of research papers is to advance methods and practice in public health surveillance. To better understand how the methods developed and proposed in this thesis fit into the current landscape of surveillance methods, I begin by placing them within their historical context.

The Father of Modern surveillance, William Farr (Langmuir, 1976), first identified and brought together the underlying concepts of surveillance: standardized data collection, taking a population health perspective using rates, comparing the observed to expected, interpreting these analyses, and calling for action. While these principles still underlie, and in fact largely define, surveillance, much has changed from its origins in communicable diseases. In the early 21st century, a shift in focus started to include chronic diseases and their risk factors into surveillance activities (Choi, 2012). Administrative health data was first used, disease specific registries for cancer came into existence, and national health surveys were implemented. These all provided rich, new sources of data in public health surveillance. Simultaneously, quantitative methods in statistics and epidemiology were also flourishing, leading to the suite of methods in use today. Surveillance activities regularly standardize and often use sophisticated methods for aberration detection (cumulative sums or segmented regressions in time, spatial scans in space). But interestingly, the disparate original data sources underlying our surveillance information systems today still have influence over the types of analysis routinely undertaken and reported. The methods described in this thesis further refine methods in aberration detection (through identification of multiple irregular spatial-temporal clusters) and bring together the many data silos in a coherent process for analysis (funnel plots and the methodology for adjusting), interpretation and reporting (founded in the use of funnel plots) the complex health outcomes of importance today.

5.2 Summary and Implications of Research

This thesis has proposed sophisticated analyses of aggregate data for inclusion in to the surveillance methods toolkit. The areas of risk factor surveillance and health service utilization surveillance are near the frontier of surveillance activities, although the state of the art remains simple descriptive statistics. For example, data on a multitude of risk factors are currently available – physical activity, fruit and vegetable consumption, drinking behaviours, obesity, and levels of stress are all available from population health surveys. Similarly, utilization of hospital resources, emergency department resources, and specific programmes like immunization are available from administrative data sources. But linking these separate resources is not a part of the norm. My thesis hopes to move public health surveillance in that direction by applying sophisticated methods to aggregate data. I now summarize how each paper addresses this objective.

5.2.1 Funnel Plot Methodology for Public Health Surveillance

The paper in Chapter 2 examining the use of funnel plots in public health surveillance developed two themes, the concept of overdispersion and the policy relevant interpretation of geographic variability. Overdispersion can occur when relevant risk factors remain unaccounted for. The aggregate data methods developed allow for the relatively straightforward accounting for risk factors with an aggregate regression methodology. This is of particular utility in public health surveillance where a large number of factors are collected across disparate systems i.e. where data is available, but often only in an aggregate fashion. The way in which these known factors are included in the aggregate regression framework leads directly to public health policy implications. When a known factor is not significant in an aggregate regression, this implies that there is no *between area* variability. This implies that policy actions related to this factor should be undertaken provincially as all geographic areas are similarly affected. Other factors are then examined sequentially, ordered by their amenability to public health intervention. In this way, priority factors can be targeted to specific areas. The identification of which areas should be targeted can then carried out based upon an iterative use the funnel plot methodology. By extending the use of the funnel plot

contours to maps, an effective visualization of the geographic distribution emerges. In total, the funnel plot methodology provides public health professionals with a rigorous, informative analysis process that is easy to implement.

5.2.2 MULTSCAN SPATIAL SCAN STATISTIC

Having dealt with systematic factors affecting health outcomes in the funnel plot methodology, the spatial scan paper in Chapter 3 addressed identification of anomalies in geographic areas. Existing cluster detection methods generally rely upon predefined shapes – cylinders (Kulldorff *et al.*, 2006), rectangles (Neill and Moore, 2004) – with the assumption that the shape of the cluster is the same through time. From a health perspective, this means that the cluster is geographically static. This assumption is clearly untenable for communicable diseases and unlikely in many other health states. A novel feature of MultScan is its ability to identify irregularly shaped clusters in both space and time. MultScan also improves the positive predictive value of the detected clusters vis-à-vis existing methods. This is important to be able to better target public health actions or better inform research hypotheses. MultScan accomplishes this improvement to positive predictive value using a data driven approach to ordering the data inspired by funnel plots. Finally, MultScan uses a methodologically rigorous method to detect simultaneous clusters.

5.2.3 IMMUNIZATION SURVEILLANCE

Geographic variation in immunization coverage for DTaP-IPV-Hib by age two in Alberta was examined in the empirical study in Chapter 4. The aggregate model for coverage was adjusted for Health Care Systems Characteristics (recent moves, immunization provider), Enabling Factors (socio-economic status, having a family doctor, number of parents), and Predisposing Characteristics (Vitamin K uptake, number of children). All factors were found significantly related to dose 1 coverage, implying public health immunization programming aimed at the first dose would need to be geographically tailored. Health Care System Characteristics have the highest potential to increase coverage as it accounted for a substantial portion of the reduction in overdispersion and is amenable to

change. The under dispersion observed in doses 2 and 3 suggests that once children begin receiving immunizations, they continue to receive immunizations (with less variability than chance would suggest). The dose 4 model displayed substantial overdispersion. The Health Care System Characteristics accounted for most of the explained variation, suggesting that there is large between provider variability dealing with dose 4. This provides the opportunity to learn from those regions that are performing particularly well. The remaining overdispersion in dose 1 and dose 4 identified an area for future investigation. It may be that known factors that were not included in the model, for example, perceptions of risks and benefits like the pain around immunization, account for this overdispersion, and it would be of value to collect indicators or identify proxies for these factors. The other possibility is that there are a number of unknown factors influencing immunization uptake, and the geographic variability might assist in designing research to identify those factors.

The known factors that were accounted for in the analysis of immunization represent a substantial contribution to public health surveillance of childhood immunization. The factors included represent a wide range of determinants of immunization. For instance, vaccine refusers have not been systematically accounted for in previous surveillance methods. With the inclusion of vitamin K uptake as a proxy for immunization related attitudes, it is now possible to accurately represent the performance of the health authority providing immunization services in regions with high numbers of refusers (low vitamin K uptake). It is also possible to now scan for clustering of immunization behaviour beyond these known determinants that have already been accounted for.

5.2.4 Summary

From a broader perspective, applying the funnel plot methodology with the MultScan spatial scan methodology provides a deeper understanding of geographic variation in immunization in the province. This type of analysis provides an example of what is possible with existing data sources in public health surveillance. Currently, immunization surveillance does not regularly use such a suite of sophisticated methods and data sources in surveillance practice; rather, studies of this scope generally are research projects. By shifting the focus from discovery of factors to explanation of variation, a wide range of

factors can now be included in our understanding of variability. The process of adjusting for these factors highlights a number of things. First, it shows which factors explain what proportions of variability in the geographic areas using overdispersion. The residual overdispersion highlights areas for surveillance system development or additional research, as this represents factors not accounted for in the modelling process. It then highlights which factors are amenable to province wide interventions and which factors are better approached using targeted methods. Again, insignificant variables in the ecologic regression are the same effect across areas while significant factors represent underlying differences in the factors. Visualizing these results with funnel plots provides and easy to comprehend visualization of the variability, and potential areas to prioritize for intervention. Finally, spatial scans such as MultScan provide information on geographic clustering due to something other than the factors included in the model. In total, this provides a very rich, comprehensive surveillance system that is sustainable and cost effective.

These methods of visualization, analysis and cluster detection all rely upon, and take advantage of, surveillance data aggregated to predefined geographic boundaries. This thesis focused on gathering many characteristics of each geography. This is a shift from the current practice focusing on boundaries alone (Yiannakoulias, 2011). This view of geographies having rich characteristics is a step toward a fuller recognition of place. Embedding these geographic characteristics into the visualization, analysis, and cluster detection in turn compels public health practitioners to carefully consider and interpret the results with attention to not just the geographic area, but the compositional and contextual characteristics accounted for, and possibly a more holistic view of place. It is my hope that the methods described result in more contemplation of the concept of place, not just geography, in future public health surveillance endeavours.

5.3 Potential Limitations

There are potential limitations within each of the studies. These limitations range from issues with interpretation to data availability to fundamentals related to the method being discussed. However, as with any research endeavour, the aim has been to incrementally

improve the field. These incremental improvements in methods can improve surveillance practices, supporting Langmuir's assertion that "Good surveillance does not necessarily ensure the making of the right decisions, but it reduces the chances of wrong ones" (Langmuir and others, 1963).

5.3.1 Funnel Plot Methodology for Public Health Surveillance

I have proposed funnel plots as a statistically robust method for visualizing and understanding aggregate data. The most obvious limitation of funnel plots has to do with their a-spatial nature. Funnel plots rely upon pre-defined aggregations. Hence, when a slight elevation is observed in one area, the funnel plot alone cannot provide useful information about the surrounding areas. I have demonstrated this limitation can be overcome using supplemental procedures such as spatial scans and maps using funnel plot contours.

Another limitation of the funnel plot methodology relates to the interpretation of results. The substantial difference in interpreting the aggregate regression vis-à-vis an individual level regression can be a challenge to communicating results to users. This challenge can be offset by the clear policy implications that result from the aggregate regression methodology.

Finally, the funnel plot methodology and its interpretations rely upon the modelling and interpretation of overdispersion. As with all statistical models, the assumptions of the model must be met. It is entirely possible that the unaccounted for overdispersion that is observed is due to misspecification of the statistical model. This limitation is not unique to the funnel plot methodology outlined.

5.3.2 MultScan Spatial Scan Statistic

MultScan begins by sorting data by the Score z-score. This data-based approach results in many of the improved performance characteristics observed with MultScan, but is also the source of its primary limitation. The ordering forces and defines the clusters. For example, the first area associated with the largest z-score in the list must be included in

the multiple clusters identified by MultScan. This lack of ability to flexibly search including or excluding any particular area has the potential to reduce the accuracy of MultScan. This restricts the set of geographies being considered as clusters. This limitation is, I believe, offset by the increases in positive predictive value attained by MultScan. MultScan single largely over comes this limitation in the single cluster detection case by, at each step going through the sorted data, identifying the single cluster with the maximal likelihood of all clusters identified at that step. In this way, any spurious z-score that is geographically isolated can be excluded while still identifying a true cluster.

MultScan shares a limitation common with all spatial scan methods using pre-defined geographic areas. The scale of the areas may not match the scale of the cluster that is trying to be detected. When the cluster occurs at the scale of one or more areas, the cluster detection methods will perform as expected and be able to detect the cluster. However, if the scale of the cluster is below the scale of the pre-defined geographies (within only a part of one area), or if the cluster occurs in multiple areas but not fully covering any of them, spatial scan statistics will have great difficulty detecting them (Assunção *et al.*, 2006). This can represent a significant challenge in public health practice. As the scale shrinks, it is increasingly difficult to obtain reliable data: privacy becomes a concern, population denominators are more difficult to accurately estimate, attributing cases to exact geographies is error prone, and few risk factors are available at very small scales. Spatial scanning when the cluster scale is below the current boundary scale will likely be ineffective. Public health practitioners need to be mindful of these issues when applying MultScan or any other cluster detection in practice.

5.3.3 IMMUNIZATION SURVEILLANCE

The availability of factors related to immunization uptake is the greatest limitation in the empirical immunization uptake study. The study was restricted to using aggregate, already available measures of the factors related to immunization uptake. Measures for many determinants of immunization uptake were not available. For example, no indicators or related proxies for fear of needles were identified. In other cases, direct

measures of factors were not available, but proxy measures were. For example, Vitamin K uptake was used as a proxy for immunization related beliefs. The accuracy of the regression model results depends upon the level of measurement error in the proxy measurement. This will usually underestimate the amount of between area variation attributable to the factor.

The Behavioural Model of Health Services Utilization was originally designed choosing constructs that "can be operationalized in social survey research" (Andersen and Newman, 1973). It provides a clear categorization and way to describe the factors influencing immunization uptake. Having the dimensions of Predisposing Characteristics, Enabling Resources and Health Care System Characteristics provides clear, meaningful and readily communicated categorizations as well as linking to interventions. However, there are a number of other taxonomies that could be employed. Tauil *et al* use a taxonomy very similar to the Behavioural Model of Health Services Utilization with Family Features, Parents' Knowledge and Attitudes, and Health Services dimensions (de Cantuária Tauil, Sato and Waldman, 2016). The Vaccine Hesitancy model focuses on separating vaccine specific from generic influences, looking at the dimensions of Individual/Social Group Influences, Contextual Influences and Vaccine and Vaccination-Specific Issues (Larson *et al.*, 2014). The social ecological model posits intrapersonal, interpersonal, institutional, community and policy levels (Kumar *et al.*, 2012) affecting immunization uptake. Thomson *et al* developed an "intuitive taxonomy", the "5As". It incorporates the five dimensions of Access, Affordability, Awareness, Acceptance, and Activation and is claimed to have "facilitated mutual understanding of the primary determinants of suboptimal coverage within inter-sectoral working groups." (Thomson, Robinson and Vallée-Tourangeau, 2016). In terms of communication of analyses and results for the purpose of public health action, any of these taxonomies could be successfully utilized. In practice, those undertaking the analysis and dissemination will have to make the determination of which taxonomy will be best understood by the public health professions they wish to inform and who will be taking the public health actions based upon this information.

5.4 Future Research

5.4.1 Funnel Plot Methodology for Public Health Surveillance

A key element of surveillance is that monitoring occurs *over time*. The funnel plots described in this thesis capture variability at only a single point in time. To be more broadly useful, the temporal dimension will have to be incorporated. This raises a number of interesting issues for future research. Consider first, the residual overdispersion observed in the fourth dose of DTaP-IPV-Hib in Alberta. If one were interested in performing the same analysis of immunization uptake in the following year, it is unclear what the best approach would be. It is possible to completely replicate the study design and analysis methods on the next years' data. This would, in all likelihood, result in a similar finding of residual overdispersion. At a high level, this is a useful approach to monitoring changes in the effects of factors and the fraction of unexplained overdispersion. However, it might be more fruitful to ask a slightly different question, along the lines of "Has anything changed?". In this case, there may be merit to assuming that each aggregate geography has a suite of unmeasured factors responsible for the residual overdispersion. This could be quantified as the area level residual from the regression model. This estimate of the net effect of unmeasured factors could then be entered in future analyses as an additional geography specific factor. This shifts the question away from identifying the unmeasured factors to identifying unexplained changes in the small areas.

Visualizing funnel plots over time is also an interesting problem. Funnel plots were designed as two-dimensional plots. A visual approach to incorporating time into them will thus be a challenge. It would be possible to imagine a three-dimensional funnel plot created as a stack of regular funnel plots. Geographies could then be identified as connected series through the funnel plots. The general adoption of this would require specialized software and not be amenable to the two-dimensional reporting common place in surveillance today. The issue of interpretation of a time series of funnel plots must be addressed. The meaning of being outside of control limits for, say, three years in a row and the interpretation of being outside control limits for only one out of three years must be clarified in an approach like this. However, replicating a funnel plot is not the

only option. A method taking into account the time series nature of the data could also be used. For example, CUSUMs – cumulative sums – are commonly used in single series outbreak detection (Lawson and Kleinman, 2005). Since CUSUMs change with additional time series data, the idea of placing CUSUMs (or similar measures) in funnel plots is feasible and could also be investigated. The extension of the funnel plot methodology to incorporate time is a promising are for continued research.

5.4.2 MULTSCAN SPATIAL SCAN STATISTIC

The first step of MultScan is to sort the aggregate data by Score z-score. Geographic clusters are then created by sequentially following this sorted list and checking adjacency conditions. Generalizing this idea of starting with an a-spatial set and then creating the clusters from it, leads to an entire class of potential clustering algorithms. A promising future avenue would be the application of sparse data methods, for example, those used to identify gene and microarray effects, to the a-spatial data. Clusters could then be created, as in MultScan, from the identified areas. One potential approach would be LASSO regression. LASSO regression, standing for least absolute shrinkage and selection operator, is a sparse data regression method that uses shrinkage to reduce over fitting and simultaneously performs covariate selection (Tibshirani, 1996). In epigenetics the LASSO is often used to identify the small number of genes that are expressing out of a large number of genes tested. Using spatial-temporal units in place of genes, LASSO regression has the potential to quickly identify the set of geographies with elevated risk. After identification by LASSO regression, these areas could then be formed into geographic clusters. This is a well-established method used in sparse data which can be generalized for most common distributions (Normal, Binary, Poisson) and can easily include covariates. While designed to deal with sparse data and therefore is naturally "looking for" the few areas of elevated risk, it may still suffer inefficiency due to not the incorporating geographic adjacency information.

Spatial scanning algorithms like MultScan depend upon the key idea of adjacency. Scanning techniques in fact started in adjacency in time (Glaz, Pozdnyakov and Wallenstein, 2009) and have since been extended to adjacency in space and time

(Kulldorff and Nagarwalla, 1995). It should be possible to extend the algorithm to additional dimensions whenever adjacency can be defined. Two public health relevant scenarios are briefly considered.

First, extending form a space-time scan to a space-time-age group scan is considered. Spatial scans today deal with age by assuming a constant age effect within space and time. This is accomplished using age-standardized rates or by controlling for age in a regression framework. Age groups have a natural adjacency: for example, 50-59 is clearly adjacent to 60-69 but not 70-79. By adding this additional dimension to the MultScan algorithm, clusters in space-time-age groups could be identified. In the same way that MultScan can identify a cluster whose shape changes over time, it would also be possible to identify changes in the age groups affected over space and time.

The second instance where adjacency can be well defined is in disease coding. International Classification of Disease codes (ICD codes), used in the coding and analysis of mortality and morbidity, are hierarchical. It would be possible to have all codes within (say) a chapter considered to be adjacent. Similarly, the Anatomical Therapeutic Chemical (ATC) classification system is a five level hierarchy designed to classify drugs into therapeutic and chemical subgroups (WHO Collaborating Centre for Drug Statistics Methodology, no date) could be used to scan for adverse drug reactions. Analysis of health outcomes along these lines can identify previously unexpected relationships that can be used in public health.

In each case, the extensions to MultScan for scanning additional dimensions are possible because a measure of adjacent-ness can be constructed for each new dimension being scanned. Many classification systems and elements of interest in epidemiology and public health are hierarchical or otherwise have obvious adjacency measures associated with them. This class of problems has been approached previously, for example Kulldorff's spatial scan has been adapted to a hierarchy of ICD codes (Kulldorff, Fang and Walsh, 2003). While a comprehensive evaluation of this method does not currently exist, it is likely to suffer the same issue with poor positive predictive value in hierarchal scanning as in spatial scanning. In this case, a MultScan based approach has the potential to address this.

5.4.3 Immunization Surveillance

The paper in Chapter 3 examining childhood DTaP-IPV-Hib immunization in Alberta illustrates how a rich immunization surveillance system can be created. The study focused on one outcome measure, population-level immunization coverage. This is an oft reported outcome as herd immunity is a key component to protecting the population from a vaccine preventable disease. However, other potential measures are possible. By using the funnel plot adjustment methodology, it will be possible to create performance measures that previously would have only been possible with expensive data collection at the individual level. Consider an example with religious beliefs greatly affecting immunization uptake in some geographic areas. It would be "unfair" to include the non-immunizing children in evaluating the efficiency or efficacy of an immunization programme. It would be possible to compute a modified coverage rate, removing this unfair component. This modified coverage could then be used to monitor performance and compare performance between areas and providers. The approach can be extended to more complex immunization related measures such as time to immunization curves and total days delayed. This opens a wide array of potential advances in immunization surveillance.

5.5 Conclusion

Public health surveillance is a fundamental activity in public health. This thesis has examined ways to strengthen the public health surveillance of immunization coverage through the use of aggregate methods to visualize data using funnel plots, adjusting immunization rates for factors with available aggregate data, and identifying areas of unusual immunization uptake using MultScan, a novel spatial scan method. Further research has been proposed to expand the scope of these projects to encompass more and varied measures of immunization performance. Putting these methods into public health surveillance practice is the next step.

5.5.1 Post Script

At the time of writing, I am pleased to report an active collaboration with an Alberta Health Services Zone. This collaboration is designed to translate the childhood immunization surveillance methodology from Chapter 4 to regional public health surveillance practice.

5.6 REFERENCES

Andersen, R. and Newman, J. F. (1973) 'Societal and Individual Determinants of Medical Care Utilization in the United States', *The Milbank Memorial Fund Quarterly. Health and Society*, 51(1), p. 95.

Assunção, R. *et al.* (2006) 'Fast detection of arbitrarily shaped disease clusters.', *Statistics in medicine*, 25, pp. 723–742.

de Cantuária Tauil, M., Sato, A. P. S. and Waldman, E. A. (2016) 'Factors associated with incomplete or delayed vaccination across countries: a systematic review', *Vaccine*. Elsevier, 34(24), pp. 2635–2643.

Choi, B. C. K. (2012) 'The past, present, and future of public health surveillance', *Scientifica*. Hindawi Publishing Corporation, 2012.

Glaz, J., Pozdnyakov, V. and Wallenstein, S. (2009) *Scan statistics: methods and applications*. 1st edn. Birkhäuser Basel.

Kulldorff, M. *et al.* (2006) 'An elliptic spatial scan statistic.', *Statistics in medicine*, 25, pp. 3929–3943.

Kulldorff, M., Fang, Z. and Walsh, S. J. (2003) 'A Tree-Based Scan Statistic for Database Disease Surveillance', *Biometrics*. Wiley Online Library, 59(2), pp. 323–331.

Kulldorff, M. and Nagarwalla, N. (1995) 'Spatial disease clusters: detection and inference.', *Statistics in medicine*, 14, pp. 799–810.

Kumar, S. *et al.* (2012) 'The social ecological model as a framework for determinants of 2009 H1N1 influenza vaccine uptake in the United States', *Health Education & Behavior*. Sage Publications Sage CA: Los Angeles, CA, 39(2), pp. 229–243.

Langmuir, A. D. (1976) 'William Farr: founder of modern concepts of surveillance', *International Journal of Epidemiology*. IEA, 5(1), pp. 13–18.

Langmuir, A. D. and others (1963) 'The surveillance of communicable diseases of

national importance.', *New England journal of medicine*, 268(4), pp. 182–192.

Larson, H. J. *et al.* (2014) 'Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007--2012', *Vaccine*. Elsevier, 32(19), pp. 2150–2159.

Lawson, A. B. and Kleinman, K. (2005) 'Spatial and Syndromic Surveillance for Public Health', *Spatial and Syndromic Surveillance for Public Health*, pp. 1–269.

Neill, D. B. and Moore, A. W. (2004) 'Rapid detection of significant spatial clusters', *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*, p. 256.

Thomson, A., Robinson, K. and Vallée-Tourangeau, G. (2016) 'The 5As: A practical taxonomy for the determinants of vaccine uptake', *Vaccine*. Elsevier, 34(8), pp. 1018–1024.

Tibshirani, R. (1996) 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR, pp. 267–288.

WHO Collaborating Centre for Drug Statistics Methodology (no date) *ATC Structure and principles*. Available athttps://www.whocc.no/atc/structure_and_principles/.

Yiannakoulias, N. (2011) 'Spatial aberration vs. geographical substance: Representing place in public health surveillance', *Health & Place*, 17(6), pp. 1242–1248.

BIBLIOGRAPHY

Abrams, A. M., Kleinman, K. and Kulldorff, M. (2010) 'Gumbel based p-value approximations for spatial scan statistics', *International Journal of Health Geographics*, 9(1), p. 61.

Acosta-Ramírez, N. *et al.* (2005) 'Determinants of vaccination after the Colombian health system reform', *Revista de Saude Publica*, 39(3), pp. 421–429.

Aday, L. A. and Andersen, R. (1974) 'A Framework for the Study of Access to Medical Care', *Health services research*, 9(3), pp. 208–220.

Alberta Health (2017) *2016/17 Influenza Surveillance Report*, *2016/17 Influenza Surveillance Report*. Available at http://www.health.alberta.ca/documents/Influenza-Surveillance-2017-01.pdf.

Alberta Ministry of Health (2017a) *Interactive Health Data Application (IHDA) - Demographics - Population Estimates - Adjusted*. Available at http://www.ahw.gov.ab.ca/IHDA_Retrieval/.

Alberta Ministry of Health (2017b) *Interactive Health Data Application (IHDA) - Immunization - Childhood Coverage Rates*. Available at http://www.ahw.gov.ab.ca/IHDA_Retrieval/.

Andersen, R. and Newman, J. F. (1973) 'Societal and Individual Determinants of Medical Care Utilization in the United States', *The Milbank Memorial Fund Quarterly. Health and Society*, 51(1), p. 95.

Angelillo, I. F. *et al.* (1999) 'Mothers and vaccination: knowledge, attitudes, and behaviour in Italy.', *Bulletin of the World Health Organization*, 77(3), pp. 224–229.

Assunção, R. *et al.* (2006) 'Fast detection of arbitrarily shaped disease clusters.', *Statistics in medicine*, 25, pp. 723–742.

Béland, Y. (2002) 'Canadian community health survey--methodological overview', *Health reports*. Statistics Canada, 13(3), p. 9.

Bell, C. A., Simmonds, K. A. and MacDonald, S. E. (2015) 'Exploring the heterogeneity among partially vaccinated children in a population-based cohort', *Vaccine*. Elsevier, 33(36), pp. 4572–4578.

Birkmeyer, J. D. (2001) 'Primer on Geographic Variation in Health Care', *Effective Clinical Practice*, 4(5), pp. 232–233.

Byrne, B. E. *et al.* (2013) 'Population-based cohort study comparing 30- and 90-day institutional mortality rates after colorectal surgery', *British Journal of Surgery*, 100(13), pp. 1810–1817.

de Cantuária Tauil, M., Sato, A. P. S. and Waldman, E. A. (2016) 'Factors associated with incomplete or delayed vaccination across countries: a systematic review', *Vaccine*. Elsevier, 34(24), pp. 2635–2643.

Center for Disease Control (CDC) (1986) 'Comprehensive plan for epidemiologic surveillance.' Atlanta , GA.

Center for Disease Control (CDC) (1990) 'Guidelines for investigating clusters of health events.', *MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports TA* -, 39(RR-11), pp. 1–23.

Center for Disease Control (CDC) (2011) 'Ten Great Public Health Achievements—United States , 2001-2010', *JAMA*, 306(1), pp. 36–38.

Choi, B. C. K. (2012) 'The past, present, and future of public health surveillance', *Scientifica*. Hindawi Publishing Corporation, 2012.

Choi, B. C. K. and Pak, A. W. P. (2001) 'Lessons for surveillance in the 21 st century: A historical perspective from the past five millennia', *Sozial-und Pr{ä}ventivmedizin/Social and Preventive Medicine*. Springer, 46(6), pp. 361–368.

Christakis, D. A. *et al.* (2000) 'The association between greater continuity of care and timely measles-mumps-rubella vaccination.', *American journal of public health*, 90(6), pp. 962–965.

Declich, S. and Carter, A. O. (1994) 'Public health surveillance: historical origins, methods and evaluation.', *Bulletin of the World Health Organization*. World Health Organization, 72(2), pp. 285–304.

Dombkowski, K. J., Lantz, P. M. and Freed, G. L. (2004) 'Risk factors for delay in age-appropriate vaccination.', *Public health reports*, 119(2), pp. 144–155.

Dover, D. C. and Schopflocher, D. P. (2011) 'Using funnel plots in public health surveillance', *Population Health Metrics*, 9(1), p. 58.

Dover, D. and Schopflocher, D. (2011) 'Using funnel plots in public health surveillance', *Population Health Metrics*, 9(1), p. 58.

Dubé, E. *et al.* (2013) 'Vaccine hesitancy: an overview', *Human vaccines & immunotherapeutics*. Taylor & Francis, 9(8), pp. 1763–1773.

Duczmal, L. *et al.* (2007) 'A genetic algorithm for irregularly shaped spatial scan statistics', *Computational Statistics and Data Analysis*, 52, pp. 43–52.

Duczmal, L. and Assuncao, R. (2004) 'A simulated annealing strategy for the detection of arbitrarily shaped spatial clusters', *Computational Statistics and Data Analysis*, 45, pp. 269–286.

Duczmal, L., Kulldorff, M. and Huang, L. (2006) 'Evaluation of Spatial Scan Statistics for Irregularly Shaped Clusters', *Journal of Computational and Graphical Statistics*, pp. 428–442.

Falagas, M. E. and Zarkadoulia, E. (2008) 'Factors associated with suboptimal compliance to vaccinations in children in developed countries: a systematic review', *Current medical research and opinion*. Taylor & Francis, 24(6), pp. 1719–1741.

Gail M, Benichou J, editors. (2000) *Encyclopedia of Epidemiologic Methods*. Wiley, pp.872

Glaz, J., Pozdnyakov, V. and Wallenstein, S. (2009) *Scan statistics: methods and applications*. 1st edn. Birkhäuser Basel.

Goh, P. *et al.* (2007) 'Safety and immunogenicity of early vaccination with two doses of tetravalent Measles-Mumps-Rubella-Varicella (MMRV) vaccine in healthy children from 9 months of age', *Infection*, 35(5), pp. 326–333.

Goujon-Bellec, S. *et al.* (2011) 'Detection of clusters of a rare disease over a large territory: performance of cluster detection methods', *International Journal of Health Geographics*, p. 53.

Government of Alberta (2011) *Guidelines for the Investigation of Clusters of Non-Communicable Health Events*. Edmonton, AB. Available at http://www.health.alberta.ca/documents/Investigation-Clusters-Guidelines-2011.pdf.

Gower, J. C. and Ross, G. J. S. (1969) 'Minimum Spanning Trees and Single Linkage Cluster Analysis', *Applied Statistics*. WileyRoyal Statistical Society, 18(1), p. 54.

Groseclose, S. L. and Buckeridge, D. L. (2017) 'Public Health Surveillance Systems: Recent Advances in Their Use and Evaluation', *Annual Review of Public Health*. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, California 94303-0139, USA, 38(0), pp. 57–79.

Hamel, M. and Laniel, N. (2014) 'Producing official statistics via voluntary surveys--the National Household Survey in Canada', *Statistical Journal of the IAOS*. IOS Press, 30(3), pp. 237–242.

Institute of Medicine (2013) *The Childhood Immunization Schedule and Safety: Stakeholder Concerns, Scientific Evidence, and Future Studies*. Washington, D.C.: National Academies Press.

Iyengar, V. S. (2005) 'Space-time clusters with flexible shapes', *MMWR Suppl*, 54, pp. 71–76.

Jung, I. (2009) 'A generalized linear models approach to spatial scan statistics for covariate adjustment', *Statistics in Medicine*. John Wiley & Sons, Ltd., 28(7), pp. 1131–1143.

Kim, S. S. *et al.* (2007) 'Effects of maternal and provider characteristics on up-to-date

immunization status of children aged 19 to 35 months.', *American journal of public health*, 97(2), pp. 259–266.

Kirkham, J. J. and Bouamra, O. (2008) 'The use of statistical process control for monitoring institutional performance in trauma care', *Journal of Trauma-Injury Infection & Critical Care*, 65(6), pp. 1494–1501.

Kulldorff, M. (1997) 'A spatial scan statistic', *Communications in Statistics - Theory and Methods*, 26(6), pp. 1481–1496.

Kulldorff, M. *et al.* (1997) 'Breast cancer clusters in the northeast United States: a geographic analysis', *American journal of epidemiology*. Oxford Univ Press, 146(2), pp. 161–170.

Kulldorff, M. *et al.* (2006) 'An elliptic spatial scan statistic.', *Statistics in medicine*, 25, pp. 3929–3943.

Kulldorff, M., Fang, Z. and Walsh, S. J. (2003) 'A Tree-Based Scan Statistic for Database Disease Surveillance', *Biometrics*. Wiley Online Library, 59(2), pp. 323–331.

Kulldorff, M. and Nagarwalla, N. (1995) 'Spatial disease clusters: detection and inference.', *Statistics in medicine*, 14, pp. 799–810.

Kumar, S. *et al.* (2012) 'The social ecological model as a framework for determinants of 2009 H1N1 influenza vaccine uptake in the United States', *Health Education & Behavior*. Sage Publications Sage CA: Los Angeles, CA, 39(2), pp. 229–243.

Kunadian, B. *et al.* (2009) 'Funnel plots for comparing performance of PCI performing hospitals and cardiologists: demonstration of utility using the New York hospital mortality data', *Catheterization and Cardiovascular Interventions*. Wiley Online Library, 73(5), pp. 589–594.

Langmuir, A. D. (1976) 'William Farr: founder of modern concepts of surveillance', *International Journal of Epidemiology*. IEA, 5(1), pp. 13–18.

Langmuir, A. D. and others (1963) 'The surveillance of communicable diseases of

national importance.', *New England journal of medicine*, 268(4), pp. 182–192.

Larson, H. J. *et al.* (2014) 'Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007--2012', *Vaccine*. Elsevier, 32(19), pp. 2150–2159.

Last, J. M. (1998) 'A dictionary of epidemiology, third edition', *American Journal of Preventive Medicine*, p. 242.

Lawson, A. B. and Kleinman, K. (2005) 'Spatial and Syndromic Surveillance for Public Health', *Spatial and Syndromic Surveillance for Public Health*, pp. 1–269.

Lee, L. *et al.* (2010) *Principles and Practice of Public Health Surveillance*. 3rd edn. Oxford University Press, USA.

Lemstra, M. *et al.* (2007) 'Disparity in childhood immunizations.', *Paediatrics & child health*, 12(10), pp. 847–852.

Li, X. Z. *et al.* (2011) 'A spatial scan statistic for multiple clusters', *Mathematical Biosciences*, 233(2), pp. 135–142.

Lieu, T. *et al.* (2015) 'Geographic clusters in underimmunization and vaccine refusal.', *Pediatrics TA* -, 135(2), pp. 280–289.

Light, R. J. and Pillemer, D. B. (1984) *Summing up: the science of reviewing research*, *Summing up: the science of reviewing research*. Cambridge, Mass., Harvard University Press.

MacDonald, S. E. *et al.* (2014) 'Parental concern about vaccine safety in Canadian children partially immunized at age 2: A multivariable model including system level factors', *Human Vaccines & Immunotherapeutics,* 10(9), pp.2603–2611.

Marshall, C. E. and Spiegelhalter, D. J. (1998) 'Reliability of league tables of in vitro fertilisation clinics: retrospective analysis of live birth rates', *BMJ*, 316, pp. 1701–1705.

Marshall, T., Mohamnmed, M. A. and Rouse, A. (2004) 'A randomized controlled trial of league tables and control charts as aids to health service decision-making', *Int J Qual*

Health Care, 16(4), pp. 309–315.

Massat, N. J. *et al.* (2015) 'Variation in cervical and breast cancer screening coverage in England: a cross-sectional analysis to characterise districts with atypical behaviour', *BMJ Open*, 5(7), p. e007735.

Matsumura, T. *et al.* (2005) 'Measles vaccine coverage and factors related to uncompleted vaccination among 18-month-old and 36-month-old children in Kyoto, Japan.', *BMC public health*, 5, p. 59.

Mazzucco, W. *et al.* (2017) 'Funnel plots and choropleth maps in cancer risk communication: a comparison of tools for disseminating population-based incidence data to stakeholders', *BMJ open*. British Medical Journal Publishing Group, 7(3), p. e011502.

Mills, E. *et al.* (2005) 'Systematic review of qualitative studies exploring parental beliefs and attitudes toward childhood vaccination identifies common barriers to vaccination', *Journal of clinical epidemiology*. Elsevier, 58(11), pp. 1081–1088.

Neill, D. B. (2012) 'Fast subset scan for spatial pattern detection', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74, pp. 337–360.

Neill, D. B. and Moore, A. W. (2004) 'Rapid detection of significant spatial clusters', *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*, p. 256.

Neutra, R. R. (1990) 'Counterpoint from a cluster buster.', *American journal of epidemiology*, 132(1), pp. 1–8.

Nur, U. *et al.* (2015) 'Inequalities in non-small cell lung cancer treatment and mortality', *Journal of Epidemiology and Community Health*, 69(10), pp. 985–992.

Ohlssen, D. I., Sharples, L. D. and Spiegelhalter, D. J. (2007) 'A hierarchical modelling framework for identifying unusual performance in health care providers', *J R Statist Soc A*, 170, pp. 865–890.

Omer, S. B. *et al.* (2008) 'Geographic clustering of nonmedical exemptions to school

immunization requirements and associations with geographic clustering of pertussis', *American Journal of Epidemiology*. Oxford Univ Press, 168(12), pp. 1389–1396.

Patil, G. and Taillie, C. (2004) 'Upper level set scan statistic for detecting arbitrarily shaped hotspots', *Environmental and Ecological statistics*, 11, pp. 189–197.

Pavlopoulou, I. D. *et al.* (2013) 'Immunization coverage and predictive factors for complete and age-appropriate vaccination among preschoolers in Athens, Greece: a cross--sectional study.', *BMC public health*. BMC Public Health, 13(1), p. 908.

Pearce, A. *et al.* (2008a) 'Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study.', *BMJ (Clinical research ed.)*, 336(7647), pp. 754–7.

Pearce, A. *et al.* (2008b) 'Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study', *bmj*. British Medical Journal Publishing Group, 336(7647), pp. 754–757.

Pearce, N. (2000) 'The ecological fallacy strikes back.', *Journal of epidemiology and community health*. BMJ Publishing Group, 54(5), pp. 326–327.

Pearson, M. *et al.* (1993) 'Primary immunisations in Liverpool. II: Is there a gap between consent and completion?', *Archives of disease in childhood*, 69(1), pp. 115–119.

Phillips, K. A. *et al.* (1998) 'Understanding the context of healthcare utilization: assessing environmental and provider-related variables in the behavioral model of utilization.', *Health services research*. Health Research & Educational Trust, 33(3 Pt 1), p. 571.

Poirier, D. J. (1995) *Intermediate Statistics and Econometrics: A Comparative Approach*. The MIT Press.

Que, J. and Tsui, F.-C. (2011) 'Rank-based spatial clustering: an algorithm for rapid outbreak detection.', *Journal of the American Medical Informatics Association : JAMIA*, 18, pp. 218–224.

Rogerson, P. and Yamada, I. (2008) *Statistical Detection and Surveillance of Geographic Clusters*, *Statistical Detection and Surveillance of Geographic Clusters*. Boca Raton, FL, Chapman and Hall.

Rothman, K. J. (2002) *Epidemiology: An Introduction*. 1st edn. Oxford University Press, USA.

Sahni, V., Lai, F. Y. and MacDonald, S. E. (2014) 'Neonatal vitamin K refusal and nonimmunization', *Pediatrics*. Am Acad Pediatrics, pp. 497-503.

De Serres, G. *et al.* (2013) 'Largest measles epidemic in North America in a decade-Quebec, Canada, 2011: Contribution of susceptibility, serendipity, and superspreading events', *Journal of Infectious Diseases*, 207(6), pp. 990–998.

Shen, W. and Louis, T. A. (2000) 'Triple-goal estimates for disease mapping', *Statistics in Medicine*, 19, pp. 2295–2308.

Smith, P. J. and Singleton, J. A. (2008) 'Vaccination Coverage Estimates for Selected Counties: Achievement of Healthy People 2010 Goals and Association with Indices of Access to Care, Economic Conditions, and Demographic Composition', *Public health reports*. Association of Schools of Public Health, 123(2), pp. 155–172.

Speakman, S., McFowland, E. and Neill, D. B. (2015) 'Scalable Detection of Anomalous Patterns With Connectivity Constraints', *Journal of Computational and Graphical Statistics*, 24(4), pp. 1014–1033.

Spiegelhalter, D. J. (2005a) 'Funnel plots for comparing institutional performance', *Statistics in Medicine*, 24, pp. 1185–2102.

Spiegelhalter, D. J. (2005) 'Funnel plots for comparing institutional performance', *Statistics in Medicine*, 24(8), pp. 1185–1202.

Spiegelhalter, D. J. (2005b) 'Handling over-dispersion of performance indicators', *Qual Saf Health Care*, 14, pp. 347–351.

Sterne, J. A. C., Egger, M. and Smith, G. D. (2001) 'Investigating and dealing with

publication and other biases in meta-analysis', *BMJ*, 323, pp. 101–105.

Suarez, L., Simpson, D. M. and Smith, D. R. (1997) 'The impact of public assistance factors on the immunization levels of children younger than 2 years.', *American journal of public health*, 87(5), pp. 845–848.

Sui, D. Z. and Holt, J. B. (2008) 'Visualizing and Analysing Public-Health Data Using Value-by-Area Cartograms: Toward a New Synthetic Framework', *Cartographica*, 43(1), pp. 3–20.

Suttorp, V. (2014) 'Outbreaks in Rural Communities with Low Immunization Rates', *Western Canadian Immunization Forum*. Edmonton, AB.

Tango, T. (2008) 'A spatial scan statistic with a restricted likelihood ratio', *Japanese Journal of Biometrics*, 29(2), pp. 75–95.

Tango, T. and Takahashi, K. (2005) 'A flexibly shaped spatial scan statistic for detecting clusters', *International Journal of Health Geographics*, 4(1), p. 11.

Tango, T. and Takahashi, K. (2012) 'A flexible spatial scan statistic with a restricted likelihood ratio for detecting disease clusters.', *Statistics in medicine*, pp. 4207–18.

Thomson, A., Robinson, K. and Vallée-Tourangeau, G. (2016) 'The 5As: A practical taxonomy for the determinants of vaccine uptake', *Vaccine*. Elsevier, 34(8), pp. 1018–1024.

Tibshirani, R. (1996) 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR, pp. 267–288.

Trauth, J. M. *et al.* (2002) 'Do beliefs of inner-city parents about disease and vaccine risks affect immunization?', *Journal of the National Medical Association*, 94(9), pp. 820–832.

Tunstall, H. V. Z., Shaw, M. and Dorling, D. (2004) 'Places and health.', *Journal of epidemiology and community health*, 58(1), pp. 6–10.

Wacholder, S. (1986) 'Binomial regression in GLIM: estimating risk ratios and risk

differences.', *American journal of epidemiology*, 123(1), pp. 174–84.

Wakefield, J. and Kim, A. (2013) 'A Bayesian model for cluster detection.', *Biostatistics (Oxford, England)*, 14, pp. 752–65.

van der Wal, M. F. *et al.* (2005) 'Vaccination rates in a multicultural population.', *Archives of disease in childhood*, 90(1), pp. 36–40.

Waller, L. A. and Gotway, C. A. (2004a) 'Applied Spatial Statistics for Public Health Data', *Environmental Health*, 100, pp. 702–703.

Waller, L. A. and Gotway, C. A. (2004b) *Applied Spatial Statistics for Public Health Data*, *Environmental Health*.

Walter, S. D. (2000) 'Disease mapping: a historical perspective', in *Spatial epidemiology: methods and applications*, pp. 223–239.

WHO Collaborating Centre for Drug Statistics Methodology (no date) *ATC Structure and principles*. Available at https://www.whocc.no/atc/structure_and_principles/.

Woodall, D. H. (2006) 'The Use of Control Charts in Health-Care and Public-Health Surveillance', *J Qual Technol*, 38(2), pp. 89–104.

World Health Organization (WHO) (2014) *Report of the SAGE working group on vaccine hesitancy*. Available at http://www.who.int/immunization/sage/meetings/2014/october/1_Report_WORKING_GROUP_vaccine_hesitancy_final.pdf.

Woudenberg, T. *et al.* (2017) 'Large measles epidemic in the Netherlands, May 2013 to March 2014: changing epidemiology', *Eurosurveillance*. European Centre for Disease Prevention and Control, 22(3).

Yiannakoulias, N. (2011) 'Spatial aberration vs. geographical substance: Representing place in public health surveillance', *Health & Place*, 17(6), pp. 1242–1248.

Yiannakoulias, N., Rosychuk, R. J. and Hodgson, J. (2007) 'Adaptations for finding irregularly shaped disease clusters.', *International journal of health geographics*, 6, p. 28.

Zhang, Z., Assuncao, R. and Kulldorff, M. (2010) ‘Spatial scan statistics adjusted for multiple clusters’, *Journal of Probability and Statistics*, 2010, pp. 1–11.

Zhao, Z., Mokdad, A. H. and Barker, L. (2004) ‘Impact of health insurance status on vaccination coverage in children 19-35 months old, United States, 1993-1996.’, *Public health reports*, 119(2), pp. 156–162.