Statistical Modeling of Pneumonia Transmission Rates in Manitoba

by

Md. Hasan

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Community Health Sciences

University of Manitoba

Winnipeg

Copyright © 2023 by Md. Hasan

Abstract

Introduction: Pneumonia is a major respiratory infection that significantly strains Manitoba's healthcare system, resulting a substantial number of hospitalizations and fatalities. Understanding the transmission dynamics and risk factors associated with pneumonia in this population is crucial for targeted interventions. Spatial variability in pneumonia hospitalizations has been observed in Manitoba, where various risk factors contribute to pneumonia infection. Therefore, it is essential to investigate the determinants of pneumonia, including spatial aspects and disease transmission rates.

Methods: We applied a spatial Poisson regression model that incorporates the Intrinsic Conditional Autoregressive model to explore region level potential risk factors. To understand the influence of comorbidity, we utilized a mixed-effects Poisson regression model. Our analysis focused on hospital data spanning the years 2015 to 2019, and encompassing 96 Manitoba health regions. Moreover, to investigate disease transition rates, we employed both the Susceptible-Exposed-Infected-Recovery (SEIR) model (excluding reinfection) and the Susceptible-Exposed-Infected-Recovery-Susceptible (SEIRS) model (including reinfection). These compartmental models considered data from both hospital and physician-reported pneumonia cases during August 2017 to July 2018.

Results: The raw incidence rate of pneumonia infection exhibited significant variation across different regions, ranging from 2 to 55 cases per 1000 population. After adjusting for potential risk factors, the incidence rate ratios ranged from 0.38 to 6.63, indicating substantial variations in incidence rates among regions. Factors such as age, immigration status, and comorbidities, notably Chronic Obstructive Pulmonary Disease (COPD) and Cardiovascular Disease (CVD), were identified as significant contributors to the risk of pneumonia infection. Conversely, vaccination was found to exert a protective effect, especially among individuals aged over 60 years. The domain-level analysis further demonstrated the significant impact of Inflammatory Bowel Diseases (IBD), COPD and CVD on pneumonia infection. The SEIR and SEIRS models were employed to analyze pneumonia transmission rates, indicated rates of 0.81 and 0.88, respectively. The SEIRS model suggested a reinfection rate of pneumonia was 0.003. The average reproduction number (R_0) was calculated at 1.03 for both models, signifying the potential for disease spread in the population.

Conclusion: The findings highlighted spatial variation in pneumonia incidence across the 96 health regions in Manitoba. Elder individuals, immigration status, and comorbidities, were identified as significant factors influencing pneumonia infections rates. Vaccination demonstrated a protective effect, particularly among the elderly population. The estimated reinfection interval for pneumonia was determined to be 333 days.

Acknowledgments

I would like to extend my heartfelt appreciation to my supervisor, Dr. Mahmoud Torabi, and my co-supervisor, Dr. Zeinab Mashreghi. Their unwavering guidance, profound expertise, and steadfast support have been the cornerstones of this research's success.

I am also profoundly grateful to my esteemed committee members, Dr. Harminder Singh and Dr. Barbara Nery Porto, for their invaluable feedback and illuminating suggestions, which significantly enriched the quality of this thesis.

I wish to express my deep gratitude and immense pleasure in acknowledging the financial support extended to me through the University of Manitoba Graduate Fellowship (UMGF), as well as the research grant generously provided by Dr. Mahmoud Torabi and Dr. Zeinab Mashreghi. Their contributions have been instrumental in making this research endeavor a reality.

Lastly, I am indebted to the unwavering support of my family members specially, my wife and two small kids who have been my pillars of strength throughout this journey.

Dedication

I would like to dedicate this work to my wife, who supported me throughout my academic journey.

Contents

List of Tables

List of Figures

Chapter 1

Background and Introduction

1.1 Introduction

Pneumonia is a lung illness caused by an acute respiratory infection. When a person develops pneumonia, the alveoli become clogged with pus and fluid, making breathing difficult and limiting oxygen intake (WHO, 2019). This is caused primarily by a variety of infectious organisms such as bacteria, viruses, and fungi (WHO, 2019). In Canada, infectious diseases such as pneumonia account for about $70,000$ hospitalizations each year, imposing a big burden on the health care system (CIHI, 2021). Manitoba is a Canadian province with a population of approximately 1.4 million (Statistics Canada, 2022). The province has a well-developed healthcare system that provides comprehensive healthcare services to its residents. Pneumonia is a leading cause of hospitalization and death, with around 3000 people admitted each year in Manitoba (CIHI, 2021), particularly among older adults and those with underlying health conditions (Statistics Canada, 2022). To reduce this burden, it is important to better understand the transmission dynamics of pneumonia in this population for developing targeted interventions. Compartmental models have widely been used to study the transmission dynamics of infectious diseases, including pneumonia (Ong'ala et al., 2015; Ndelwa et al., 2015; Tilahun et al., 2017). These models divide the population into compartments based on their infection status, allowing for a more detailed analysis of transmission rates and disease spread.

By applying a compartmental modeling techniques, we aim to improve our understanding of the epidemiology of pneumonia transmission rates in Manitoba and inform public health interventions to reduce its impact.

The burden of disease may not be equally distributed across space and time. Previous research using a time series method in Finland indicated that pneumonia cases were extremely seasonal, with high numbers in the winter and low numbers in the summer (Saynajakangas et al., 2001). Spatial variability in pneumonia hospitalizations has widely been reported (Kim et al., 1996; Crighton et al., 2008; Blain et al., 2014; Beck et al., 2015; Iroh et al., 2017; Kassam et al., 2021). Furthermore, there are numerous risk factors such as environmental, cultural, ecological, and behavioural factors that contribute significantly to the development of pneumonia. Some chronic conditions such as asthma, inflammatory bowel disease (IBD), chronic obstructive pulmonary disease (COPD), and cardiovascular disease (CVD) can compromise the immune system and increase the susceptibility to pneumonia infection. On the other hand, the pneumonia vaccine has been found to play a protective role in preventing pneumonia, particularly in children and older adults. In the course of literature review, we discovered a paucity of articles that concentrate on the spatial variation of pneumonia hospitalization in Manitoba, as well as the identification of its determinants. Therefore, it is scientific interest to identify the determinants of pneumonia while considering spatial factors. A retrospective cohort design has been used to examine the pneumonia transmission and its determinants in the province of Manitoba, Canada.

1.2 Literature Review

According to the World Health Organization (WHO), the most common cause of death among children was pneumonia (WHO, 2019). In 2017, almost $808,000$ children under five years old died from this disease, which accounts for 15% of all annual deaths of this age group globally (WHO, 2019). A total of 1.4 million Americans visited emergency hospital department for ambulatory pneumonia care in 2021 (Cairns and Kang, 2021). In the United States and Canada, pneumonia and influenza were the eighth and seventh-positioned causes of mortality, respectively (File Jr and

Marrie, 2010). A list of factors contributing to pneumonia morbidity and mortality has already been notified. Gender, chronic illnesses including asthma, CVD, and antibiotic use were considered risk factors for pneumonia (Lange et al., 1995; Plouffe et al., 1996; Kaplan et al., 2003). Elderly people were identified as being the most likely to acquire illness and experience the worst outcomes as a result of their advanced ages (Loeb, 2003). Smoking, obesity, excessive alcohol drinking, and lifestyle factors were also associated with pneumonia (Nuorti et al., 2000; Baik et al., 2000; Almirall et al., 2000).

The burden of this disease has been studied in different ways. Mathematical models have been used to measure the burden and transmission rate for respiratory infectious disease such as severe acute respiratory syndrome (Lipsitch et al., 2003), tuberculosis (Kumar et al., 2013) and pneumonia. Tilahun et al. (2017) introduced and examined a non-linear mathematical model for cost-effective strategies in pneumonia disease transmission dynamics. The study employed a deterministic susceptible, vaccinated, carrier, infected and recovered compartmental model and differential equation stability theory. To investigate the transmission dynamics of pneumonia and the impacts of screening and treatment interventions, Ndelwa et al. (2015) developed and studied a mathematical model considering the dynamics of pneumonia infection among four sub-populations, namely susceptible, asymptomatic infectives (or simply carriers), symptomatic infectives, and treated infectives. In Ong'ala et al. (2015), a mathematical model was developed to study the transmission dynamics of pneumonia in children under the age of five . They categorized the disease status into four compartments, namely susceptible, carriers, infectious and recovered. They used the theory of ordinary differential equations, dynamical systems, and the possibility of bifurcation to analyze the model. In this study, streptococcus pneumonia was only considered since it is the most common among children under five years of age. Kizito and Tumwiine (2018) investigated a mathematical model to understand the spread and control of bacterial pneumonia under public health interventions, including treatment and vaccination. They analyzed the effect of these interventions on the dynamics of pneumonia by conducting sensitivity analysis on the effective reproduction number. They used numerical simulations to illustrate the analytical results and determine the long-term behavior of the disease. The results showed that while treatment intervention alone may not eliminate pneumonia, a combination of treatment and vaccination interventions can effectively wipe

out the disease. Temime et al. (2004) constructed a mathematical model of selection for resistance to penicillin G in streptococcus pneumonia in an age-structured population, incorporating vaccination to understand the epidemiological characteristics of the disease in a vaccinated population. The model included three age groups, and the results suggested that the effects of vaccination may not be sustained in the long term due to serotype replacement. In a study by Sutton et al. (2010), they developed a partial differential equation model for pneumococcal infection with vaccination, focusing on evaluating the impact of vaccines at the population level and estimating parameters using data.

The spatio-temporal association with pneumonia incidence was also reported in different literature. In the United States, all hospital admission records data showed regional elevation, particularly in southern and northern zones (File Jr and Marrie, 2010). A spatial analysis conducted in Alberta, Canada, reported that the highest hospitalization rates were found in the northern part compared to the southern part, and in the remote regions compared to its urban counterparts (Kassam et al., 2021). In Ontario, Canada, spatio-temporal interaction was found in several studies, with marked spatial variability in temporal patterns, and temporal variability in spatial patterns (Crighton et al., 2007a,b, 2008). A similar variation in hospitalization rate over the region was observed in England, Minnesota (USA), and Korea (Blain et al., 2014; Iroh et al., 2017; Kim et al., 2019). Studies also showed that there was a clear seasonality with consistent summer troughs and winter peaks for both sexes and all ages combined (Crighton et al., 2004; Benincà et al., 2017).

In Manitoba, total respiratory morbidity showed statistically significant variation among the Manitoba regions (Torabi and Jiang, 2020). According to the Manitoba health annual statistics, pneumonia has been playing the most significant role in hospitalization among all ambulatory care-sensitive conditions. From 2013 to 2019, almost 16 to 18 percent of patients were admitted to the hospital due to pneumonia (Manitoba Health, 2014, 2015, 2016, 2017, 2018, 2019). In $2005-2006$, the Manitoba Center for Health Policy (MCHP) showed that the hospitalization rate varied over the Manitoba health regions, the highest percentage was found in Assiniboine residents while the lowest was found in Winnipeg and Parkland residents (Finlayson et al., 2009).

Since the impact of pneumonia infection is severe on a domain level, finding the spatial

variation could help to design a new epidemiological study that may help to identify hidden risk factors and trends of these factors. To monitor geographic regional trends over time, public health agencies generally use disease rates (ratios) maps. These agencies require reliable maps based on sound methodology. The fields of statistics and epidemiology have paid considerable attention to this area, as maps depicting morbidity and mortality rates for evolution that can provide sufficient information to researchers about spatio-temporal patterns of diseases. Disease rates in different geographical regions may indicate that there exist region-specific factors, which could explain the geographic pattern. It might be possible to isolate the causal factors behind these differences.

Poisson regression model is commonly used to analyze disease rates (ratios) by assuming that neighbouring regions are independent and that variances are equal in each region. However, the model may produce extra-Poisson variations in the absence of unidentified causal factors. We may need to incorporate some spatial correlation into this model, depending on the degree of smoothness of the omitted factors in each region. A spatially correlated random effects mixed model was first developed by Clayton and Kaldor to accommodate extra-Poisson variation (Clayton and Kaldor, 1987).

Statistical inference can be drawn in different ways from Generalized Linear Mixed Models (GLMMs). Markov chain Monte Carlo (MCMC) method via the Bayesian approach may be used through Gibbs's sampler or the Metropolis–Hasting's algorithm, though the monitoring of the convergence of this model is difficult (Bernardinelli and Montomoli, 1992; Elliott et al., 1996; Waller et al., 1997; Torabi and Rosychuk, 2012). The Penalized Quasi-Likelihood (PQL) method is also a useful technique for estimating the parameter of GLMM. Breslow and Clayton (1993) suggested this approach, and illustrated how to utilize the PQL for an estimate in mapping studies. For longitudinal data analysis, one may use the Generalized Estimating Equation (GEE) approach for drawing inferences from the Generalized Linear Model (GLM) (Liang and Zeger, 1986; Prentice and Zhao, 1991).

The Intrinsic Conditional Autoregressive (ICAR) model proposed by Besag et al. (1991) is a popular statistical tool for spatial analysis, which allows us to analyze spatial data where observations are not independent but exhibit spatial correlation or dependence. The ICAR model is

based on the Conditional Autoregressive (CAR) model that models the spatial dependence between neighboring regions in a spatial data set. For smoothing the temporal rates of count data, an Autoregressive (AR) model was developed by Zeger (1988). A priori hierarchical Bayes spatial model was extended by Waller et al. (1997) and Knorr-Held (2000) to incorporate time-related effects and spatio-temporal interactions. MacNab and Dean (2001) and Silva et al. (2008) proposed spatio-temporal models that smooth spatial and temporal dimensions with AR local smoothing and cubic B-spline smoothing, respectively. Based on the PQL estimation techniques, Torabi and Rosychuk (2011) also developed a new spatio-temporal model using B-spline for disease mapping. In Torabi and Rosychuk (2010), a method was established for dealing with the seasonal impacts of spatio-temporal illness rate modeling. The GEE method was used to estimate the model parameters.

Therefore, strong methodologies have been developed for analyzing the spatial variation of pneumonia, its determinants as well as understanding the disease dynamics. However, there have been few applications in the field of infectious diseases, particularly in tracking pneumonia trends in Manitoba. To fill this gap, this research first aims to measure spatial variation as well as the determinants of pneumonia in Manitoba at the regional level. For this purpose, we have utilized a spatial Poisson regression model, in which the spatial variability was measured using an ICAR model, incorporating 96 Manitoba health regions. We also investigated the impact of socioeconomic factors and comorbidity on pneumonia infection at domain level. Secondly, to predict disease transmission dynamics, compartmental models have been employed. In this research, we utilized the Susceptible-Exposed-Infected-Recovered (SEIR) model as well as the Susceptible-Exposed-Infected-Recovered-Susceptible (SEIRS) model to investigate pneumonia transmission rates in Manitoba. These models were applied to assess scenarios without and with reinfection cases.

1.3 Research Questions and Hypotheses

Based on the literature review, we decided to explore three research questions in my thesis which can be expressed as:

1. What are potential risk factors of pneumonia infection in Manitoba?

2. What is the impact of comorbidity on pneumonia infection in Manitoba?

3. What is the transmission rate of pneumonia in Manitoba?

Chapter 2

Materials and Methods

2.1 Research Design

This study employed a quantitative research approach utilizing a retrospective cohort study design. Hospital admission and physician visit records were utilized, and all exposure information was assessed retrospectively. Study participants were selected based on the criterion of residing in Manitoba. A retrospective cohort design was a preferred alternative when a randomized control design was not feasible.

2.2 Study Area

This study was conducted in the province of Manitoba. According to the last published census 2021 report, 1,342,150 people were living in Manitoba (Statistics Canada, 2021b). The male and female ratio were almost equal, 675,660 (50.34%) identified as female and 666,490 (49.66%) as male (Statistics Canada, 2021b). A total of 55.58% population lived in the capital city, Winnipeg. This province is culturally diverse with a higher proportion of Indigenous population compared with other provinces (Coates et al., 2023). Manitoba was divided into 5 Regional Health Authorities

(RHAs) after merging the 11 RHAs which took place in 2012 (MCHP, 2013). The health regions were further divided into sub-regions called Regional Health Authority Districts (RHADs). The RHADs were considered as a single-level aggregated spatial units in this study. In this thesis, we used 96 RHADs in Manitoba for the data analysis. *QGIS* (version 3.26.3) and *R* software were used to create choropleth maps that were displayed in the Results section. All maps presented in this study were created from the shape files which form a part of the data set used in the analysis and no copyright permission was required.

2.3 Data Extraction and Management

2.3.1 Data Extraction Process

We extracted our required data file from the Population Health Research Data Repository (PHRDR) that is established at the MCHP housed at the University of Manitoba. The PHRDR is a data repository hub containing all refined virtual records of the provincial health care system, including physician records, hospital records, home care records, and prescriptions for medications from all registered patients (MCHP, 2023). It is a comprehensive repository of historical information about Manitoba residents, including administrative, registry, survey, and other data. To use this data, approval from MCHP has been taken following proper procedure. A confidentiality agreement has been signed between MCHP and the principal investigator of this research, in order to protect the data and ensure all team members follow the approved protocol. The data files have been prepared and analyzed within the secure system of the MCHP web portal.

For this study, we extracted the "hospital abstract" data files that contain patient descriptions (e.g., date of admission, diagnosis, duration of hospital stay, etc.) with the residential postal code. The physician visit record file utilized for obtaining pneumonia diagnoses by physicians. The "Insurance Registry" data file used to extract the patient's age and sex history. The vaccination history of the respondent has been extracted from Manitoba Immunization Monitoring System record database. Socioeconomic Factor Index version 2 (SEFI-2) and the immigrant status of the

population data were extracted from the 2021 Canadian census.

2.3.2 Data Preparation Process

The data have been processed and organized to align with the specific requirements of each objective. To investigate the spatial variability of pneumonia incidence and risk factors at regional level, we included data for all patients admitted to health facilities due to pneumonia as well as selected comorbidities between January $1, 2015$, and December $31, 2019$. For regional-level modeling, we structured the data according to 96 RHAD levels. This process entailed merging municipality codes and postal codes to create unique identifiers known as RHAD codes, enabling the effective integration of various datasets. To assess the impact of comorbidity on pneumonia infection, we restricted our analysis to patients who were admitted with pneumonia during the same time period and organized the data at the domain level. In this context, we considered age-sex grouping and age-specific data to tailor it for different modeling approaches. We defined a population cohort based on individuals' residency status between 2015 and 2019. Specifically, individuals who resided in Manitoba for one or more days during this period were categorized as Manitoba residents and were included in the cohort population.

To delve into the dynamics of disease transmission, we focused on a one-year period, from August $1, 2017$, to July $31, 2018$. This period encompassed pneumonia cases admitted to hospitals and those diagnosed by physicians. Upon analyzing data spanning the past 20 years, we noted recurrent peaks in pneumonia infections during winter seasons. The most recent peak was observed in the winter of $2017 - 18$. This observation led to our selection of a one-year data set starting in mid-2017. This dataset was structured on a daily basis, with each row representing a single day and containing the number of active pneumonia cases for that day. The active pneumonia case was defined based on a recovery period of 7 days.

2.4 Outcome of Interest

The outcome of interest, pneumonia cases, was defined using the international classification of disease version 10 (ICD-10 CM) code (WHO, 2019). Table 2.1 shows the details of the pneumonia cases definition that was considered for this study.

Table 2.1: Description of all pneumonia cases with sub-types using ICD-10 CM

ICD-10 CM	Sub-types
J12	Viral pneumonia, not elsewhere classified
J13	Pneumonia due to Streptococcus pneumonia
J14	Pneumonia due to Hemophilus influenza
J15	Bacterial pneumonia, not elsewhere classified
J16	Pneumonia due to other infectious organisms, not elsewhere classified
J17	Pneumonia in diseases classified elsewhere
J18	Pneumonia, unspecified organism

We selected those patients who were admitted to the hospital with pneumonia as their most responsible diagnosis (MRDx) (Crighton et al., 2007a, 2008).

Reinfection pneumonia was formally defined as the occurrence of discrete episodes of pneumonia in individuals where the symptoms of a subsequent episode manifest at least 30 days after the resolution of their previous pneumonia episode. This is also called the early recurrent pneumonia. For the purpose of classification, individuals who seek medical attention for pneumonia more than once within a 30-day period were considered to have experienced a single pneumonia episode (Winterbauer et al., 1969; Dang et al., 2014, 2015).

2.5 Independent/Exposure Variables

For this study, we considered age and sex of the respondents, pneumonia vaccination coverage, immigration status, the SEFI-2 patients and comorbidity including Asthma, IBD, COPD, and CVD as independent covariates.

Age of respondents is an important factor as the distribution of pneumonia is heterogeneous over the age groups, with younger and older people being more vulnerable than young adults (Government of Manitoba, 2022). We categorized age into three groups for this analysis: 0 to 5 years, 6 to 59 years, 60 years and above.

The distribution of the immigrant population across the province is also an important factor that was considered in our analysis to measure the geographical variation of pneumonia. There may be a potential link between the spatial distribution of immigrants and the corresponding variation in the incidence of pneumonia across the province. In this study, we used the distinct percentages of the population who identified themselves as immigrants during the study period in Manitoba, as defined by the 2021 Canadian census (Statistics Canada, 2021a).

The SEFI-2 score was derived through a linear combination of various factors such as the unemployment rate of age 15+, average household income of age 15+, the proportion of single-parent households, and the proportion of the population age 15+ without high school graduation. This was achieved using factor analysis as described in detail in the MCHP protocol (Metge et al., 2009). The range of this score based on our data was between -1.53 to 2.96. The scores less than zero indicate more favourable socioeconomic conditions, while scores greater than zero indicate less ideal socioeconomic conditions.

IBD is a chronic inflammatory condition that primarily affects the gastrointestinal tract, including the colon and small intestine. It encompasses two main forms: Crohn's disease and ulcerative colitis. IBD was defined using the algorithm developed by Lix et al. (2021). According to this algorithm, individuals with 5 or more inpatient hospital or physician visits, if they have 2 years or more of insurance coverage, or 3 or more visits for those with less than 2 years of coverage, were classified as IBD cases.

CVD is a medical condition that affects the heart and blood vessels. CVD was defined as meeting one of the following criteria: having one or more hospitalizations with a diagnosis of Ischemic Heart Disease as indicated by ICD-10-CA codes I20-I22, I24, I25 (Fransoo et al., 2009; Martens et al., 2010; Chartier et al., 2012; Fransoo et al., 2013; Martens et al., 2015; Fransoo et al., 2019).

COPD is a chronic inflammatory lung disease that causes obstructed airflow to the lungs, making it difficult to breathe. It typically included conditions such as chronic bronchitis and emphysema. The diagnosis codes (ICD-10-CA: J40 - J44) indicate that individuals who received a primary diagnosis falling within these codes were classified as having COPD (Chateau et al., 2019).

Asthma is a chronic respiratory condition characterized by inflammation and narrowing of the airways, leading to recurring episodes of wheezing, breathlessness, chest tightness, and coughing. The diagnosis code (ICD-10-CA: J45) indicates that individuals who received a primary hospitalization diagnosis falling within this code were classified as having asthma (Brownell et al., 2012; Chartier et al., 2016; Fransoo et al., 2019).

Another factor that has been established as a preventive measure against pneumonia is the pneumonia vaccine. The full vaccinated population has been determined by identifying individuals who received 1 dose of pneumococcal conjugate vaccine (PCV-23) or $4+$ doses of PCV-7/13 between 2000 and 2019. The analysis considered the use of PCV vaccines for pneumonia prevention; however, influenza vaccine was not included in the study. The exclusion of influenza vaccine was deliberate as influenza cases were not part of the designated outcomes. Influenza cases were defined using the ICD-10 codes J10-J11(WHO, 2019). Vaccination coverage for a particular year was calculated based on the population of that year and the available history of getting vaccines up to that year. For example, the numerator for the vaccination coverage of the year 2015 was the population who received the vaccine from 2000 to 2015 among the population in 2015.

2.6 Statistical Models

This section provided a detailed description of the statistical models employed in this thesis. Given that the outcome of interest was a count variable, a spatial Poisson regression model was deemed appropriate for identifying potential risk factors associated with pneumonia infection. To capture the spatial random effect, we utilized the ICAR model. The parameters of this model were estimated using a Bayesian approach, with the MCMC technique being employed for this purpose. To assess the influence of comorbidity, a domain-level mixed-effects Poisson regression model was employed to quantify its impact. We measured pneumonia transmission using both the SEIR compartmental

model, which excludes reinfection, and the SEIRS model, which accounts for pneumonia disease transmission and reinfection.

2.6.1 Spatial Model

For the count data we generally consider the Poisson regression model which assumes that these events occur independently across different regions. Let y_i be the observed count of events in the ith region, and x_i be a vector of covariate values for that region. The Poisson regression model assumes that:

$$y_i \sim \text{Poisson}(\mu_i); \quad i = 1, 2, \ldots, n,$$

where,

$$\log(\mu_i) = \log(E_i) + x_i^\top \beta,$$

β is a vector of regression coefficients and E_i is the expected number of pneumonia cases admitted in hospital in region i, n is the number of regions. The sex standardization is done through the E_i term, which is defined as:

$$E_i = \sum_{j=1}^{2} \frac{y_j^s}{n_j^s} n_{ij},$$

where, y_j^s and n_j^s are the pneumonia counts and population total in the provincial population in the jth sex group, respectively, and n_{ij} is the total population in the jth sex group in the ith health region.

However, in many cases, the events may be spatially dependent, i.e., the occurrence of an event in one region may affect the probability of an event occurring in a neighboring region. To account for this spatial dependence, the CAR model was developed, which assumes that the response variable in each region is conditionally dependent on the response variables in neighboring

regions. This model introduces a spatial random effect term that captures the dependence between the response variables in different regions. If we rewrite the above-mentioned model again, then it looks like

$$y_i \sim \text{Poisson}(\mu_i),$$

where,

$$log(\mu_i) = \log(E_i) + x_i^\top \beta + U_i + V_i, \tag{2.1}$$

V_i is the random effect for region i that follows a normal distribution with mean zero and variance σ^2, $V_i \sim \text{N}(0, \sigma^2)$, representing non-spatial over-dispersion and U_i is the spatially correlated random effect for region i. There are several forms for spatial random effects U_i in the context of the CAR model. We provided two common types of the CAR model below.

Join Model Approach

In this approach, we assume that $U = (U_1, \ldots, U_N)$ follows a multivariate normal distribution with mean 0, $\text{var}(U_i) = \sigma_U^2$, and correlations between U_i and U_j is $\text{corr}(U_i, U_j) = \exp(-\phi d_{ij}) = \rho^{d_{ij}}$, where d_{ij} is the distance between the centrists of regions i and j and $\rho > 0$ is a parameter that determines the extent of the correlation. There are other forms of correlation structure like spherical, Gaussian, etc. If we express a more general form of correlation, then it can be shown as

$$corr(U_i, U_j) = \frac{\exp(-(\phi d_{ij})^\kappa)}{\tau_U},$$

where τ_U is the marginal precision parameter of the CAR model, which controls the overall precision of the spatial random effects, ϕ is the spatial decay parameter, which controls the rate at which spatial correlation declines with distance, and κ is a smoothing parameter, which controls the rate at which spatial correlation changes with distance. The range of κ spans from 0 to 2, ensuring that the correlation function remains smooth and differentiable (Diggle et al., 1998). We use the term "join model" to refer to this statistical model as it involves the specification of the joint distribution of U.

Conditional Model Approach

Another way to model spatial dependence is the conditional model approach, where the distribution of each U_i is determined based on the known values of spatial random effects in neighboring regions, which are selected based on certain rules. However, selecting neighbors can be challenging, especially when the regions are irregular in shape, as in epidemiological contexts. One common approach is to define neighboring regions as those that share a common boundary, which works well for equally sized regions arranged in a regular pattern. In this scenario, an ICAR prior can be used to model the spatial random effects.

ICAR Model

In this approach, the U_i is the spatially correlated heterogeneity described by an ICAR distribution which is defined as

$$U_i|U_j, j\epsilon\partial_i \sim N(\bar{U}_i, \sigma_u^2/m_i),$$

where, m_i, $i = 1, \ldots, 96$, is the number of neighbors of region i, ∂_i indicates the set of neighbours of region i, $\bar{U}_i$ is the mean of the spatial random effects of these neighbours and σ_u^2 is a conditional variance that determines the amount of spatial dispersion. To define neighbours, we used the common boundary criterion which determines neighboring regions based on their shared boundaries. The model parameters were estimated via the Bayesian approach using the MCMC technique with a non-informative prior.

Prior selection: In this model, we have two variance components. One can choose a prior for each component separately, but it is difficult to control the overall residual relative risk. Therefore, it is more feasible to express the overall variability using a single prior. Thus, we adopted the approach of defining the total precision $\tau_T = (\sigma_v^2 + \sigma_u^2)^{-1}$, where τ_T was assumed to follow a Gamma distribution with scale and shape parameters 1 and 0.0005, respectively i.e. $\tau_T \sim \text{Gamma}(1, 0.0005)$, which implies a $\log t$ distribution for the overall residual relative risks. Let $p = \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$ representing the proportion of the total residual variation attributable to

the spatial component, follow a beta prior, i.e. $p \sim \text{beta}(1,1)$, and transform from (σ_T^2, p) to (σ_v^2, σ_u^2) via $\sigma_u = \sqrt{p/\tau_t}$; $\sigma_v = \sqrt{(1-p)/\tau_t}$; $\tau_u = 1/\sigma_u^2$; $\tau_v = 1/\sigma_v^2$. We also did a sensitivity analysis using flat uniform distributions for $\sigma_u \sim \text{uniform}(0.001, 10)$ and $\sigma_v \sim \text{uniform}(0.001, 10)$ to ensure the priors did not influence subsequent results.

The results have been presented in tabular format, including the mean, median, and a 95% credible interval. Notably, the highest posterior density credible interval has been employed in this model over other types of credible intervals.

During the analysis, we employed multiple diagnostic methods to ensure that the model parameters converged. The trace plots, MCMC error and the autocorrelation of generated samples of model parameters from the posterior distribution were used to evaluate the model's convergence. If the MCMC error is less than 5% of their standard deviation then model seems to be converge. We also utilized Gelman and Rubin (1992) convergence diagnostic test. The Gelman test score of 1 indicates perfect convergence of the model, although a score up to 1.10 is considered permissible. None of the diagnostic tests indicated non-convergence of the model parameters.

Selection of ICAR model: Spatial models have two aspects: the strength of dependence and the total amount of spatial dependence. In the ICAR model, there is only one parameter controlling both aspects, which may seem counterintuitive. In contrast, the joint model uses ρ to determine the strength and σ_U^2 to determine the total amount of dependence. However, since the ICAR model cannot accommodate non-spatial variability, a non-spatial random effect should always be included along with the ICAR random effect. This is achieved in the joint model with only u_i when $\rho = 0$. It is important to note that if the majority of the variability is non-spatial, inference for the joint model may suggest the presence of spatial dependence when it is not actually present. Another drawback of the join model is the computational cumbersome and it requires a significant amount of time even for a study with a moderate sample size.

2.6.2 Mixed Effect Poisson Regression

To quantify the impact of comorbidity (IBD, asthma, COPD and CVD), a domain-level mixed-effects Poisson regression model was utilized. The data were organized in two formats: the first involved grouping by age and sex, consisting of 107 age groups and two sex groups, while the second focused solely on age groups.

Let Y_{ij} be the count of pneumonia cases for the ijth age-sex group, where $i = 0, 1, \ldots, 107$ and $j = 0, 1$, where 0 for female and 1 for male. The number of pneumonia cases in each group follows an independent Poisson distribution with mean λ_{ij}, while this λ_{ij} is associated to the systematic components through a log-link function:

$$Y_{ij} \sim \text{Poisson}(\lambda_{ij}); \quad i = 0, 1, \ldots, 107; j = 0, 1$$

$$\begin{aligned} \log(\lambda_{ij}) = \log(k_{ij}) + \beta_0 + \beta_1 \text{IBD}_{ij} + \beta_2 \text{Vaccine}_{ij} + \beta_3 \text{Asthma}_{ij} \\ + \beta_4 \text{COPD}_{ij} + \beta_5 \text{CVD}_{ij} + u_i + v_{ij}, \end{aligned} \tag{2.2}$$

where $\log(k_{ij})$ represents the logarithm of the population size for the ij-th group, IBD_{ij}, Vaccine_{ij}, Asthma_{ij}, COPD_{ij}, and CVD_{ij} represent the count of individuals with IBD, the count of individuals received full doses of vaccine, the count of individuals with asthma, the count of individuals with COPD and the count of individuals with CVD, all within the ijth age-sex group. u_i represents the random effect for the age group i, and v_{ij} represents the random effect for the combination of age group i and sex group j. The random effect u_i follows a normal distribution with mean zero and variance σ_u^2, capturing the variability between different age groups. Similarly, the random effect v_{ij} follows a normal distribution with mean zero and variance σ_v^2, representing the variability between different sex groups within age groups.

In the case of focusing only on the age groups, we arrange data as follows: Y_i for the count of pneumonia cases in the ith age group and Male_i for the count of males in the ith age group, IBD_i, Vaccine_i, Asthma_i, COPD_i, and CVD_i for the count of individuals with IBD, the

count of individuals with full vaccinated people, the count of individuals with asthma, the count of individuals with COPD, and the count of individuals with CVD, all within the ith age group, and k_i as the total population size of the ith age group. Then model can be written as:

$$Y_i \sim \text{Poisson}(\lambda_i); \quad i = 0, 1, \ldots, 107$$

$$\begin{aligned} \log(\lambda_i) = \log(k_i) + \beta_0 + \beta_1 \text{Male}_i + \beta_2 \text{IBD}_i \\ + \beta_3 \text{Vaccine}_i + \beta_4 \text{Asthma}_i + \beta_5 \text{COPD}_i + \beta_6 \text{CVD}_i + u_i \end{aligned} \tag{2.3}$$

where u_i represents the random effect term capturing the random variation between different age groups.The random effect term u_i is assumed to follow a normal distribution with mean zero and variance σ_u^2.

2.6.3 Compartmental Models

Compartmental models are widely used in epidemiology to study the dynamics of infectious diseases. The concept of compartmental modeling was first introduced by Kermack and McKendrick (1927) to describe the spread of infectious diseases in a population. The basic idea is to divide the population into different compartments or classes based on their disease status, and then to model the flow of individuals between these compartments over time.

Susceptible-Infected-Recovered (SIR) Model

The most commonly used compartmental model is the SIR model, which stands for Susceptible-Infected-Recovered. In this model, individuals in the population are classified into three compartments based on their disease status. The susceptible compartment (S) represents individuals who are susceptible to the disease but have not yet been infected. The infected compartment (I) represents individuals who are infected with the disease and are capable of transmitting it to others. The recovered compartment (R) represents individuals who have recovered from the disease and are no

longer infectious. Then, the total number in the population can be written as

$$N = S + I + R$$

The dynamics of the disease are modeled using a system of ordinary differential equations that describe the flow of individuals between these compartments. The dynamics of the SIR model can be described by the following set of differential equations:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\frac{\beta S(t) I(t)}{N} \\
\frac{dI(t)}{dt} &= \frac{\beta S(t) I(t)}{N} - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t)
\end{aligned}$$

where

- $S(t)$ is the number of susceptible individuals at time t
- $I(t)$ is the number of infectious individuals at time t
- $R(t)$ is the number of recovered individuals at time t
- β is the transmission rate (rate at which susceptible individuals become infected upon contact with infected individuals)
- γ is the recovery rate (rate at which infected individuals recover and become immune)

The model assumes that individuals move between these compartments according to certain rates, which can be affected by factors such as the transmission rate of the disease, the duration of the infectious period, and the efficacy of interventions such as vaccination or treatment. By simulating the movement of individuals between the compartments, the model can provide insights into the number of individuals who will be infected, the time course of the epidemic, and the effectiveness of various control measures.

Since the introduction of the SIR model, many variations of the basic compartmental model

have been developed to account for more complex disease dynamics, such as the SEIR (Susceptible-Exposed-Infectious-Recovered) model and the SIS (Susceptible-Infectious-Susceptible) model.

SEIR Model

The SEIR model is an extension of the SIR model, which adds a new compartment for individuals who are exposed to the infectious agent but have not yet become infectious. The SEIR model is particularly useful in studying infectious diseases that have a prolonged latent period, during which individuals are exposed to the infected agent but are not yet infectious themselves. The population is divided into four compartments which is similar to SIR model. The additional compartment is Exposed (E) defined as individuals who have been exposed to the disease but are not yet infectious. Therefore, the total number in the population can be written as

$$N = S + E + I + R$$

Each individual in the population belongs to one, and only one, of the aforementioned compartments at any given time point. Let us consider the disease transmission rate β, the average duration of the latent period before the infection σ and the recovery rate from the infection γ. Then, combining all the definitions and assumptions, the model for the transmission dynamics of pneumonia is given by the following system of differential equations:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\frac{\beta S(t) I(t)}{N} \\
\frac{dE(t)}{dt} &= \frac{\beta S(t) I(t)}{N} - \sigma E(t) \\
\frac{dI(t)}{dt} &= \sigma E(t) - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t)
\end{aligned}$$

In the context of our research, the SEIR model may fit well for measuring pneumonia transmission rate as it considers the latent period. The SIR model assumes that individuals move directly from the susceptible compartment to the infectious compartment, and from the infectious

compartment to the recovered compartment, without any delay or latency period. However, for many infectious diseases, including pneumonia, there is a period of time between exposure to the infected agent and the onset of infectiousness. This is known as the latent period, during which individuals may be exposed to the infected agent but are not yet infectious themselves. Therefore, a model that includes a latent period, such as the SEIR model, may be a better fit for studying the transmission dynamics of pneumonia.

SEIRS Model

The SEIRS (Susceptible-Exposed-Infectious-Recovered-Susceptible) model is an epidemiological model used to describe the spread of infectious diseases and also its reinfection in a population. It builds upon the SEIR model by introducing the possibility of individuals becoming susceptible again after recovering from the disease. This captures scenarios where immunity is not permanent, and individuals can be reinfected over time. The SEIRS model is described by the following set of differential equations:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\frac{\beta S(t) I(t)}{N} + \omega R(t) \\
\frac{dE(t)}{dt} &= \frac{\beta S(t) I(t)}{N} - \sigma E(t) \\
\frac{dI(t)}{dt} &= \sigma E(t) - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t) - \omega R(t)
\end{aligned}$$

where ω is the rate at which immunity wanes, making recovered individuals become susceptible again. The SEIRS model allows for a more realistic representation of certain diseases where immunity is not lifelong, such as seasonal flu or some respiratory disease like pneumonia, COVID-19, etc. It captures the dynamic interplay between infection, recovery, and susceptibility to reinfection, making it a valuable tool for studying and predicting the spread of such diseases.

2.7 Analysis Plan

The extracted data has been checked to see whether there was any abnormality. We have measured the crude pneumonia rates for 96 Manitoba regions and presented this rate in a graph to get a clear indication of any potential spatial pattern. As sex was known to contribute to the risk of pneumonia infection, a standardization technique has been used considering the Provincial population as the reference. This enabled the effects of sex to be accounted for and the risk ratios to be compared across different groups. The standardized expected count has been used as the outcome measure for subsequent statistical models. We fitted the spatial Poisson regression model considering all covariates that mentioned in Equation (2.1) where the spatial variability was measured by the ICAR model. The adjusted risk ratio based on the fitted model has been presented in the Manitoba geographical maps for making a comparison with the crude estimates maps. The estimated spatial effect has also been presented in Manitoba maps to understand the spatial variation of pneumonia incidence. To rigorously assess the impact of comorbidity, a mixed-effects Poisson regression model was employed, considering two distinct scenarios. The model equations were presented as Equation (2.2) and Equation (2.3) in the research study.

We employed the ordinary differential equation technique to estimate model parameters for both the SEIR and SEIRS models, optimizing the parameters using the Nelder-Mead algorithm. A total of 1000 iterations were conducted to refine the parameter estimates. To evaluate the accuracy of the fitted models, we employed mean square error (MSE), defined as:

$$\text{MSE} = \frac{1}{T}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2$$

In this expression, T represents the total number of time points, y_i represents the actual number of reported cases on a daily basis, and $\hat{y}_i$ represents the number of predicted cases. The quality of model fitting is assessed by identifying the set of parameters that minimizes MSE resulting in the best fit to the data.

A list of statistical software including *R, SAS, QGIS*, and *WinBuGS* has been used for data management, cleaning, data manipulation, final analysis, and mapping.

2.8 Ethical Consideration

The data for this research was already collected by the Manitoba Health. So, there is no active participation of the respondent in this research. In this regard, participants had not been facing any risk in terms of psychological, social, cultural, behavioural, or economic aspects. We did not need to obtain any written consent from the participants in the study because we analyzed recorded administrative health data. Data privacy and confidentiality have been maintained as per the MCHP protocol. The principal investigator has completed research ethics training provided by the Tri-Council Policy Statement Ethical Conduct, and privacy training for the researcher and Winnipeg Regional Health Authority Personal Health Information Act. To ensure confidentiality, all data from the PHRDR has fully been anonymized before being extracted for this research. The research protocol has been approved by the Research Ethics Board at the University of Manitoba and the Provincial Health Research Privacy Committee at Manitoba Health.

Chapter 3

Results

3.1 Data Description

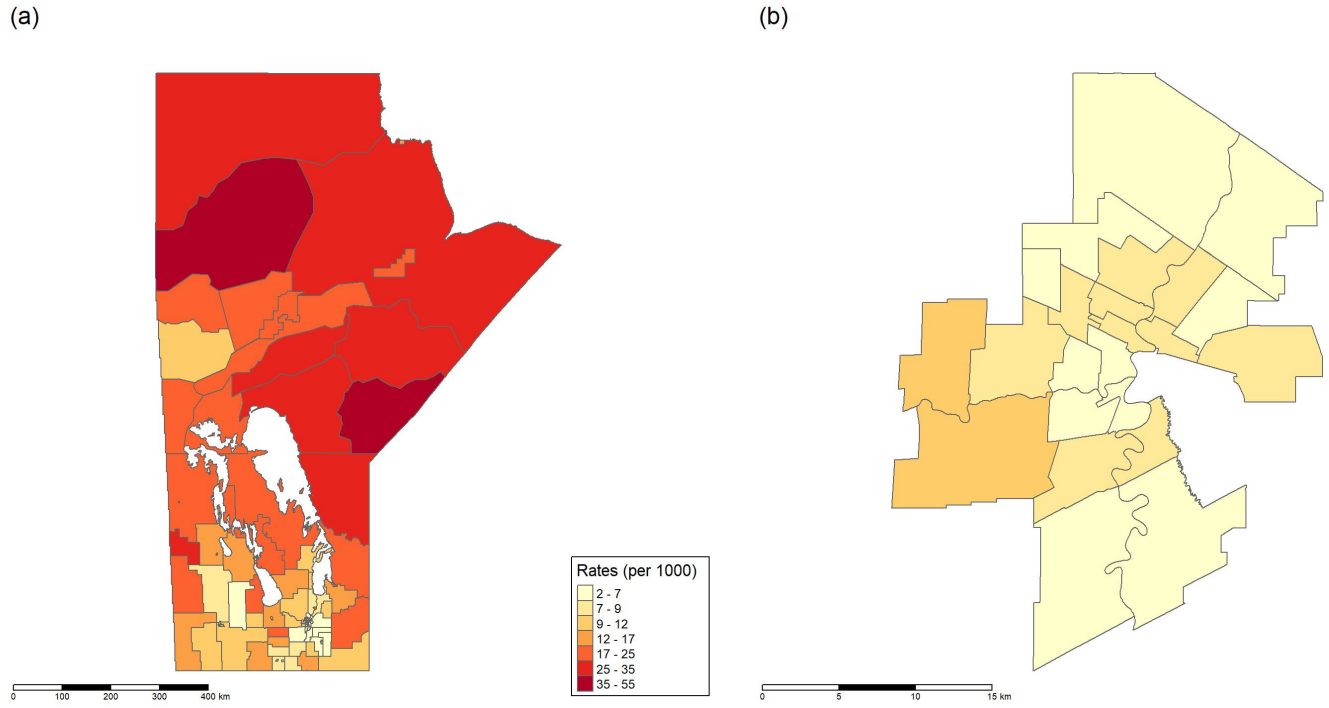
During the five-year span from 2015 to 2019, Manitoba hospitals admitted a significant number of patients with various health conditions. The admissions included $1{,}636$ cases of asthma, $1{,}615$ cases of IBD, $13{,}988$ cases of COPD, and $18{,}627$ cases of CVD. Concurrently, there were $13{,}561$ patients were admitted to hospital, emphasizing the burden on the healthcare system. Among these pneumonia cases, 10.39% were children under the age of 5, and 26.43% were in the age group of 6 to 59 years, while the highest number of cases (63.18%) occurred in people aged 60 years and over. Table 3.1 gives description of covariates among $13{,}561$ pneumonia cases.

Table 3.1: Descriptive statistics of covariates among $13{,}561$ pneumonia cases

Independent variable	Frequency	Percentage
Asthma	353	2.60%
IBD	186	1.37%
COPD	248	1.83%
CVD	593	4.37%
Full doses of vaccine	7,993	58.94%

In order to calculate the rate of pneumonia transmission, we utilized a dataset spanning one year. During this period, we documented a total of $16{,}079$ cases of pneumonia. These cases were

Figure 3.1: Geographical distribution of raw rates (per 1000) of pneumonia cases across the 96 RHADs, (a)Manitoba, (b)Winnipeg



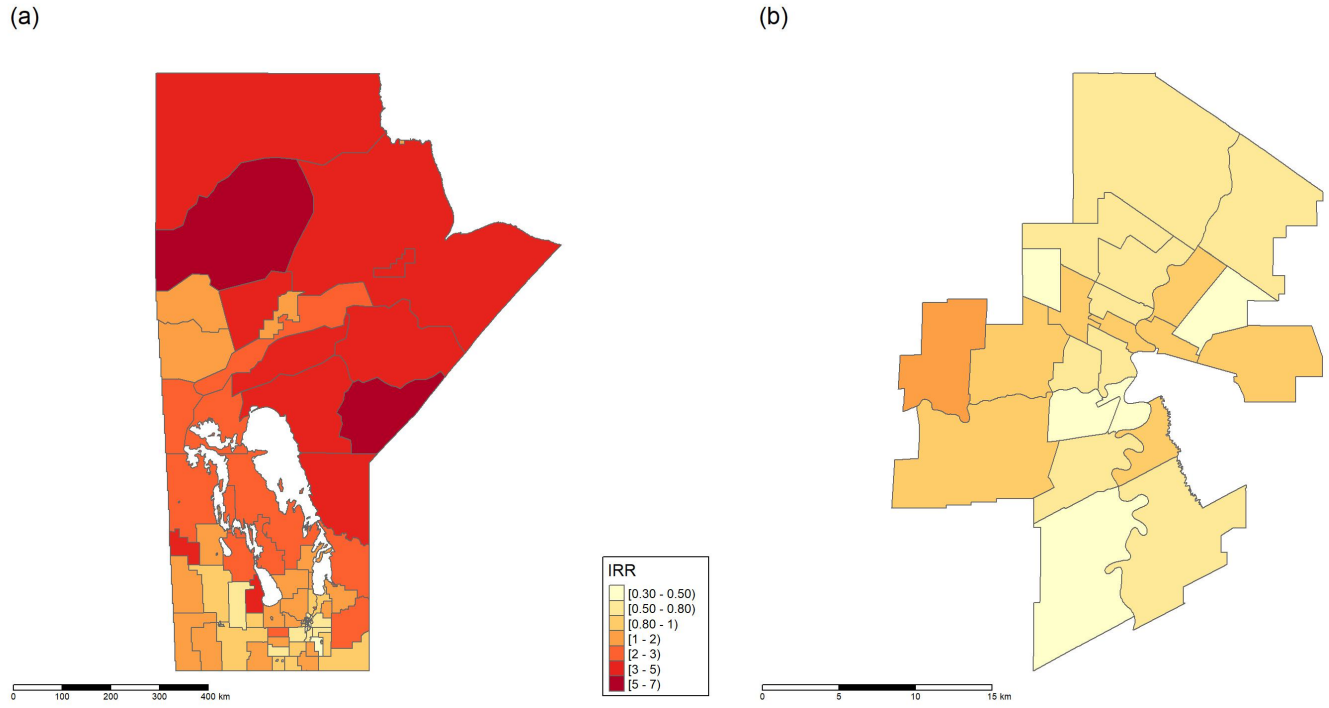
identified either within a hospital setting or diagnosed by physicians. A few cases experienced reinfection within this time period, accounting for 1,228 cases (7.10%), characterized as reinfection pneumonia.

3.2 Spatial Variation of Pneumonia Infection

To assess the spatial variability of pneumonia in Manitoba, we analyzed the raw incidence rate of pneumonia and present the results in Figure 3.1. The incidence rate ranges from 2 to 55 per 1000 population across different regions. The Northern RHA exhibited the highest incidence rate, whereas the lowest incidence rate was observed in Winnipeg RHA, followed by the Southern RHA. These findings demonstrated a significant spatial variability of pneumonia incidence in Manitoba.

Figure 3.2 presents the incidence rate ratio (IRR) for the fitted model, which shows that there is spatial variability in pneumonia incidence even after adjusting for potential risk factors.

Figure 3.2: Maps of sex standardized IRR of overall pneumonia cases using spatial Poisson regression model, (a)Manitoba, (b)Winnipeg



The range of the IRR was between 0.37 and 6.39, indicating that some regions have a much higher incidence rate than other regions. The highest IRR was still observed in the Northern RHA, followed by the Northern part of the Interlake-Eastern region. In contrast, the Southern RHA and the south side of the Prairie Mountain RHA showed the lowest IRR of hospitalization due to pneumonia. Interestingly, despite Winnipeg city being the most populated city in the province, the pneumonia IRR was similar to that of the Southern RHA region.

Figure 3.3 illustrates the spatial autocorrelation of the model, taking into account the spatial structure incorporated in the model. The autocorrelation ranged between -0.095 and 0.055, indicated a stronger presence of negative autocorrelation compared to positive values. This suggested that there was an inverse relationship between spatial location and the incidence of pneumonia, particularly in the Northern RHA. This means that closer locations in these regions were less likely to have similar rates of pneumonia incidence.

The finding from the spatial Poisson regression model utilized in this study is presented in

Figure 3.3: Maps of spatial variability in overall pneumonia cases, results from ICAR model, (a)Manitoba, (b)Winnipeg

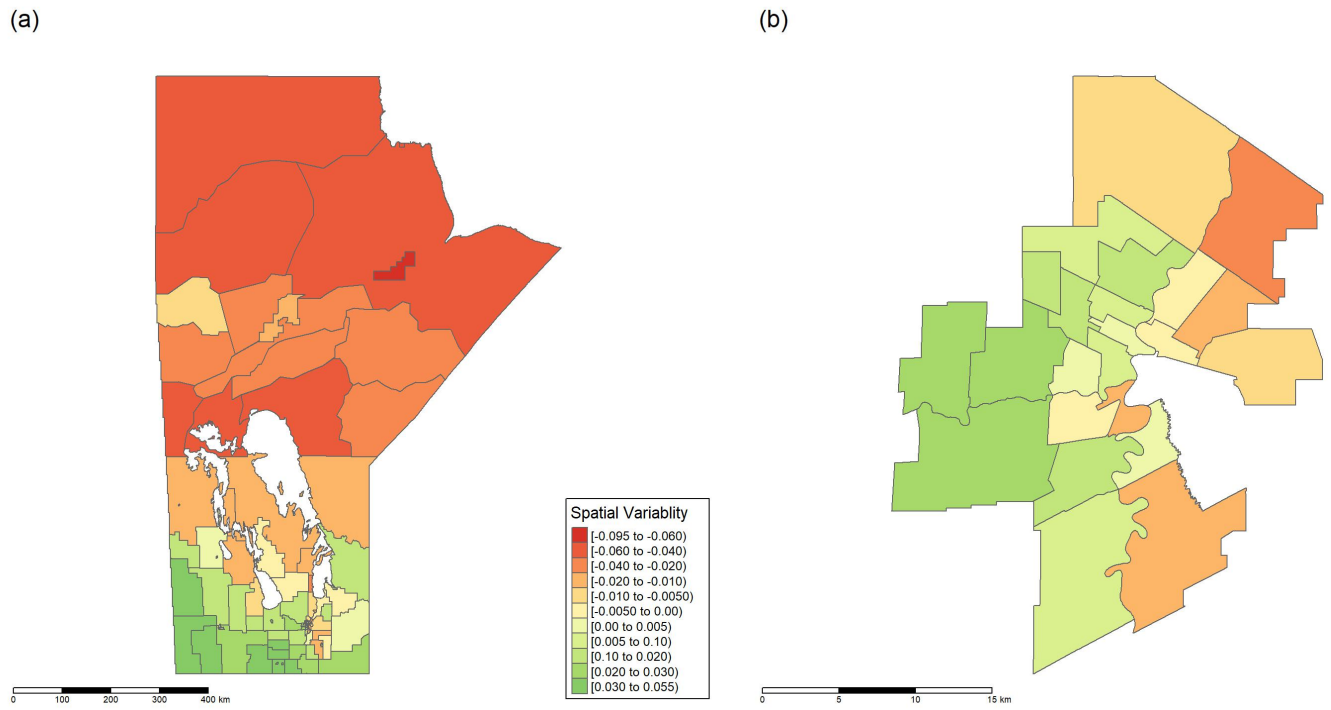


Table 3.2: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	95% CI
Intercept	-1.80	[-2.22, -1.38]
Age $(0-5)$	5.64	[2.03, 9.22]
Age $(6-59)$	13.19	[0.64, 25.67]
Age (60 and above)	30.57	[22.19, 38.78]
SEFI 2	0.04	[-0.02, 0.10]
Immigration Status	-4.37	[-6.97, -1.77]
Full Vaccination History	4.83	[3.26, 6.51]
Age $(>= 60)$*Immigration	1.81	[0.89, 2.72]
Age$(>= 60)$*Full vaccine	-0.74	[-1.01, -0.46]
Comorbidity: IBD	-89.74	[-211.30, 33.07]
Comorbidity: COPD	12.82	[3.69, 21.69]
Comorbidity: CVD	11.83	[3.81, 19.97]
IBD*Age(6 to 59)	3.13	[-3.50, 9.63]
p (ratio of spatial variability)	0.21	[0.01, 0.61]

Table 3.2. Prior to the determination of the final model, an array of potential risk factor combinations was contemplated and applied to the data. In total, twelve potential model combinations were assessed. The ultimate model, as depicted in Table 3.2, is chosen based on its statistical significance and the goodness of fit. The remaining results for the other considered models are provided in Appendix A [Table A.1 - A.11]. All the covariates selected for the study had a significant impact on the pneumonia infection except SEFI 2, and patients' history of IBD.

The results showed that the posterior mean IRR for children below the age of 5 was 5.64, and the 95% credible interval ranged from 2.03 to 9.22, which was the lowest among the three age groups studied. In contrast, the posterior mean for the second age group ($6 - 59$ years) was 13.19 with 95% credible interval was between 0.64 and 25.67. It was observed that the third age group had the highest IRR with a mean of 30.57 and 95% credible interval of $(22.19, 38.78)$, in comparison to the other two age groups.

This study also found that the immigrant population was more vulnerable to pneumonia than non-immigrants, particularly among older people aged 60 years over. Moreover, vaccination coverage played a protective role in preventing pneumonia among older people. The increase in full doses of vaccine coverage for older adults (age ≥ 60) led to a reduction in the mean posterior distribution of the IRR by 0.74 units. The credible interval for this estimate was between -1.01 and -0.46 units.

The analysis revealed a significant positive association between comorbidity conditions and the likelihood of pneumonia infection. Among all the comorbidity conditions examined, COPD exhibited the highest posterior mean IRR. The mean posterior IRR for COPD was estimated to be 12.82, with a 95% credible interval ranging from 3.69 to 21.69. This indicates that with the increase of a single unit of COPD in a region, then the incidence rate of pneumonia infection increased 12.82 units, after adjusting for other covariates in the model.

Furthermore, CVD was also found to significantly contribute to the development of pneumonia, similar to COPD. The mean posterior IRR for CVD was estimated to be 11.83, with a 95% credible interval ranging from 3.81 to 19.97. The analysis revealed that the ratio of spatial variability, p, was found to be 21% that suggested that spatial variation accounts for 21% of the total

variability observed in the data, with the remaining 79% attributed to other sources of variation.

Table 3.3: Pneumonia infection and comorbidity: results from a mixed-effects Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	- 5.022	[- 5.253, - 4.790]
IBD	0.034	[0.007, 0.061]
Asthma	0.009	[- 0.009, 0.027]
COPD	0.028	[0.003, 0.053]
CVD	0.052	[0.037, 0.067]
σ_u	1.297	[0.923, 1.672]
σ_v	0.014	[0.004, 0.024]
AIC	2001	
BIC	2020	

3.3 Influence of Comorbidity on Pneumonia Infection

Table 3.3 presents the results of the final mixed-effects Poisson regression model for age-sex group data assessing the influence of comorbidity on pneumonia infection. Out of the 16 models, this model demonstrated the best fit based on AIC and BIC values. Appendix B [Table B.1 - B.15] provides the results for the other scenarios.

The results revealed that IBD, COPD, and CVD had a significant positive impact on pneumonia infection, while asthma had no impact despite having a positive coefficient. Among these three conditions, CVD exhibited the highest impact, with a hospitalization rate of 0.052 (95% CI: $0.037-0.067$), IBD followed with a hospitalization rate of 0.034 (95% CI: $0.007-0.061$), while COPD had the lowest rate of hospitalization at 0.028 (95% CI: $0.003-0.053$). The impact of asthma on pneumonia infection has been explored for children (Age $\leq$ 17 years) and adults (Age $\geq$ 18 years) separately. Among all possible combinations, results from a mixed-effect Poisson regression model with age as a random effect showed a positive association [0.023 (95% CI: $0.004 - 0.042$)] between asthma and pneumonia infection. The results are presented in Appendix B (Table B.13).

In the second scenario, where data was arranged based on age groups only, the best model's

Table 3.4: Pneumonia infection and comorbidity: results from a mixed-effects Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	- 5.580	[- 6.109, - 5.59]
Male	0.007	[0.003, 0.011]
IBD	0.002	[- 0.093, 0.096]
CVD	0.102	[0.068, 0.134]
σ	0.644	[0.456, 0.833]
AIC	1225	
BIC	1238	

final results are presented in Table 3.4. The remaining model results can be found in Appendix B [Table B.16 - B.26]. Similar to the previous scenario, the model indicated a positive association between CVD and pneumonia infection, after adjusting for other covariates. The infection rate was estimated to be 0.102, with a 95% CI of 0.068 to 0.134. Furthermore, male residents exhibited a higher infectiousness compared to females. The analysis also indicated that IBD contribution to pneumonia infection was not statistically significant, despite having a positive coefficient. An age-segregated analysis was conducted to assess the impact of asthma on pneumonia infection. However, none of the model results show a positive impact of asthma on pneumonia infection.

3.4 SEIR and SEIRS Model Results

SEIR Model

The pneumonia transmission rate was assessed through the SEIR model. In the SEIR model, each individual in the population is assigned to one of the mentioned compartments at any given time. To initiate the SEIR model, we required initial values for its parameters. Given the limited literature on pneumonia transmission rates specific to Manitoba, we employed the transmission rate of COVID-19 as an initial approximation, considering the similarities between pneumonia and COVID-19 (Amiri et al., 2023). The incubation and recovery periods for pneumonia were reported as ranging from 1 to 3 days (CDC, 2022; Public Health, 2023) and 7 to 14 days (NIH, 2022; Panel,

2023), respectively. We adopted initial values of $\beta = 0.24$ for the transmission rate, $\sigma = 0.33$ for the exposure rate, and $\gamma = 0.14$ for the recovery rate.

Under these assumptions, we fit the SEIR model, yielding the parameter estimates detailed in Table 3.5.

Table 3.5: Results of SEIR and SEIRS model fitting for Manitoba

Parameter	SEIR	SEIRS
β	0.81	0.88
σ	0.48	0.57
γ	0.78	0.85
ω	-	0.003
R_0	1.03	1.03

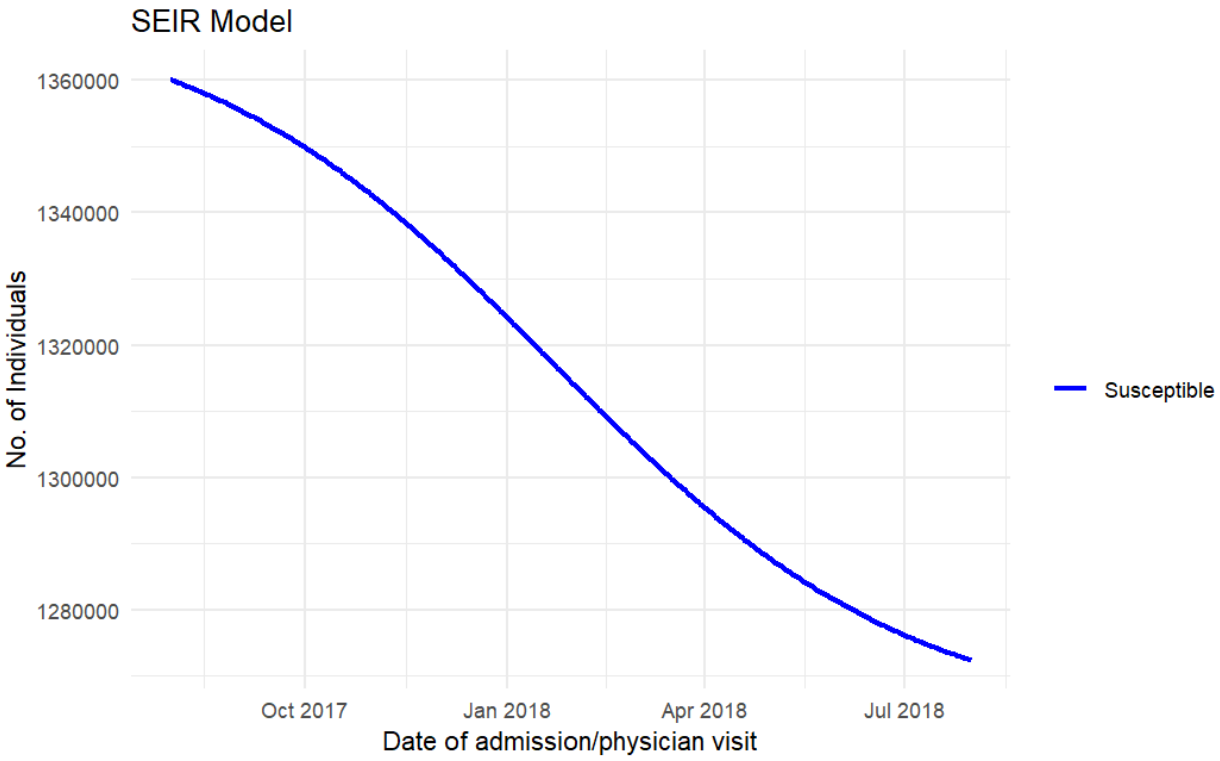
Multiple combinations of initial values for exposure and infected cases were considered, with the selected combination yielding the best-fitting model, supported by the lowest MSE of 0.078. The pneumonia transmission rate was estimated at 0.81, with the exposure rate at 0.48 and the recovery rate at 0.78. Figure 3.4 illustrates the results, depicting a comparison between observed and fitted data. Figure 3.5 shows the graphical presentation of susceptible cases and recovery cases of SEIR model.

SEIRS Model

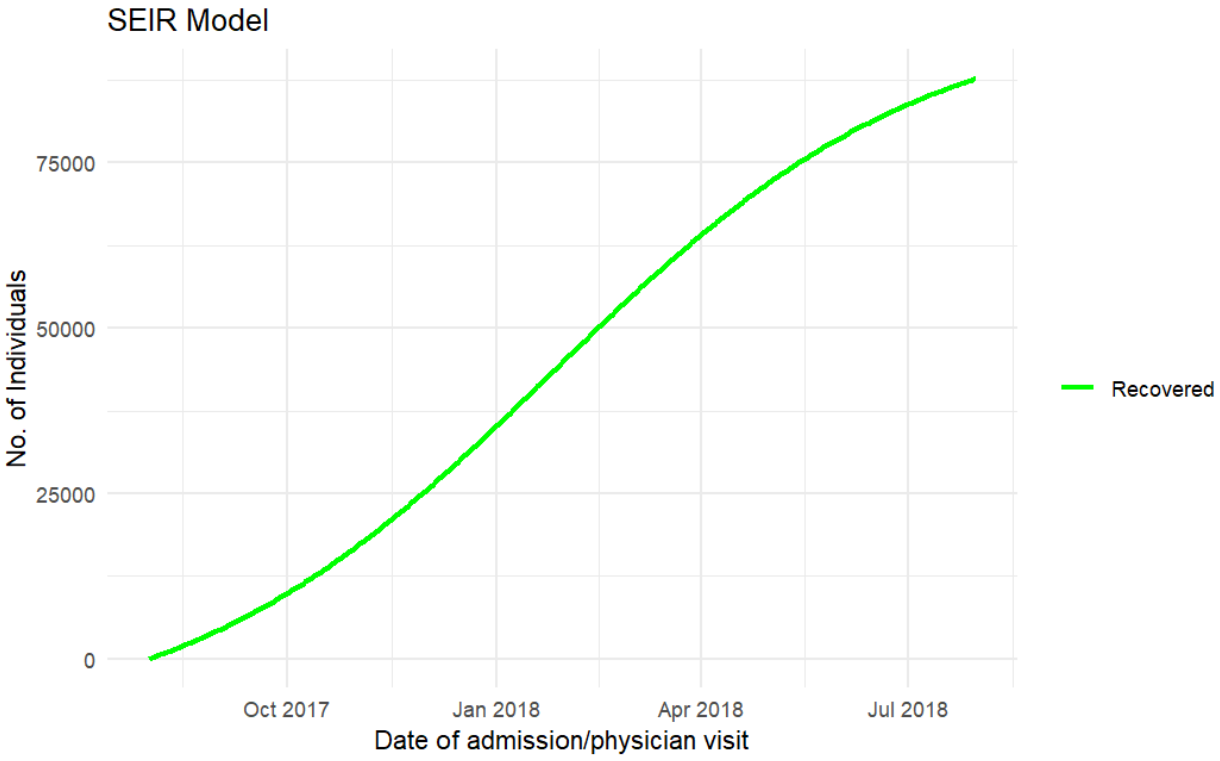
Inherent to the nature of the disease, individuals previously afflicted with pneumonia face the possibility of experiencing another infection, a condition referred to as reinfection pneumonia.

To comprehend the reinfection dynamics, we employed the SEIRS model, which aptly accounts for reinfection phenomena while estimating model parameters. In this modeling framework, we utilized the parameter estimates obtained from the SEIR model as initial values for the SEIRS model parameters. The resulting fitted model parameters are presented in Table 3.5, where it is evident that the reinfection rate stands at 0.003. Furthermore, the analysis yielded adjustments in other parameter values, notably indicating a transmission rate of 0.88, an exposure rate of 0.57, and a recovery rate of 0.85. We assessed the model's goodness of fit by calculating the MSE, which

Figure 3.4: Results from SEIR model: graphical presentation of the observe line and fitted line.



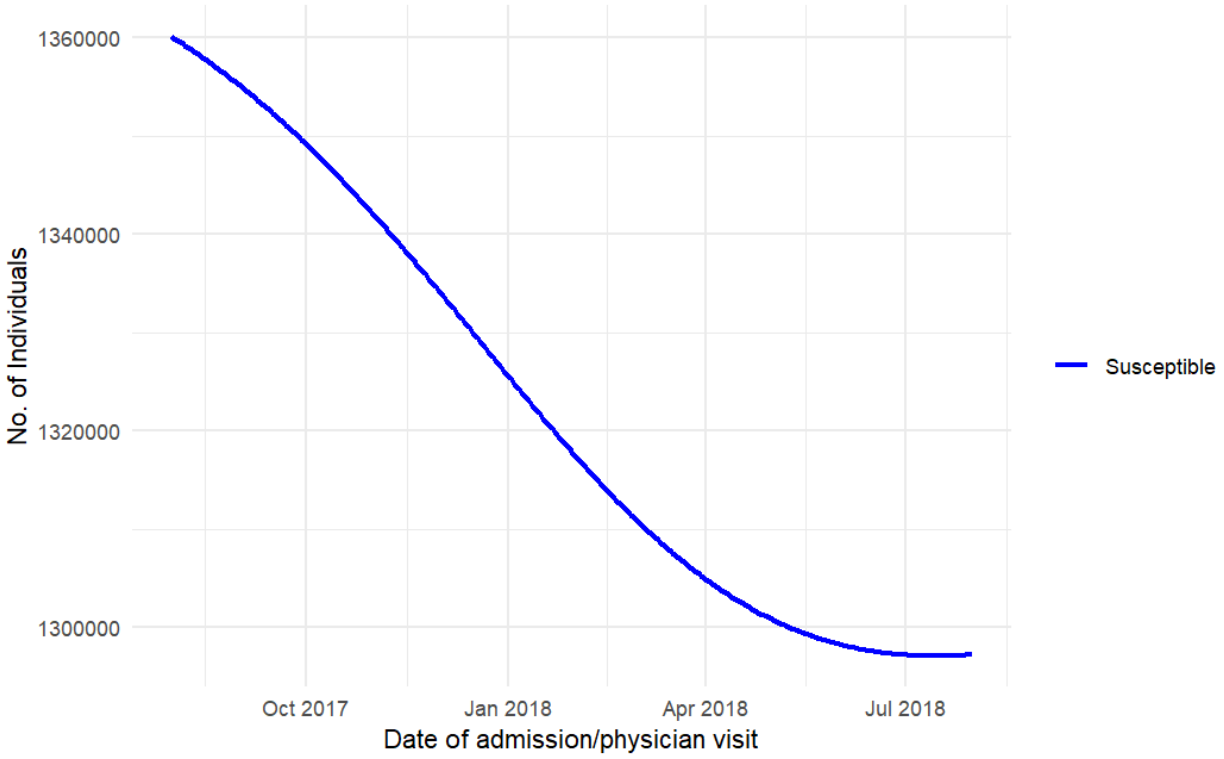
(a) Graphical presentation of the susceptible cases for SEIR model



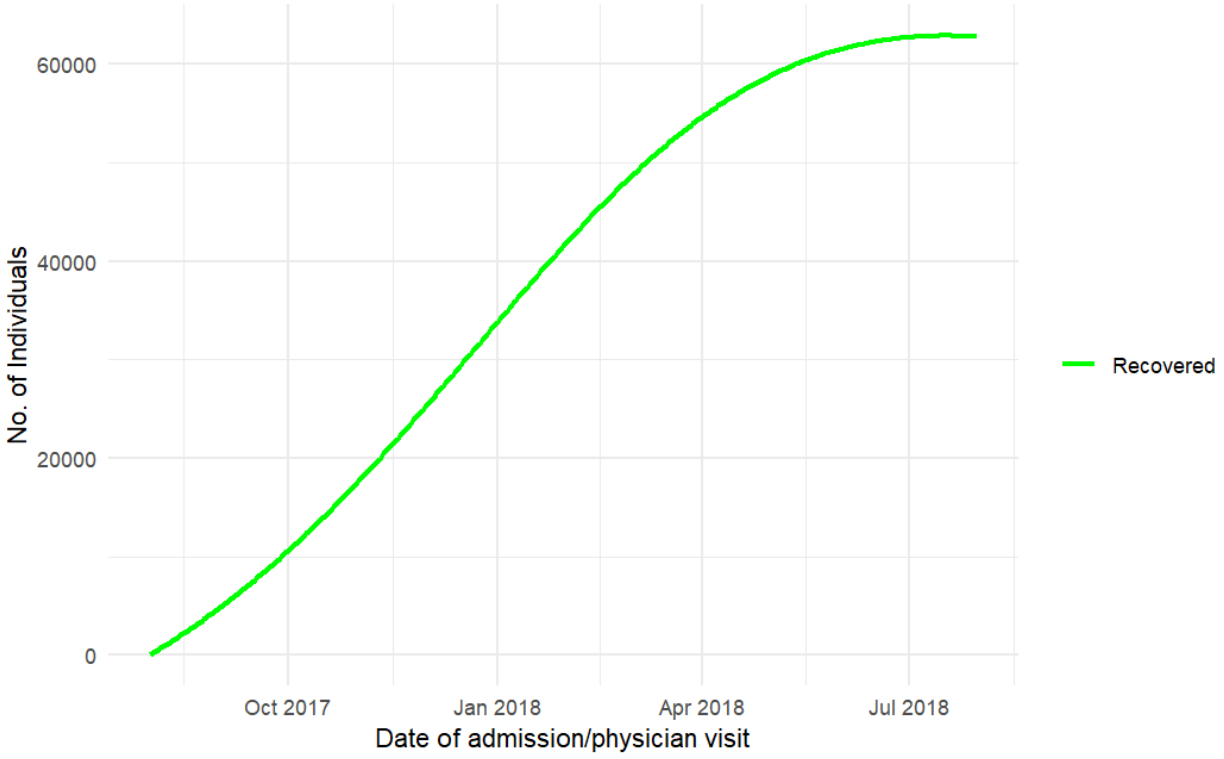
(b) Graphical presentation of the recovery cases for SEIR model

Figure 3.5: Graphical presentation of the susceptible and recovery cases for SEIR model

Figure 3.6: Results from SEIRS model: graphical presentation of the observe line and fitted line.



(a) Graphical presentation of the susceptible cases for SEIRS model



(b) Graphical presentation of the recovery cases for SEIRS model

Figure 3.7: Graphical presentation of the susceptible and recovery cases for SEIRS model

yielded a value of 1.36. Figure 3.6 shows the graphical presentation of observed and fitted cases while Figure 3.7 shows the susceptible and recovery cases for the SEIRS model.

An R_0 value of 1.03 in the SEIR and SEIRS model suggested that, on average, each individual infected with pneumonia transmits the disease to approximately 1.03 other individuals during their period of infectiousness. In the context of infectious diseases, R_0 served as an important epidemiological indicator. When R_0 is greater than 1, it indicates that the disease has the potential to spread within the population, leading to sustained transmission. The R_0 of 1.03 is moderately above 1, which indicated that there was some potential for sustained transmission of pneumonia within the population, although it may not lead to large-scale outbreaks.

Chapter 4

Discussion and Conclusion

4.1 Discussion

The primary objectives of this thesis revolved around gaining a deeper comprehension of the spatial patterns in pneumonia hospitalizations. We aimed to identify potential covariates at both regional levels and domains that play a role in influencing pneumonia infections. Additionally, our focus extended to measuring the disease transmission mechanisms, encompassing single infections as well as reinfection scenarios. Regional level modeling is a type of ecological modeling that focuses on the spatial distribution of environmental factors and how they affect human health. It is used to identify areas with high levels of environmental risk and to develop strategies to reduce exposure to these risks. While there is no direct link between domain-level modeling and area-level modeling, they can be used in conjunction with each other to provide a more comprehensive understanding of the complex systems that influence human health and behaviour. On the other hand, the present study was an ecological analysis and additional studies are needed to confirm its findings. The subsequent sections present various findings derived from the results.

We observed significant variations in pneumonia hospitalization rate ratios across the Manitoba 96 health regions, with regions exhibiting similar ratios clustering together, suggesting common spatial processes influencing hospitalization patterns. The highest rates were concentrated

in the Northern regions, particularly in districts 3 and 15, where the estimated IRR ranged from 5 to 7. Other northern regions also reported elevated IRR, except for districts 1, 2, and 4. The Interlake Eastern region showed a similar IRR range between 1 to 3. Conversely, the lowest IRR values were observed in various point of Southern regions and the Whitemud region within the Prairie Mountain RHA, where IRR ranged from 0.30 to 0.50. Within Winnipeg, consisting of 25 health regions, most reported low IRR, less than 0.80. Notably, the St. James-Assiniboia West area displayed the highest IRR, estimated between 1 and 2. Other regions, including Assiniboine South, St. James-Assiniboia East, Inkster East, Point Douglas South, River East South and West, Transcona, and St. Boniface, reported IRR between 0.8 to 1. These findings align with the Epidemiology of Communicable Diseases in the Winnipeg Health Region, $2013 - 2018$ report, which indicated that the age-standardized rate for Invasive Pneumococcal Disease (IPD) was highest in Point Douglas, followed by downtown, Assiniboine South, St. James-Assiniboia, Inkster, and River East (Population Health Surveillance Team, WRHA, 2020). Annual Statistics for Manitoba also revealed that hospitalization for Ambulatory Care, a group of diseases, was most frequent in the Northern regions, where bacterial pneumonia infections ranked second in terms of disease prevalence (Manitoba Health, 2019). Spatial variation in pneumonia hospitalization rates is not unique to Manitoba. Similar patterns have been observed in other Canadian provinces, such as Ontario and Newfoundland and Labrador. Research conducted in Ontario supports these findings, indicating that hospitalization rates for pneumonia are notably higher in Northern territories when compared to other regions (Crighton et al., 2007b, 2008).

We also examined spatial autocorrelation (Figure 3.4), which revealed an inverse association between IRR and spatial location. The majority of negative values were concentrated in the Northern RHA, particularly in the Interlake Eastern region, as well as in certain parts of Winnipeg. This suggests a tendency for dissimilar health regions to cluster together in these areas. In contrast, most of the positive correlations were observed in the Southern RHAs and the Prairie Mountain RHA, indicating that regions with similar characteristics tended to be in close proximity to each other.

This research highlighted the heightened vulnerability of older individuals, specifically those aged 60 and above, to pneumonia infection compared to other age groups, irrespective of migration status. This heightened susceptibility in older adults can be attributed to several age-

related physiological changes in the respiratory system. These changes include reduced pulmonary reserve, decreased mucociliary transport, a less effective cough reflex, diminished alveolar elasticity, and compromised ventilation (Mouton et al., 2001). Furthermore, the aging process leads to a progressive decline in immunological responses, further increasing the risk of pneumonia in older populations (Chebib et al., 2021).

Moreover, individuals already grappling with chronic conditions like COPD and CVD face an elevated risk of hospitalization due to pneumonia. Specially, patients with COPD who contracted pneumonia experience exacerbated breathing difficulties, often necessitating hospitalization. Studies have indicated that a substantial proportion of first-time hospitalizations for COPD exacerbations are directly related to pneumonia, accounting for 36.1% of such cases (Søgaard et al., 2016). Additionally, research has shown that the use of inhaled corticosteroids (ICS) in COPD patients is associated with a $20\% - 30\%$ increase in the risk of developing pneumonia (Janson et al., 2018).

Vaccination coverage emerges as a pivotal factor in safeguarding individuals against pneumonia. Our analysis revealed the pronounced effectiveness of vaccination in preventing pneumonia among older adults. However, its impact appeared statistically insignificant for other age groups, including children. Pneumococcal conjugate vaccines (PCVs) have proven to be instrumental in providing protection against vaccine-type pneumococcal disease in both pediatric and adult populations (Dunne et al., 2023). Research also suggested that PCVs extend their protective benefits beyond pneumonia, encompassing a reduction in lower respiratory tract infections, including defense against respiratory diseases associated with viral agents (Dunne et al., 2023). This broad spectrum of protection underscores the value of vaccination. In line with these findings, the Canadian Immunization Guide advocates the routine immunization of adults aged 65 and older with pneumococcal vaccines. This preventive measure aims to combat diseases stemming from Streptococcus pneumoniae, pneumonia being a prominent concern among them (Government of Canada, 2023, 2019).

From the perspective of our domain-level model, we unveiled an additional comorbidity factor: IBD, alongside COPD and CVD. IBD displayed a substantial and significant impact on pneumonia infection risk. A study featured in the American Journal of Gastroenterology delved

into the realm of pneumonia risk in IBD patients, along with the potential influence of specific medications on this risk (Long et al., 2013). The findings of this study revealed that individuals with IBD face an elevated risk of pneumonia when compared to the general population. Moreover, the study pinpointed corticosteroids and narcotics as medications particularly associated with pneumonia incidence in this specific population (Long et al., 2013; Okafor et al., 2013; Sierra and Daiya, 2022).

Asthma, too, emerged as a significant factor influencing pneumonia infection, with a more pronounced effect observed among the younger age group (less than 18 years). A comprehensive meta-analysis revealed a connection between asthma and an elevated risk of IPD. This study further illuminated that children within the age range of $0 - 17$ years who had asthma faced a greater susceptibility to IPD in comparison to adult patients aged 18 years and above (Li et al., 2020).

In this research, we determined the pneumonia transmission rates, which were found to be 0.81 for the SEIR compartmental model and 0.88 for the SEIRS model. It's worth noting that these rates are higher than the COVID-19 transmission rate in Manitoba, as reported by Amiri et al. (2023). The estimated incubation period and recovery period were reported as 0.48 and 0.57 for the SEIR model and 0.78 and 0.85 for the SEIRS model, respectively. This suggests that the incubation period is approximately 2 days. According to the Government of Canada, the incubation period for pneumococcal disease is not precisely defined but may be as short as 1 to 3 days (Public Health, 2023).

For certain infectious diseases, such as pneumonia, immunity is not lifelong and may wane over time. In our analysis, we found that 7.10% of cases became reinfected within one year of follow-up. This finding aligns with the result of other studies. For instance, a 5-year follow-up study reported a recurrence rate of pneumonia of 9% (Dang et al., 2014), while a 3-year follow-up study found a 9.4% rate of reinfection pneumonia (Garcia-Vidal et al., 2009). In our analysis, we calculated a pneumonia reinfection rate of 0.003, indicating that, on average, reinfections occurred every 333 days. A study conducted in Nova Scotia, Canada showed that the median time to reinfection pneumonia was 317 days, with an interquartile range from 177 days to 569 days (Dang et al., 2014).

4.2 Potential Implications

Research invariably yields benefits, either directly or indirectly, for humanity and society by advancing knowledge for future generations. This study, in particular, has potential implications at the community level. The findings help to take decisive action against the significant burden of disease within each RHAD community. The five major regional health departments in Manitoba can gain a better understanding of the distribution and spatial pattern of pneumonia. Healthcare providers can utilize the results of this study to plan and target high-risk regions within each RHAD, thereby providing intensive care to prevent the disease at the community level.

The age-specific findings assist health authorities in safeguarding vulnerable groups, such as children under five years of age and older demographics. Additionally, the identification of other potential risk factors guide the development of early prevention guidelines. This research also pave the way for designing new studies aimed at identifying zone-specific driving factors contributing to seasonal and spatial variations.

4.3 Strength and Limitation

The robustness of this research is underscored by a few key strengths. The use of high-quality, population-based data from the Manitoba province during the period of five years (2015 - 2019) provides an accurate and comprehensive picture of pneumonia infection rates. The inclusion of recent census data further bolsters the validity and relevance of our findings. The spatial Poisson regression model used for regional-level analysis allows for precise stratification of pneumonia rates. This result enhances our understanding of the geographical distribution and patterns of pneumonia, thereby informing targeted public health interventions. In parallel, the mixed-effect Poisson regression model applied to domain level data yields rigorous results, particularly in assessing the impact of comorbidities. These strengths collectively contribute to the depth and rigour of our research.

While conducting this study, certain limitations must be acknowledged. Our study was

constrained by the unavailability of laboratory-confirmed pneumonia cases. Access to such data would have allowed us to differentiate between viral and bacterial pneumonia, as well as to identify dominant serotypes. The absence of this information limits our ability to delve deeper into the specific pathogens responsible for pneumonia cases, which is a notable limitation. Furthermore, when measuring the disease transmission rate, we based our calculations on active cases. We considered a seven days recovery period for categorizing cases as active, which may not accurately represent all infected cases. Some individuals might experience early or late recoveries, and this could lead to variations in our transmission rate estimations.

4.4 Future Plan

In alignment with our research objectives, we have conducted a thorough analysis of the extracted data. However, our study was limited by the lack of access to laboratory-confirmed data. As part of our forthcoming research agenda, we intend to acquire and scrutinize laboratory-validated data, with a specific emphasis on discriminating between viral and bacterial pneumonia. The examination of serological trends and distribution is another critical aspect of our research, especially in the context of vaccine effectiveness. We intend to delve deeper into these trends, taking into account the spatial context.

In our current study, we estimated the disease transmission rate without considering individual-level characteristics. Moving forward, we plan to incorporate Geographically Dependent Individual Level Modeling into our research. This approach will allow us to consider the SEIRS compartmental model, thereby providing a more comprehensive understanding of disease transmission dynamics taking into account reinfection histories.

4.5 Conclusion

The study's findings unveiled substantial spatial variation in pneumonia incidence across Manitoba. Specifically, districts 3 and 15 in the Northern regions exhibited the highest incidence rates. Of note, individuals aged over 60 years were more vulnerable to pneumonia compared to other age groups. This vulnerability extended to immigrants, particularly those aged 60 and above. Moreover, the study highlighted that individuals with compromised immune systems due to comorbidities like COPD, CVD, IBD, and asthma faced an elevated risk of pneumonia. The research delved into assessing both the pneumonia transmission rate and reinfection rate. The pneumonia transmission rate was found to be higher than that of COVID-19. Conversely, the reinfection rate indicated an average reinfection interval of 333 days. These collective findings significantly contribute to our enhanced understanding of pneumonia dynamics within Manitoba.

Bibliography

Almirall, J., Bolibar, I., Vidal, J., Sauca, G., Coll, P., Niklasson, B., Bartolome, M., and Balanzo, X. (2000). Epidemiology of community-acquired pneumonia in adults: a population-based study. *European Respiratory Journal*, 15(4):757–763.

Amiri, L., Torabi, M., and Deardon, R. (2023). Analyzing COVID-19 data in the Canadian province of Manitoba: A new approach. *Spatial Statistics*, 55:100729.

Baik, I., Curhan, G. C., Rimm, E. B., Bendich, A., Willett, W. C., and Fawzi, W. W. (2000). A prospective study of age and lifestyle factors in relation to community-acquired pneumonia in US men and women. *Archives of Internal Medicine*, 160(20):3082–3088.

Beck, A. F., Florin, T. A., Campanella, S., and Shah, S. S. (2015). Geographic variation in hospitalization for lower respiratory tract infections across one county. *JAMA pediatrics*, 169(9):846–854.

Benincà, E., van Boven, M., Hagenaars, T., and van der Hoek, W. (2017). Space-time analysis of pneumonia hospitalisations in the Netherlands. *PloS One*, 12(7):e0180797.

Bernardinelli, L. and Montomoli, C. (1992). Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Statistics in Medicine*, 11(8):983–1007.

Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1):1–20.

Blain, A., Thomas, M., Shirley, M., Simmister, C., Elemraid, M., Gorton, R., Pearce, M., Clark, J., Rushton, S., and Spencer, D. (2014). Spatial variation in the risk of hospitalization with childhood pneumonia and empyema in the North of England. *Epidemiology & Infection*, 142(2):388–398.

Breslow, N. E. and Clayton, D. G. (1993). Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association*, 88(421):9–25.

Brownell, M., Chartier, M., Santos, R., Ekuma, O., Au, W., Sarkar, J., MacWilliam, L., Burland, E., Koseva, I., and Guenette, W. (2012). How are Manitoba's Children Doing? Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Cairns, C. and Kang, K. (2021). National hospital ambulatory medical care survey: 2021 emergency department summary tables. Available from: https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS/doc21-ed-508.pdf.

CDC (2022). Clinical Features of Pneumococcal Disease: CDC. https://www.cdc.gov/pneumococcal/clinicians/clinical-features.html. Accessed on October 16, 2023.

Chartier, M., Brownell, M., MacWilliam, L., Valdivia, J., Nie, Y., Ekuma, O., Burchill, C., Hu, M., Rajotte, L., and Kulbaba, C. (2016). The Mental Health of Manitoba's Children. Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Chartier, M., Finlayson, G., Prior, H., McGowan, K., Chen, H., de Rocquigny, J., Walld, R., and Gousseau, M. (2012). Health and Healthcare Utilization of Francophones in Manitoba. Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Chateau, D., Doupe, M., Prior, H., Soodeen, R., Sarkar, J., Dragan, R., Stevenson, D., and Rajotte, L. (2019). The Health Status of Community-Dwelling Older Adults in Manitoba. Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Chebib, N., Cuvelier, C., Malézieux-Picard, A., et al. (2021). Pneumonia prevention in the elderly patients: the other sides. *Aging Clinical and Experimental Research*, 33:1091–1100.

CIHI (2021). Hospital stays in Canada, Canadian Institute of Health Science. https://www.cihi.ca/en/hospital-stays-in-canada. Accessed on October 30, 2021.

Clayton, D. and Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, pages 671–681.

Coates, K. S., Bumsted, J. M., and McLintock, P. (2023). Manitoba. https://www.britannica.com/place/Manitoba.

Crighton, E., Elliott, S., Moineddin, R., Kanaroglou, P., and Upshur, R. (2007a). An exploratory spatial analysis of pneumonia and influenza hospitalizations in Ontario by age and gender. *Epidemiology & Infection*, 135(2):253–261.

Crighton, E. J., Elliott, S. J., Kanaroglou, P., Moineddin, R., and Upshur, R. E. (2008). Spatio-temporal analysis of pneumonia and influenza hospitalizations in Ontario, Canada. *Geospatial Health*, 2(2):191–202.

Crighton, E. J., Elliott, S. J., Moineddin, R., Kanaroglou, P., and Upshur, R. (2007b). A spatial analysis of the determinants of pneumonia and influenza hospitalizations in Ontario (1992–2001). *Social Science & Medicine*, 64(8):1636–1650.

Crighton, E. J., Moineddin, R., Mamdani, M., and Upshur, R. E. (2004). Influenza and pneumonia hospitalizations in Ontario: a time-series analysis. *Epidemiology & Infection*, 132(6):1167–1174.

Dang, T., Eurich, D., Weir, D., Marrie, T., and Majumdar, S. (2014). Rates and risk factors for recurrent pneumonia in patients hospitalized with community-acquired pneumonia: population-based prospective cohort study with 5 years of follow-up. *Clinical Infectious Diseases*, 59(1):74–80.

Dang, T. T., Majumdar, S. R., Marrie, T. J., and Eurich, D. T. (2015). Recurrent Pneumonia: A Review with Focus on Clinical Epidemiology and Modifiable Risk Factors in Elderly Patients. *Drugs & Aging*, 32(1):13–19.

Diggle, P. J., Tawn, J. A., and Moyeed, R. A. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 47(3):299–350.

Dunne, E., Nunes, M., Slack, M., et al. (2023). Effects of pneumococcal conjugate vaccines on reducing the risk of respiratory disease associated with coronavirus infection. *Pneumonia*, 15:10.

Elliott, P. et al. (1996). *Geographical and Environmental Epidemiology: Methods for Small Area Studies*. Oxford University Press, online edn, Oxford Academic, 1 sept. 2009 edition.

File Jr, T. M. and Marrie, T. J. (2010). Burden of community-acquired pneumonia in North American adults. *Postgraduate Medicine*, 122(2):130–141.

Finlayson, G. S., Reimer, J., Dahl, M., Stargardter, M., and McGowan, K.-L. (2009). The direct cost of hospitalizations in Manitoba. *Policy*.

Fransoo, R., Mahar, A., The Need to Know Team, Anderson, A., Prior, H., Koseva, I., McCulloch, S., Jarmasz, J., and Burchill, S. (2019). The 2019 RHA Indicators Atlas. Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Fransoo, R., Martens, P., Burland, E., to Know Team, T. N., Prior, H., and Burchill, C. (2009). Manitoba RHA Indicators Atlas 2009. Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Fransoo, R., Martens, P., to Know Team, T. N., Prior, H., Burchill, C., Koseva, I., Bailly, A., and Allegro, E. (2013). The 2013 RHA Indicators Atlas. Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Garcia-Vidal, C., Carratala, J., Fernandez-Sabe, N., Dorca, J., Verdaguer, R., Manresa, F., and Gudiol, F. (2009). Aetiology of, and risk factors for, recurrent community-acquired pneumonia. *Clinical Microbiology and Infection*, 15(11):1033–1038.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Government of Canada (2019). Update on the Use of Pneumococcal Vaccines in Adults. https://www.canada.ca/en/public-health/services/publications/healthy-living/update-on-the-use-of-pneumococcal-vaccines-in-adult.html. Accessed on October 23, 2023.

Government of Canada (2023). Canadian Immunization Guide - Part 4: Active Vaccines. https://www.canada.ca/en/public-health/services/publications/healthy-living/

canadian-immunization-guide-part-4-active-vaccines/page-16-pneumococcal-vaccine.html. Accessed on October 23, 2023.

Government of Manitoba (2022). Annual Report 2021-2022. https://www.gov.mb.ca/health/annualreports/docs/2122.pdf.

Iroh, T. P., Krzyzanowski, B., Oakes, J., Kne, L., and Manson, S. (2017). Spatial variation of pneumonia hospitalization risk in twin cities metro area, minnesota. *Epidemiology & Infection*, 145(15):3274–3283.

Janson, C., Johansson, G., Ställberg, B., et al. (2018). Identifying the associated risks of pneumonia in COPD patients: ARCTIC an observational study. *Respiratory Research*, 19:172.

Kaplan, V., Clermont, G., Griffin, M. F., Kasal, J., Watson, R. S., Linde-Zwirble, W. T., and Angus, D. C. (2003). Pneumonia: Still the Old Man's Friend? *Archives of Internal Medicine*, 163(3):317–323.

Kassam, S., Serrano-Lomelin, J., Hicks, A., Crawford, S., Bakal, J. A., and Ospina, M. B. (2021). Geography as a Determinant of Health: Health Services Utilization of Pediatric Respiratory Illness in a Canadian Province. *International Journal of Environmental Research and Public Health*, 18(16).

Kermack, W. and McKendrick, A. (1927). A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A*, 115(772):700–721.

Kim, A. M., Kang, S., Park, J. H., Yoon, T. H., and Kim, Y. (2019). A spatial analysis of geographic variation and factors associated with hospitalization for bacterial pneumonia in Korea. *BMC Pulmonary Medicine*, 19(1):45.

Kim, P. E., Musher, D. M., Glezen, W. P., Rodriguez-Barradas, M. C., Nahm, W. K., and Wright, C. E. (1996). Association of invasive pneumococcal disease with season, atmospheric conditions, air pollution, and the isolation of respiratory viruses. *Clinical Infectious Diseases*, 22(1):100–6.

Kizito, M. and Tumwiine, J. (2018). A mathematical model of treatment and vaccination interventions of pneumococcal pneumonia infection dynamics. *Journal of Applied Mathematics*, 2018.

Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine*, 19(17-18):2555–2567.

Kumar, P., Balasubramanian, R., Chellappan, S., and Ramalingam, S. (2013). A mathematical model for the transmission dynamics of tuberculosis with case detection rate. *Computational and Mathematical Methods in Medicine*, 2013:816390.

Lange, P., Vestbo, J., and Nyboe, J. (1995). Risk factors for death and hospitalization from pneumonia. a prospective study of a general population. *European Respiratory Journal*, 8(10):1694–1698.

Li, L., Cheng, Y., Tu, X., Yang, J., Wang, C., Zhang, M., and Lu, Z. (2020). Association between asthma and invasive pneumococcal disease risk: a systematic review and meta-analysis. *Allergy, Asthma & Clinical Immunology*, 16:1–9.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22.

Lipsitch, M., Cohen, T., Cooper, B., Robins, J. M., Ma, S., James, L., Gopalakrishna, G., Chew, S. K., Tan, C. C., Samore, M. H., Fisman, D. N., and Murray, M. (2003). Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 300(5627):1966–1970.

Lix, L., Derksen, S., Sirski, M., et al. (2021). Gastrointestinal Endoscopy (GIE) Utilization in Manitoba. Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Loeb, M. B. (2003). Community-acquired pneumonia in older people: the need for a broader perspective. *Journal of the American Geriatrics Society*, 51(4):539–43.

Long, Millie D MD, M., Martin, C. M., Sandler, Robert S MD, M., and Kappelman, Michael D MD, M. (2013). Increased Risk of Pneumonia Among Patients With Inflammatory Bowel Disease. *American Journal of Gastroenterology*, 108(2):240–248.

MacNab, Y. C. and Dean, C. B. (2001). Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. *Biometrics*, 57(3):949–56.

Manitoba Health (2014). Annual statistics 2013–2014. Accessed on April 5, 2023.

Manitoba Health (2015). Annual statistics 2014–2015. Accessed on April 5, 2023.

Manitoba Health (2016). Annual statistics 2015–2016. Accessed on April 5, 2023.

Manitoba Health (2017). Annual statistics 2016–2017. Accessed on April 5, 2023.

Manitoba Health (2018). Annual statistics 2017–2018. Accessed on April 5, 2023.

Manitoba Health (2019). Annual statistics 2018–2019. Accessed on April 5, 2023.

Martens, P., Bartlett, J., Burland, E., Prior, H., Burchill, C., Huq, S., Romphf, L., Sanguins, J., Carter, S., and Bailly, A. (2010). Profile of Metis Health Status and Healthcare Utilization in Manitoba: A Population-Based Study. Technical report, Manitoba Centre for Health Policy, Winnipeg, MB.

Martens, P., Nickel, N., Forget, E. L., Lix, L., Turner, D., Prior, H., Walld, R., Soodeen, R.-A., Rajotte, L., and Ekuma, O. (2015). *The Cost of Smoking: A Manitoba Study*. Manitoba Centre for Health Policy, Winnipeg, MB.

MCHP (2013). Concept: Regional Health Authorities (RHA) in Manitoba. http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?conceptID=1218. Accessed on October 30, 2021.

MCHP (2023). Manitoba Population Research Data Repository Data Descriptions: Hospital Abstract Data. http://mchp-appserv.cpe.umanitoba.ca/dataDescriptions.php?ds=Hospital. Accessed on October 30, 2023.

Metge, C., Chateau, D., Prior, H., Soodeen, R., De Coster, C., and Barré, L. (2009). Composite Measures/Indices of Health and Health System Performance. Technical report, Manitoba Centre for Health Policym Winnipeg, MB.

Mouton, C. P., Bazaldua, O. V., Pierce, B., and Espino, D. V. (2001). Common infections in older adults. *Amirican Family Physician*, 63(2):257–269.

Ndelwa, E., Kgosimore, M., Massawe, E., and Namkinga, L. (2015). Mathematical modelling and analysis of treatment and screening of pneumonia. *Mathematical Theory and Modeling*, 5(10):21–39.

NIH (2022). National Heart, Lung, and Blood Institute: Recovery after Pneumonia. https://www.nhlbi.nih.gov/health/pneumonia/recovery. Accessed on October 16, 2023.

Nuorti, J. P., Butler, J. C., Farley, M. M., Harrison, L. H., McGeer, A., Kolczak, M. S., Breiman, R. F., and Team, A. B. C. S. (2000). Cigarette smoking and invasive pneumococcal disease. *New England Journal of Medicine*, 342(10):681–689.

Okafor, P. N., Nunes, D. P., and Farraye, F. A. (2013). Pneumocystis Jiroveci Pneumonia in Inflammatory Bowel Disease: When Should Prophylaxis be Considered? *Inflammatory Bowel Diseases*, 19(8):1764–1771.

Ong'ala, J. O., Mugisha, J., and Odhiambo, J. W. (2015). Mathematical model for pneumonia dynamics among children. *Southern Africa Mathematical Sciences Association Conference*.

Panel, E. R. (2023). Pneumonia treatment and recovery: American lung assocation. https://www.lung.org/lung-health-diseases/lung-disease-lookup/pneumonia/treatment-and-recovery. Accessed on October 16, 2023.

Plouffe, J. F., Breiman, R. F., Facklam, R. R., Baird, I., Barnishan, J., Porterfield-Baxa, J., Best, M., Dalieh, S., Emerick, J., Fass, R. J., et al. (1996). Bacteremia with Streptococcus pneumoniae: implications for therapy and prevention. *Journal of the American Medical Association*, 275(3):194–198.

Population Health Surveillance Team, WRHA (2020). The Epidemiology of Communicable Diseases in the Winnipeg Health Region, 2014-2019. Report.

Prentice, R. L. and Zhao, L. P. (1991). Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics*, 47(3):825–839.

Public Health (2023). Invasive Pneumococcal Disease (IPD) in Canada: Information for Health Professionals. https://www.canada.ca/en/public-health/services/immunization/vaccine-preventable-diseases/invasive-pneumococcal-disease/health-professionals.html. Accessed on October 16, 2023.

Saynajakangas, P., Keistinen, T., and Tuuponen, T. (2001). Seasonal Fluctuations in Hospitalisation for Pneumonia in Finland. *International Journal of Circumpolar Health*, 60:34–40.

Sierra, C. and Daiya, K. (2022). Prophylaxis for Pneumocystis jirovecii pneumonia in patients with inflammatory bowel disease: A systematic review. *Pharmacotherapy*, 42:858–867.

Silva, G. L., Dean, C. B., Niyonsenga, T., and Vanasse, A. (2008). Hierarchical Bayesian spatiotemporal analysis of revascularization odds using smoothing splines. *Statistics in Medicine*, 27(13):2381–401.

Statistics Canada (2021a). Population and dwelling counts, for Canada, provinces and territories, 2021 and 2016 censuses. https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/az/definition-eng.cfm?ID=pop148. Accessed on October 30, 2022.

Statistics Canada (2021b). Profile table, Census Profile, 2021 Census of Population. https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/details/page.cfm?Lang=E&SearchText=Manitoba&DGUIDlist=2021A000246&GENDERlist=1,2,3&STATISTIClist=1&HEADERlist=0. Accessed on April 4, 2023.

Statistics Canada (2022). Population by year, by province and territory. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501. Accessed on February 22, 2022.

Sutton, K. L., Banks, H., and Castillo-Chavez, C. (2010). Public vaccination policy using an age-structured model of pneumococcal infection dynamics. *Journal of Biological Dynamics*, 4(2):176–195.

Søgaard, M., Madsen, M., Løkke, A., Hilberg, O., Sørensen, H., and Thomsen, R. (2016). Incidence and outcomes of patients hospitalized with COPD exacerbation with and without pneumonia. *International Journal of Chronic Obstructive Pulmonary Disease*, 11(1):455–465.

Temime, L., Guillemot, D., and Boëlle, P. Y. (2004). Short- and long-term effects of pneumococcal conjugate vaccination of children on penicillin resistance. *Antimicrobial Agents and Chemotherapy*, 48(6):2206–2213.

Tilahun, G. T., Makinde, O. D., and Malonza, D. (2017). Modelling and optimal control of pneumonia disease with cost-effective strategies. *Journal of Biological Dynamics*, 11(sup2):400–426.

Torabi, M. and Jiang, J. (2020). Estimation of mean squared prediction error of empirically spatial predictor of small area means under a linear mixed model. *Journal of Statistical Planning and Inference*, 208:82–93.

Torabi, M. and Rosychuk, R. J. (2010). Spatio-temporal modelling of disease mapping of rates. *Canadian Journal of Statistics*, 38(4):698–715.

Torabi, M. and Rosychuk, R. J. (2011). Spatio-temporal modelling using b-spline for disease mapping: analysis of childhood cancer trends. *Journal of Applied Statistics*, 38(9):1769–1781.

Torabi, M. and Rosychuk, R. J. (2012). Hierarchical bayesian spatiotemporal analysis of childhood cancer trends. *Geographical Analysis*, 44(2):109–120.

Waller, L. A., Carlin, B. P., Xia, H., and Gelfand, A. E. (1997). Hierarchical Spatio-Temporal Mapping of Disease Rates. *Journal of the American Statistical Association*, 92(438):607–617.

WHO (2019). ICD-10 Version: 2019 - Pneumonia (J12-J18). https://icd.who.int/browse10/2019/en#/J12. Accessed on October 30, 2021.

WHO (2019). News room, pneumonia. https://www.who.int/news-room/fact-sheets/detail/pneumonia. Accessed on October 24, 2021.

Winterbauer, R. H., Bedon, G. A., and Ball, J. W. C. (1969). Recurrent Pneumonia Predisposing Illness and Clinical Patterns in 158 Patients. *Annals of Internal Medicine*, 70(4):689–700.

Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika*, 75(4):621–629.

Appendices

Appendix A: Regional Level Analysis

Table A.1: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	0.270	0.270	[0.169, 0.369]
Age (0-5)	- 1.286	- 1.299	[- 8.481, 5.818]
Age (6-59)	17.030	18.290	[- 18.290, 47.730]
Age (60 and above)	25.710	25.650	[18.590, 33.280]
SEFI 2	0.1002	0.0998	[0.0123, 0.1912]
Immigration Status	2.467	2.467	[-0.1687, 5.031]
Full Vaccination History	1.287	1.304	[-0.2217, 2.758]
Age (0-5)*Immigration Status	- 1.758	- 1.774	[- 4.407, 1.015]
Age (6-59)*Immigration Status	- 2.434	- 2.008	[- 15.600, 8.955]
Age (60 and above)*Immigration Status	5.647	5.625	[3.071, 8.254]

Table A.2: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	-1.49	-1.49	[-1.91, -0.99]
Age (0-5)	35.55	35.75	[22.32, 47.22]
Age (6-59)	23.83	23.83	[8.061, 39.77]
Age (60 and above)	9.042	9.029	[5.159, 12.97]
SEFI 2	0.04595	0.04542	[-0.02811, 0.122]
Immigration Status	-9.919	-9.894	[-13.76, -6.398]
Full Vaccination History	4.538	4.563	[2.568, 6.273]
Age (0-5)*Immigration Status	-3.28	-3.29	[-5.75, -0.77]
Age (6-59)*Immigration Status	-0.69	-0.67	[-11.37, 9.62]
Age (60 and above)*Immigration Status	5.46	5.45	[3.33, 7.68]
Age (0-5)*Full Vaccination History	-1.01	-1.02	[-1.36, -0.62]

Table A.3: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	-2.17	-2.15	[-3.02, -1.55]
Age (0-5)	0.33	1.09	[-34.50, 28.96]
Age (6-59)	27.26	27.25	[12.23, 42.49]
Age (60 and above)	47.39	47.14	[19.86, 88.64]
SEFI 2	0.05	0.05	[-0.02, 0.12]
Immigration Status	-7.01	-6.94	[-11.45, -2.75]
Full Vaccination History	6.96	6.89	[4.70, 10.01]
Age (0-5)*Immigration Status	-1.13	-1.18	[-4.00, 1.96]
Age (6-59)*Immigration Status	-3.08	-2.98	[-13.61, 7.14]
Age (60 and above)*Immigration Status	3.71	3.68	[1.20, 6.29]
Age (0-5)*Full Vaccination History	0.29	0.27	[-0.72, 1.54]
Age (60 and above)*Full Vaccination History	-1.38	-1.36	[-2.84, -0.40]

Table A.4: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	-2.25	-2.26	[-2.76, -1.71]
Age (0-5)	-15.43	-15.80	[-34.80, 4.12]
Age (6-59)	86.23	84.47	[18.83, 167.10]
Age (60 and above)	49.42	49.41	[25.62, 72.45]
SEFI 2	0.05	0.05	[-0.02, 0.12]
Immigration Status	-5.11	-5.12	[-8.31, -1.91]
Full Vaccination History	7.12	7.14	[4.97, 9.11]
Age (60 and above)*Immigration Status	2.12	2.12	[1.00, 3.25]
Age (0-5)*Full Vaccination History	0.84	0.85	[0.11, 1.56]
Age (6-59)*Full Vaccination History	-1.89	-1.82	[-4.38, 0.17]
Age (60 and above)*Full Vaccination History	-1.44	-1.44	[-2.27, -0.59]

Table A.5: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	-2.27	-2.26	[-2.85, -1.71]
Age (0-5)	-9.42	-10.37	[-26.64, 10.12]
Age (6-59)	25.76	25.79	[11.50, 39.79]
Age (60 and above)	56.10	57.16	[37.13, 73.47]
SEFI 2	0.06	0.06	[-0.01, 0.13]
Immigration Status	-5.41	-5.41	[-8.45, -2.20]
Full Vaccination History	7.28	7.27	[5.13, 9.45]
Age (60 and above)*Immigration Status	2.15	2.16	[1.01, 3.26]
Age (0-5)*Full Vaccination History	0.62	0.66	[-0.11, 1.26]
Age (60 and above)*Full Vaccination History	-1.66	-1.70	[-2.30, -0.99]

Table A.6: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	0.33	0.33	[0.26, 0.39]
Age (0-5)	3.18	3.18	[-2.47, 8.77]
Age (6-59)	25.48	25.49	[11.15, 39.87]
Age (60 and above)	25.33	25.35	[19.98, 30.49]
SEFI 2	0.05	0.05	[-0.02, 0.12]
Immigration Status	2.06	2.08	[0.00, 4.07]
Full Vaccination History	2.77	2.76	[1.52, 4.04]
Age (0-5)*Immigration Status	-1.49	-1.50	[-3.76, 0.76]
Age (60 and above)*Immigration Status	3.23	3.24	[1.11, 5.28]
Age (0-5)*Full Vaccination History	0.31	0.32	[-0.62, 1.15]
Age (60 and above)*Full Vaccination History	-1.40	-1.41	[-2.23, -0.50]

Table A.7: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	0.326	0.326	[0.258, 0.398]
Age (0-5)	3.235	3.277	[-2.709, 9.164]
Age (6-59)	25.430	25.280	[11.500, 39.370]
Age (60 and above)	25.320	25.310	[20.220, 30.360]
SEFI 2	0.047	0.048	[-0.029, 0.121]
Immigration Status	2.054	2.041	[-0.042, 4.161]
Full Vaccination History	2.771	2.774	[1.520, 4.003]
Age (0-5)*Immigration Status	-1.483	-1.495	[-3.778, 0.843]
Age (60 and above)*Immigration Status	3.257	3.262	[1.068, 5.409]
Age (0-5)*Full Vaccination History	0.302	0.307	[-0.655, 1.187]
Age (60 and above)*Full Vaccination History	-1.398	-1.401	[-2.273, -0.445]
SEFI 2*Immigration Status	-0.001	-0.001	[-0.022, 0.019]

Table A.8: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	-2.11	-2.10	[-2.66, -1.59]
Age (0-5)	5.84	5.84	[1.01, 10.73]
Age (6-59)	23.52	23.53	[8.79, 38.27]
Age (60 and above)	40.56	40.58	[30.22, 52.03]
SEFI 2	0.05	0.05	[-0.016, 0.12]
Immigration Status	-6.13	-6.14	[-9.48, -2.79]
Full Vaccination History	6.74	6.72	[4.73, 8.85]
Age (60 and above)*Immigration	246.10	246.80	[124.90, 366.60]
Age(60 and above)*Full vaccine	-109	-109.1	[-147.10, -75.09]
Comorbidity: IBD	-3.46	-3.55	[-104.40, 95.04]
Comorbidity: Asthma	24.59	24.29	[-25.47, 75.41]

Table A.9: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	-1.98	-2.01	[-2.47, -1.34]
Age (0-5)	6.61	6.62	[1.79, 11.40]
Age (6-59)	16.19	16.05	[1.42, 31.44]
Age (60 and above)	40.30	40.69	[28.46, 49.47]
SEFI 2	0.04	0.04	[-0.03, 0.11]
Immigration Status	-5.78	-5.71	[-9.16, -2.82]
Full Vaccination History	6.82	6.88	[4.46, 8.71]
Age (60 and above)*Immigration	2.26	2.23	[1.21, 3.40]
Age(60 and above)*Full vaccine	-1.08	-1.10	[-1.40, -0.70]
Comorbidity: IBD	-148.00	-150.30	[-285.30, -3.99]
Comorbidity: Asthma	14.00	13.63	[-33.54, 62.15]
IBD*Age(6 to 59)	9.61	9.66	[2.05, 16.67]

Table A.10: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	-1.77	-1.78	[-2.22, -1.33]
Age (0-5)	5.71	5.734	[1.931, 9.562]
Age (6-59)	10.3	10.32	[-4.759, 25.41]
Age (60 and above)	29.9	29.94	[21.35, 38.86]
SEFI 2	0.0376	0.03735	[-0.02355, 0.09883]
Immigration Status	-4.261	-4.276	[-6.884, -1.652]
Full Vaccination History	4.777	4.776	[3.127, 6.521]
Age (60 and above)*Immigration	178.1	178.4	[84.96, 269.9]
Age(60 and above)*Full vaccine	-71.11	-71.27	[-101.2, -42.11]
Comorbidity: IBD	-97.37	-98.39	[-216.6, 24.3]
Comorbidity: Asthma	-10.31	-10.12	[-52.81, 31.64]
Comorbidity: COPD	12.68	12.69	[3.324, 21.69]
Comorbidity: CVD	12.4	12.35	[3.979, 20.78]
Comorbidity: IBD*Age (6 to 59)	9.045	9.158	[-6.846, 24.48]

Table A.11: IRR of covariates and corresponding 95% credible interval of spatial Poisson regression model

Variable	Mean	Median	95% CI
Intercept	-1.77	-1.78	[-2.22, -1.31]
Age (0-5)	5.23	5.21	[1.66, 8.86]
Age (6-59)	12.70	12.69	[0.00, 25.69]
Age (60 and above)	30.27	30.27	[21.74, 39.20]
SEFI 2	0.04	0.04	[-0.02, 0.10]
Immigration Status	-4.50	-4.50	[-7.21, -1.81]
Full Vaccination History	4.76	4.78	[2.93, 6.41]
Age (60 and above)*Immigration	1.85	1.85	[0.90, 2.80]
Age(60 and above)*Full vaccine	-0.73	-0.73	[-1.02, -0.43]
Comorbidity: IBD	-93.20	-92.94	[-220.90, 32.01]
Comorbidity: COPD/Asthma	12.30	12.35	[2.99, 21.51]
Comorbidity: CVD	11.57	11.58	[2.95, 20.15]
IBD*Age(6 to 59)	3.22	3.21	[-3.48, 10.02]

Appendix B: Domain Level Analysis

Table B.1: Pneumonia infection and IBD: results from a Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.011	[-5.242, -4.779]
IBD	0.273	[0.168, 0.379]
Scale	10.884	[10.884, 10.884]
AIC	1965	
BIC	1987	

Table B.2: Pneumonia infection and vaccine: results from a Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.868	[-5.025, -4.711]
vaccine	0.026	[0.023, 0.028]
Scale	5.102	[5.102, 5.102]
AIC	4244	
BIC	4251	

Table B.3: Pneumonia infection and asthma: results from a Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-4.849	[-5.094, -4.604]
Asthma	0.0.054	[-0.006, 0.114]
Scale	11.827	[11.827, 11.827]
AIC	18260	
BIC	18267	

Table B.4: Pneumonia infection and COPD: results from a Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.151	[-5.378, -4.924]
COPD	0.311	[0.233, 0.388]
Scale	10.113	[10.113, 10.113]
AIC	13578	
BIC	13585	

Table B.5: Pneumonia infection and CVD: results from a Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.302	[-5.475, -4.129]
CVD	0.188	[0.164, 0.213]
Scale	7.498	[7.498, 7.498]
AIC	9348	
BIC	9355	

Table B.6: Pneumonia infection, vaccine and comorbidity: results from a Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.863	[-6.031, -5.694]
IBD	-0.01	[-0.069, 0.049]
Full vaccination	0.0239	[0.021, 0.027]
Asthma	0.007	[-0.022, 0.036]
COPD	-0.015	[-0.068, 0.037]
CVD	0.032	[0.000, 0.065]
Scale	5.034	[5.034, 5.034]
AIC	4143	
BIC	4163	

Table B.7: Pneumonia infection and comorbidity: results from a Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.695	[-5.918, -5.472]
IBD	0.077	[-0.003, 0.157]
Asthma	0.101	[0.062, 0.140]
COPD	0.102	[0.032, 0.171]
CVD	0.170	[0.140, 0.200]
Scale	7.054	[7.054, 7.054]
AIC	7645	
BIC	7662	

Table B.8: Pneumonia infection, vaccine, and comorbidity: results from a negative binomial regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.347	[-5.538, -5.156]
IBD	-0.083	[-0.174, 0.007]
Full vaccination	0.0256	[0.019, 0.032]
Asthma	-0.065	[-0.111, -0.020]
COPD	-0.084	[-0.171, 0.002]
CVD	0.049	[-0.008, 0.107]
Dispersion	0.578	[0.468, 0.712]
AIC	2191	
BIC	2214	

Table B.9: Pneumonia infection and comorbidity: results from a negative binomial regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.116	[-5.332, -4.900]
IBD	0.010	[-0.093, 0.114]
Asthma	0.033	[-0.007, -0.074]
COPD	0.043	[-0.054, 0.141]
CVD	0.216	[0.168, 0.264]
Dispersion	0.781	[0.646, 0.944]
AIC	2245	
BIC	2265	

Table B.10: Pneumonia infection, vaccine, and comorbidity: results from a mixed-effect a Poisson regression model for age-sex group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.848	[-6.011, -5.686]
IBD	-0.018	[-0.107, 0.070]
Full vaccination	0.025	[0.019, 0.032]
Asthma	-0.019	[-0.065, 0.027]
COPD	-0.038	[-0.120, 0.042]
CVD	0.047	[-0.003, 0.099]
σ	0.545	[0.424, 0.666]
AIC	2123	
BIC	2247	

Table B.11: Pneumonia infection, vaccine, and comorbidity: results from a mixed-effect Poisson regression model for age-group, nested sex data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.483	[-5.724, -5.242]
IBD	-0.013	[-0.018, 0.045]
Full vaccination	0.013	[0.008, 0.017]
Asthma	0.006	[-0.015, 0.028]
COPD	0.023	[-0.006, 0.052]
CVD	0.047	[0.030, 0.064]
σ_u	0.718	[0.483, 0.954]
σ_v	0.024	[0.008, 0.040]
AIC	1965	
BIC	1987	

Table B.12: Pneumonia infection with asthma and COPD in children up to 17 years: mixed-effect Poisson regression model for age-group, nested sex data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-6.032	[-6.527, -5.537]
Asthma	0.053	[-0.014, 0.120]
COPD	0.023	[-0.006, 0.052]
σ_u	0.751	[0.096, 1.407]
σ_v	0.033	[-0.031, 0.099]
AIC	312	
BIC	316	

Note: No IBD and CVD cases in children up to 17 years.

Table B.13: Pneumonia infection with asthma and COPD in children up to 17 years: mixed-effect Poisson regression model for age-sex group data, with age as a random variable

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.092	[-6.420, -5.429]
Asthma	0.023	[0.004, 0.042]
COPD	0.005	[-0.142, 0.152]
σ	0.942	[0.254, 1.629]
AIC	316	
BIC	320	

Note: No IBD and CVD cases in children up to 17 years.

Table B.14: Pneumonia infection with asthma and COPD in children aged 18 years and above: results from a mixed-effect Poisson regression model for age-grouped sex data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-4.822	[-5.066, -4.577]
Asthma	-0.0002	[-0.022, 0.022]
COPD	0.026	[0.003, 0.051]
IBD	0.033	[0.008, 0.059]
CVD	0.049	[0.035, 0.063]
σ_u	1.190	[0.816, 1.564]
σ_v	0.012	[0.002, 0.022]
AIC	1681	
BIC	1699	

Table B.15: Pneumonia infection with asthma and COPD in children aged 18 years and above: results from a mixed-effect Poisson regression model for age-sex group data, where age is a random variable

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-4.780	[-5.021, -5.438]
Asthma	-0.005	[-0.021, 0.009]
IBD	0.020	[0.005, 0.036]
COPD	0.028	[0.011, 0.044]
COPD	0.043	[0.034, 0.052]
σ	1.232	[0.852, 1.622]
AIC	1696	
BIC	1711	

Table B.16: Pneumonia infection with asthma and COPD in children up to 17 years: results from a mixed-effect Poisson regression model for the age group, where age is a random variable

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-6.619	[-6.918, -5.320]
Male	0.010	[0.006, 0.015]
Asthma	0.027	[-0.011, 0.066]
COPD	0.114	[-0.231, 0.460]
σ	0.143	[0.029, 0.257]
AIC	171	
BIC	175	

Note: No IBD and CVD cases for children up to 17 years old

Table B.17: Pneumonia infection with asthma and COPD for children aged 18 and above: results from a mixed-effect Poisson regression model for age group data, where age is a random variable

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.267	[-5.625, -4.910]
Male	-0.006	[-0.018, 0.005]
Asthma	-0.055	[-0.142, 0.032]
COPD	0.032	[-0.068, 0.133]
IBD	0.074	[-0.038, 0.188]
CVD	0.147	[-0.091, 0.204]
σ	0.601	[0.410, 0.791]
AIC	1034	
BIC	1051	

Table B.18: Pneumonia infection, vaccine, and comorbidity: results from a Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.493	[-6.141, -5.643]
Male	0.002	[-0.003, 0.006]
IBD	0.012	[-0.050, 0.074]
Full vaccination	0.013	[0.007, 0.017]
Asthma	-0.010	[-0.045, 0.024]
COPD	-0.036	[-0.092, 0.019]
CVD	0.010	[-0.027, 0.048]
Scale	6.653	[6.653, 6.653]
AIC	3257	
BIC	3276	

Table B.19: Pneumonia infection, sex, and comorbidity: results from a Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-6.022	[-6.332, -5.712]
Male	0.009	[0.004, 0.014]
IBD	0.044	[-0.034, 0.123]
Asthma	0.005	[-0.039, 0.050]
COPD	-0.021	[-0.092, 0.048]
CVD	0.083	[-0.051, 0.115]
Scale	8.295	[8.295, 8.295]
AIC	4373	
BIC	4389	

Table B.20: Pneumonia infection, sex, vaccine, and comorbidity: results from a mixed-effect Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.776	[-6.021, -5.530]
Male	-0.0001	[-0.0075, 0.0072]
IBD	-0.013	[-0.107, 0.060]
Full vaccination	0.014	[0.007, 0.021]
Asthma	-0.024	[-0.075, 0.026]
COPD	-0.057	[-0.145, 0.029]
CVD	0.018	[-0.040, 0.077]
σ	0.541	[0.380, 0.702]
AIC	1214	
BIC	1235	

Table B.21: Pneumonia infection, sex, and comorbidity: results from a mixed-effect Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.851	[-6.114, -5.589]
Male	0.008	[0.001, 0.014]
IBD	0.012	[-0.088, 0.114]
Asthma	-0.003	[-0.0588, 0.050]
COPD	-0.036	[-0.130, 0.058]
CVD	0.107	[0.062, 0.151]
σ	0.640	[0.450, 0.828]
AIC	1228	
BIC	1247	

Table B.22: Pneumonia infection with sex, IBD, COPD, and CVD: results from a mixed-effect Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.854	[-6.114, -5.595]
Male	0.008	[0.003, 0.013]
IBD	0.014	[-0.086, 0.114]
COPD	-0.035	[-0.129, 0.058]
CVD	0.108	[0.070, 0.146]
σ	0.640	[0.453, 0.828]
AIC	1227	
BIC	1242	

Table B.23: Pneumonia infection with sex, IBD, asthma, and CVD: results from a mixed-effect Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.849	[-6.112, -5.586]
Male	0.007	[0.000, 0.014]
IBD	0.013	[-0.095, 0.098]
Asthma	-0.001	[-0.055, 0.053]
CVD	0.109	[0.059, 0.142]
σ	0.644	[0.456, 0.833]
AIC	1227	
BIC	1243	

Table B.24: Pneumonia infection with sex, vaccine, IBD, asthma, and CVD: results from a mixed-effect Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.775	[-6.022, -5.537]
Male	0.0008	[-0.0008, 0.0066]
IBD	-0.030	[-0.122, 0.060]
Full vaccine	0.014	[0.007, 0.021]
Asthma	-0.019	[-0.070, 0.031]
CVD	0.011	[-0.046, 0.070]
σ	0.551	[0.387, 0.714]
AIC	1214	
BIC	1232	

Table B.25: Pneumonia infection with sex, vaccine, IBD, COPD, and CVD: results from a mixed-effect Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.797	[-6.039, -5.555]
Male	-0.002	[-0.008, 0.004]
IBD	-0.005	[-0.098, 0.087]
Full vaccine	0.014	[0.007, 0.020]
COPD	-0.052	[-0.319, 0.034]
CVD	0.032	[-0.018, 0.083]
σ	0.545	[0.383, 0.707]
AIC	1213	
BIC	1232	

Table B.26: Pneumonia infection with sex, IBD, vaccine and CVD: results from a mixed-effect Poisson regression model for age group data

Parameter	Estimate, $\hat{\beta}$	95% CI
Intercept	-5.793	[-6.036, -5.549]
Male	-0.002	[-0.008, 0.004]
IBD	-0.022	[-0.111, 0.066]
Full vaccine	0.013	[0.006, 0.020]
CVD	0.023	[-0.025, 0.073]
σ	0.553	[0.389, 0.718]
AIC	1212	
BIC	1228	