INVESTIGATION OF PERTUSSIS RESURGENCE IN KING COUNTY, WASHINGTON

Madhura S. Rane

A dissertation
submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

University of Washington
2020

Reading Committee:
M. Elizabeth Halloran, Chair
Pejman Rohani
Amanda Phipps

Program Authorized to Offer Degree:
Epidemiology

University of Washington

ABSTRACT

INVESTIGATION OF PERTUSSIS RESURGENCE IN KING COUNTY, WASHINGTON

Madhura S. Rane

Chair of the Reading Committee:
Professor M. Elizabeth Halloran
Departments of Biostatistics and Epidemiology

Introduction: Pertussis or "whooping cough" is an infectious disease caused by bacterium *Bordetella pertussis*. Despite widespread vaccination since the 1950s, it continues to cause epidemics in several countries. In fact, in the last three decades, many countries with high vaccination coverage have reported a resurgence in pertussis activity. Understanding the factors that led to pertussis resurgence and its persistence is important because pertussis remains a leading cause of infant mortality worldwide with close to 100,000 deaths annually (WHO, 2018). This study examines the role of the Diphtheria-Tetanus-acellular-Pertussis (DTaP) vaccination coverage and vaccine failure in pertussis resurgence and persistence. Specifically, we described long-term, fine-scale spatio-temporal trends of pertussis incidence and characterized spatial dependence in pertussis cases within King County (Aim 1); estimated the association between pertus-

sis epidemics and area-level vaccination coverage and socio-demographic factors (Aim 2); assessed the role of vaccine schedules and vaccine timeliness on pertussis incidence (Aim 3); estimated the population-level effects of acellular pertussis vaccines (Aim 4); and estimated waning effects of the acellular pertussis vaccine using Schoenfeld residuals (Aim 5)

Methods: We obtained pertussis incidence data for all ages for the period between 1999 and 2017 from the Communicable Disease and Immunizations Department within Public Health Seattle and King County. Diphtheria-Tetanus-acellular-Pertussis (DTaP) vaccination records, as well as records for all pediatric vaccinations, for all children born or living in King County, WA, between 2008 and 2017 were obtained from the Washington State Immunization Information System (WA-IIS) maintained by the Washington State Department of Health. All five aims use information from either one or both datasets. These datasets were linked using probabilistic linkage methods to obtain DTaP vaccination and pertussis status for the study participants. Information on census-tract level and school-district level socio-demographic factors was obtained from the US Census and the National Historical Geographic Information System (NHGIS) databases. Aim 1 included pertussis incidence data for all ages reported between 1999 and 2017 and used Bayesian hierarchical disease mapping models and the tau statistic to characterize spatio-temporal dependence between pertussis cases. The Kulldorff spatial scan statistic was used to examine location of pertussis clusters and their overlap with clusters of non-medical vaccine exemptions. For aim 2, we used pertussis incidence data for all ages reported between 2010 to 2017 and estimated annual school-district level vaccination coverage as proportion of 19-35 month old children who received ≥ 4 DTaP doses using immunization data from the WA-IIS. Association between pertussis epidemics and vaccination coverage and other socio-demographic factors was estimated using epidemic-endemic models and the ecological vaccine model. For aims 3, 4, and 5, we used the linked dataset with individual level vaccination and pertussis status for all children born

or living in King County between 2008 and 2017. Log binomial models were used to estimate the association between DTaP vaccination schedules and age-specific pertussis incidence. Cox proportional hazards models were used to estimate vaccine direct effects among children older than 3 months and population-level effects among children older than 7 months. Schoenfeld residuals obtained from fitting Cox proportional hazards models were used to estimate waning of vaccine effectiveness among 5-9 year old children.

Results: There was no overall increase in pertussis incidence between 1999 and 2017, but we found spatial dependence between pertussis cases at very small spatial scales. Pertussis clusters overlapped with clusters of vaccine refusal suggesting an association between the two. We estimated the vaccine effectiveness of DTaP vaccine to be 83% (95% credible intervals: 63%, 95%) using the ecological vaccine model but found no correlation between the effective reproduction number of pertussis and area-level vaccine coverage. The association between area-level under-vaccination and pertussis epidemics was statistically significant as estimated using the epidemic-endemic models (adjusted Relative Risk, aRR: 2.76; 95% confidence interval: 1.44, 16.6), suggesting areas with low vaccination coverage had higher risk of experiencing pertussis outbreaks. We found significant association between under-vaccination and age-specific pertussis risk, but a short delay of a few weeks in receiving DTaP doses did not significantly alter pertussis risk. Using the Cox proportional hazards models and DTaP vaccine series as a time-dependent exposure, direct vaccine effects were estimated to be 72% (95% CI: 65%, 77%) comparing vaccinated time at risk to under-vaccinated time at risk for the entire cohort. The estimated indirect protection for the 3-dose primary series was 45% (95% CI: 1%, 70%), total protection was 94% (95% CI: 91%, 96%), and overall protection was 42.2% (95% CI: 19%, 60%). We found no evidence of waning of vaccine effectiveness after 5 doses of DTaP among 5-9 year old children. Vaccine effectiveness remained high at 83% (95% CI: 39%, 95%) four years after vaccination with 5th DTaP dose.

Conclusion: Our findings show that although pertussis transmission is ongoing in King County, there is no clear evidence for resurgence between 1999 and 2017 as seen in the rest of the country. We found that the current schedule for the 5-dose childhood DTaP vaccine series effectively reduces pertussis risk and adding or delaying booster doses may not be required. We estimated direct vaccine effectiveness (VE) of the acellular pertussis vaccine using different models in this dissertation and found that the estimates were high and consistent across analyses. However, direct VE estimates from the statistical models used in this dissertation do not provide information about mechanism of vaccine failure (i.e. leaky vs. all-or-none). We found significant vaccine indirect effects for the acellular vaccine suggesting that vaccination with DTaP may contribute to herd effects, although we cannot deduce from these data if the indirect effects are due to decrease in susceptibility to infection or decrease in infectiousness after exposure among vaccinated individuals. We found no evidence of rapid waning of vaccine effects among children who were fully vaccinated with 5 doses of DTaP, suggesting that immediate waning of vaccine effects is likely not the mechanism of vaccine failure for the DTaP vaccine. Understanding the nature of the acellular pertussis vaccine failure in shaping the epidemiology of pertussis is challenging and deserves continued research.

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I want to thank my advisor Prof. Betz Halloran for her unwavering patience, mentorship, and encouragement throughout my doctoral training. Betz fully supported my research endeavors, helped me collaborate with different public health agencies, and pulled out all the stops to get me the data I needed to see this project through. Every time I hit a roadblock, she had my back, and for that I'm incredibly grateful. Betz's grant R37-AI032042 with the National Institutes of Allergy and Infectious Diseases at the National Institute of Health fully funded this research.

I also want to thank my stellar committee members Prof. Pejman Rohani, Prof. Jon Wakefield, and Prof. Amanda Phipps, who were so generous with their time and provided invaluable technical guidance that helped me through some challenging areas of my research. From Pej, I learned how to always keep my eye on the scientific question at hand which improved my writing process. From Jon, I learned to always check the distribution of my data and know the assumptions of my models. I also learned from them to take time out to follow soccer and cricket, which is as important as anything pertussis-related they taught me! I also want to thank Prof. Steven Goodreau for serving as the GSR on my committee.

From the departments of Epidemiology and Biostatistics at the University of Washington, I would like to thank Dr. Ali Rowhani-Rahbar, Dr. Steve Hawes, Dr. Paul Drain, Dr. Marco Carone, Dr. Steve Mooney, and of course, Amanda again, for their mentorship and instruction during my time as a student at the department. I would also like to thank Julie Nevins, John Paulson, and, Kevin Schuda from the epidemiology department and Rebecca Allen and Stephanie Shadbolt at Fred Hutch for their administrative support over the last five years.

I also want to thank my former advisor from Johns Hopkins, Prof. Derek Cummings.

Derek introduced me to quantitative epidemiology and was insistent that I pursue a PhD, which turned out pretty great for me.

I want to thank my fellow classmates and friends in the department of Epidemiology- Lola Arakaki and Rachel Kubiak for your constant support and friendship, Natasha Wenzel for keeping me fed during my Oral exam and for all the hand-made birthday gifts and cakes, Tigran Avoundjian, Julianne Meisner, Sarah Romano, Maayan Simckes, and Kennedy Muni, for all the stimulating discussions we have had over the years about life and epi and spatial statistics. I'm proud to be a part of a cohort of such brilliant and passionate epidemiologists.

A huge shout out to all my friends from over the years and across the world. Thank you Tanwee, Vajini, and Gayatri, for the most wonderful two years at Hopkins, for being the most crackin' roommates, and for always rooting for my success. Vajini, if I did indeed inspire you switch to epidemiology for your doctoral studies, that will be my greatest contribution to the field of epidemiology! Thank you Anandi, Shamik, and Sindhu for being my forever cheerleaders. You guys can make me laugh, no matter what. Thank you Shashi, for feeding me so much Indian food during my time in New York. I am lucky to have such brilliant and motivated friends who I can look up to.

I want to thank my parents, Anuradha and Suhas Rane, for their support and sacrifice. My education was always the top priority for them and they encouraged me to follow my dreams. I simply could not have achieved any of this without them. I wish my dad were here to watch me defend my thesis (he would definitely have asked some tough pertussis questions!), but I take solace in thinking he would have been proud of me.

Finally, I want to thank my partner, Sri, who for the past 10 years has been my number one fan. Because why else would one put up with a mostly grumpy grad student? If I have made it this far, it is largely thanks to his support, encouragement, and terrible jokes. Thank you for believing in me, for making me a kinder, more patient person, and most importantly, for showing me how to run simulations on a cluster.

DEDICATION

I dedicate this thesis to the loving memory of my father Suhas Rane, and to my mother, Anuradha Rane.

Chapter 1

Background

Pertussis, also known as whooping cough" is a highly contagious respiratory disease caused primarily by the bacterium *Bordetella pertussis*. In its early stages, it is characterized by flu-like symptoms such as a mild cough, cold, runny nose, and a low grade fever (catarrhal stage), followed by paroxyms of rapid coughs ending in a "whoop" sound, vomitting, and exhaustion (paroxysmal stage). The coughing can last up to 6 weeks. This is followed by a convalescence stage of gradual recovery [1]. Pertussis can be particularly severe in infants and small children and is still one of the leading causes of infant mortality, with upward of 100,000 deaths annually around the world [2]. Complications in infants include pneumonia, apnea, convulsions, and encephalopathy. Pertussis disease is only found in humans [3] and spreads from person to person via coughing or sneezing or sharing breathing spaces for long periods.

1.1 Pertussis through the ages

In the pre-vaccine era, pertussis was very common and most people got infected during their lifetime [4]. Widespread use of the whole-cell pertussis vaccine in combination with the Diphtheria and Tetanus vaccines (DTwP) in 1948 resulted in a large decline in annual pertussis incidence and mortality in the US, from 157 cases per 100,000 population in the pre-vaccine era to fewer than 1 case per 100,000 in the early 1970s [4]. Due to variable effectiveness of the DTwP vaccine and rising public concerns about its safety and reactogenecity, many countries switched to the acellular pertussis vaccine (DTaP) in the 1990s [5]. The acellular pertussis vaccine contains only purified pertussis antigens [4]. During the 1980s and 1990s, even as pertussis incidence declined globally with

increasing vaccination coverage with both DTwP and DTaP vaccines, incidence started increasing in some countries such as the US, UK, and Australia, that had consistently high vaccine coverage [6] [7]. Today, pertussis is still an endemic disease in all countries, with periodic outbreaks every 2-5 years [4]. In 2014, there were an estimated 24.1 million cases (95% confidence interval (CI): 7-40 million) and 161,000 deaths (95% CI: 38,000, 671,000) in children younger than 5 years, largely in low income countries [2]. Thus, despite achieving high vaccination coverage and an overall decrease in pertussis incidence globally, we are far from eradicating it with current vaccination strategies [8].

1.2 Resurgence of pertussis

Following an increase in pertussis in some countries that had switched to DTaP vaccines, the World Health Organization (WHO) reviewed pertussis trends from 19 high and middle income countries with different pertussis vaccines and vaccination schedules [6]. They found no evidence of global resurgence. The increasing number of pertussis cases in some countries was attributed to naturally-occuring disease cycles and increase in disease awareness and more sensitive detection methods. However, they did find evidence of true resurgence in Australia, Chile, Portugal, UK, and USA. All of these countries used the DTaP vaccine for routine immunization, but differed in their vaccine schedules and reporting fidelity, making direct comparisons difficult. Interestingly, nationwide resurgence in the US started in the 1970s, before it switched from DTwP to DTaP vaccine [7]. In addition to increase in mean incidence, there has been a shift in average age of incidence, with adults and adolescents forming a large proportion of reported cases [9].

1.3 Drivers of resurgence

Because pertussis causes severe disease and mortality in infants, it is critical to understand the drivers of resurgence of pertussis in the US and other countries [8]. In recent times, research surrounding pertussis resurgence has increased and several hypotheses have been proposed to explain it. However, there is little consensus on what might be causing resurgence in specific countries. Some of the main candidate explanations are: increased disease awareness and introduction of more sensitive molecular detection techniques (e.g. enzyme-linked immunoassays (ELISA) and PCR based tests); shift from DTwP to DTaP vaccine and differences in primary vaccine schedules; waning

vaccine-induced immunity; evolution of *B. pertussis* to escape vaccine-induced immunity; inability of DTaP and DTwP vaccines to completely block disease transmission and contribute to vaccine herd effects; asymptomatic transmission from infectious individuals; and circulation of congeners, such as *B. holmesii* and *B. parapertussis*, that cause pertussis-like disease [10] [11]. In all likelihood, the resurgence is a result of a combination of these mechanisms [12]. In addition to estimating vaccine effectiveness (or vaccine failure), elucidating the mechanism of vaccine failure might provide a deeper understanding of pertussis epidemiology [10]. There are three possible mechanisms of vaccine failure: primary vaccine failure (or all-or-none), leakiness, and waning of vaccine-induced immunity over time [13]. Primary vaccine failure occurs when the vaccine does not "take" for some proportion of individuals that receive it. In this case, the vaccine fully protects a proportion of population while does not offer any protection to the remaining individuals in whom it did not take. When a vaccine is "leaky", it reduces the disease rate in all vaccinated individuals by a constant proportion [13]. Vaccine failure due to waning occurs when the vaccine-induced protection decreases over time. In this dissertation, we used various epidemiological and statistical models to shed some light on what the pertussis incidence data in King County tells us about the roll of vaccination coverage and DTaP vaccine failure on pertussis resurgence and persistence.

1.3.1 Impact of local heterogeneity in vaccination coverage on pertussis trends

Vaccination coverage estimates reported at the national or state level could hide local heterogeneities. Coverage estimates at granular spatial scales, such as census tracts, school districts, and city blocks are rarely reported. Vaccination coverage in some areas of King County may be lower than the herd immunity threshold value for pertussis due to vaccine refusal or lower access to healthcare facilities or other health inequities [14]. Local circulation of pertussis in King County and occasional epidemics might occur in pockets of low vaccination coverage. Hence, we assessed the association between area-level vaccination coverage and sociodemographic factors and pertussis epidemics in school districts in King County [Chapter 3].

1.3.2 Vaccination timing and schedules

In the US, childhood DTaP series is recommended at months 2, 4, and 6, the second-year booster between ages 15-18 months and the school-age booster between ages 4-6 years. The number and schedule of DTaP primary as well as booster doses varies between countries and can influence disease burden in younger ages. We assessed whether the US DTaP vaccine schedule is effective in reducing age-specific pertussis risk among children under 10 years of age in King County [Chapter 4].

1.3.3 Population-level effects of the acellular pertussis vaccine

Animal studies have demonstrated that acellular pertussis vaccines may not be effective in reducing transmissibility [15]. One study found evidence for asymptomatic transmission using wavelet analysis and genomic data [16]. On the other hand, epidemiologic data from Sweden where resuming vaccination with the DTaP vaccine after a gap of 17 years resulted in decrease in incidence in the unvaccinated populations provides evidence for vaccine herd effects [17]. Directly estimating indirect effects using individual level vaccination and disease status would provide evidence for the transmission reduction potential of the acellular pertussis vaccine. To our knowledge, only one randomized controlled trial so far has directly measured DTaP vaccine indirect effects. The study compared the DTaP vaccine to DT vaccine and found that vaccination with DTaP reduces transmission to close contacts [18]. Another prospective cohort study in Senegal showed pertussis vaccination is effective in reducing transmission from vaccinated breakthrough cases [19]. To examine whether acellular pertussis vaccines contribute to herd effects, we estimated the indirect, total, and overall vaccine effects for the 3-dose DTaP primary series using routinely collected immunization data for children living in King County from the Washington State Immunization Information System [Chapter 5].

1.3.4 Waning effects of the acellular pertussis vaccine

The waning effects of the acellular pertussis vaccine have to be studied indirectly by estimating temporal trends in infection rates among vaccinated individuals or by fitting dynamic transmission models to population-level pertussis time-series data because there does not exist a correlation of protection for the pertussis vaccine [20]. Evidence that DTaP vaccine effectiveness wanes over time is strong, but the estimates of duration of immunity vary widely [21] [22] [23] [24] [25]. Rate of waning of vaccine-derived immunity

can have an impact on pertussis transmission and age distribution, and have important policy implications, such as determining vaccine schedules or increasing the number of boosters. It is important to resolve and explain the discrepancies in the various estimates of waning effects obtained using different methods and data. In this dissertation, we estimated waning of immunity after 5 doses of DTaP vaccine in children 5 through 9 years of age using Schoenfeld residuals, a statistical method that has been used to estimate waning of vaccine effects for cholera vaccines [26] [Chapter 6].

These analyses were conducted in King County, WA, which is the most populous county in Washington, which experienced a large pertussis outbreak in 2012, and has high reported vaccination coverage. Because we focussed on one county, we were able to get individual-level data for all reported cases and vaccination status for all children residing in King County, forming a rich population-based cohort spanning ten years. With these data, we were able to address important questions about effects of vaccination on pertussis epidemiology using individual level data.

Chapter 2

Epidemiology of pertussis and spatio-temporal trends within King County, WA

Abstract

Background: Pertussis epidemiology has been extensively studied in the last decade. However, the spatial dynamics of pertussis are poorly understood, especially the scale of transmission and clustering.

Methods: We geocoded home locations of 3,895 pertussis cases diagnosed in King County, Washington and reported to Public Health Seattle and King County between 1999 and 2017. We used Bayesian models with spatial and temporal random effects to describe space-time dynamics of pertussis over a 19-year period. We used the tau clustering statistic to measure spatial dependence between individual cases. We calculated correlation in pertussis epidemic curves between school districts to see if clustering of cases led to synchronized outbreaks between districts. Finally, we used the Kulldorff spatial scan statistic to identify locations of pertussis and non-medical vaccine exemption clusters within King County and measured the geographic overlap between them.

Results: We found census-tract level spatial dependence in pertussis incidence in outbreak years but no clear temporal trends in annual incidence. In the period 2009-2012, which included the largest pertussis outbreak, pertussis cases were 1.22 times (95% Confidence Interval (CI): 1.16, 1.27) more likely to occur within 0.01 km of each other in

the same month than expected given the underlying spatial and temporal distribution of cases. Significant clustering was observed up to 0.1 km in all time periods. Pertussis epidemics in neighboring school districts were synchronized and correlation decreased with distance. There were 6 statistically significant census tract clusters for reported pertussis cases between 1999 and 2017 and 7 significant clusters for vaccine exemption, and the geographical overlap between the two was statistically significant (OR: 2.07; 95% CI 1.29, 3.35).

Conclusion: There was no evidence for increase in pertussis incidence within King County, WA since 1999. Pertussis cases cluster over very small spatial scales suggesting immediate neighbor contacts are important contributors to transmission. Census-tract level pertussis clustering might indicate clustering of vaccine refusal.

2.1 Introduction

Pertussis, or whooping cough, is a serious, highly contagious respiratory infection primarily caused by the bacterium *Bordetella pertussis* [27]. Following large-scale implementation of a pediatric vaccination program with whole-cell pertussis vaccines in the 1950s, pertussis activity decreased dramatically in the US [7], UK [28], Australia [29] , and Canada [30]. Some countries have experienced a resurgence in pertussis incidence in the last few decades, especially among adolescents, despite high vaccination coverage [11] [31]. In 2012, there was an unexpectedly large epidemic in the US, resulting in 48,277 reported cases and 20 deaths. The state of Washington alone recorded around 5,000 cases, the second highest number of cases in the country [32]. As pertussis causes severe disease among infants too young to be vaccinated [33], it is crucial to understand the complexities of pertussis epidemiology.

Although pertussis has been studied widely, the spatial scale of transmission and spatial clustering of pertussis have not been well-characterized. Spatio-temporal modeling of infectious diseases has become increasingly possible with availability of surveillance data with precise locations of cases [34]. Global and local clustering in pertussis incidence has been studied before with the aim to explore relationships between vaccine refusal and pertussis incidence and to identify areas for intervention [35] [36] [37] [38] [39]. However, these studies measured spatial dependence using aggregated data at the state, county, or community level. Important insights about spatial scale of transmission can be lost

due to aggregation. Understanding the extent of spatial clustering in cases and the spatial scale of transmission can help elucidate mechanisms of pertussis spread and inform control and surveillance activities [40].

The goal of this analysis is to describe the spatio-temporal epidemiology of pertussis at a fine spatial scale within King County, Washington. We used detailed geo-referenced surveillance data collected over a 19-year period for analyzing spatio-temporal trends. To our knowledge, ours is the first study to have explored high resolution spatial dependence in pertussis cases using point pattern data. Using Bayesian hierarchical spatio-temporal models [41], we first mapped the spatio-temporal variation in pertussis risk at census-tract level over the entire 19 year time period. We then characterized short-term spatial dependence between individual pertussis cases using the tau statistic measure for global clustering [40]. We quantified correlation between pertussis time-series in King County school districts to check if short-term spatial dependence between individual cases led to synchrony in pertussis epidemics between neighboring areas [42]. Finally, we used the Kulldorff scan statistic to find the locations of space-time clusters within King County and measured their overlap with vaccine exemption clusters [43].

2.2 Methods

2.2.1 Study Area

King County is the largest and the most populous county in Washington State, with an area of 2,307 mi^2 (5,975 km^2) and population of 2.1 million as of 2017 [44]. The Puget Sound surrounds the west and the Cascade Range is located in the eastern part of the county. The largest city is Seattle with a population of 725,000 (2017) [44]. About two-thirds of King County's population lives in Seattle and surrounding metropolitan areas. With the exception of Vashon Island, western King County is largely urban and suburban.

King County has 18 school districts, 82 zip codes, and 397 census tracts. Analysis in this paper is done at the school district and census tract level, with zip code information used for geocoding cases when information on census tracts was missing. There is considerable variability in the geographic area and population size of the census tracts and schools districts. The 397 census tracts in King County range from 0.35 km^2 to

2,181 km^2 in area and the 18 school districts range from 51 km^2 to 1,817 km^2 in area. Most census tracts (n=317) have area <10 km^2 and only 18 tracts are larger than 50 km^2. The median population of census tracts in 2010 was 4,777 (range: 1,287-11,041). Census tracts vary in area depending on population density, so census tracts in Seattle and surrounding metropolitan areas are small but densely populated while those in the Cascades are geographically large but sparsely populated. The median population of school districts is 91,356 (range: 627 – 609,471). Seattle school district is the most populated, while Enumclaw, Skykomish, and Snoqualmie school districts, which are located near the Cascades, are the least populated. All urban school districts are, on average, within 35 kms (21 miles) of Seattle. Rural school districts are on average, 55 kms (34 miles), from urban school districts.

2.2.2 Pertussis Cases

In King County, 4,173 pertussis cases were reported between 1st January, 1999 and 31st December 2017, of which 3,904 were residents of King County. In this period, two large outbreaks occurred: one in 1999 with 456 cases and another in 2012 with 895 cases. Pertussis is a nationally notifiable disease and cases are reported to Public Health Seattle and King County (PHSKC) by primary care providers and diagnostic laboratories within 24 hours of detection. The clinical case definition of pertussis is a cough illness lasting $\geq$ 2 weeks with at least one of the following: paroxysms of coughing or inspiratory "whoop" or post-tussive vomiting, or apnea (with or without cyanosis) for infants up to 1 year of age. Cases are classified as suspected, probable, or confirmed based on Centers for Disease Control (CDC) case definitions of pertussis (Supplementary Material 2.9.1) [45]. Pertussis cases in this study include only ones that were symptomatic, sought medical help, and were ultimately reported to the PHSKC surveillance system. One challenging aspect of using surveillance data to study pertussis is the high level of under-reporting, especially among adolescents and adults [8].

PHSKC additionally collects information on patient demographics, home address, dates of disease diagnosis and onset, site of exposure, clinical symptoms, laboratory test results, vaccination status, number of Diptheria-Tetanus-acellular Pertussis (DTaP) doses, date of last DTaP dose, hospitalization status, antibiotics administered, vaccine exemption status, and reasons for exemption. Washington State Department of Health (WA-DOH) estimated that in 2017, 74% of children aged 19-35 months were fully vaccinated with

$\geq$ 4 DTaP vaccine doses in King County [46].

2.2.3 Geocoding of pertussis cases

We used ArcGIS 10.1 for geocoding case locations [47]. Using TIGER/Line® King County census tract shapefiles with address range features as reference data, we geocoded street addresses (and zip codes when street addresses were unavailable) of pertussis cases [48]. Address ranges are a unique collection of addresses that fall along a road or path. They provide a way of locating homes and businesses based on their street addresses when no other location information is available. Each potential match between a case's residence address and a street address from the reference file was scored between 0 and 100 by ArcGIS. A match was scored low when addresses were misspelled, incorrect, or missing some elements. According to the software, a match score between 85 and 99 can be considered a good match. Based on this, a threshold score of 85 was set to classify a match as a "true" match. Addresses from partial matches were manually edited to find true matches. The geocoding process is described in Supplementary Material (2.9.2). After geocoding addresses, ArcGIS created a new point shapefile that included census tracts of residence appended to geocoded home addresses. We spatially overlaid the point shapefile on the King County school district shapefile to get the school district of residence for cases. The maps presented in this paper do not identify home locations of cases.

Census tract- and school-district level population estimates for years 2000 and 2010 were obtained from the National Historical Geographical Information System [49]. Shapefiles for King County census tracts for years 2000 and 2010, and King County school districts for 2010 were obtained from the US Census [50]. Information on school-level vaccine exemptions is made publicly available by Washington Department of Health [51].

2.3 Descriptive Analysis

2.3.1 Pertussis case characteristics by vaccination status

The DTaP vaccine is administered as a primary series at months 2, 4, and 6 with a second-year booster at age 15-18 months and a preschool booster at age 4-6 years. An adolescent booster is recommended at age 11-12 years. Cases over 6 months of age were

determined to be up-to-date (UTD) for the pertussis vaccine series if they received age-appropriate number of DTaP doses at the time of diagnosis. Assuming that 3 doses of DTaP are required for adequate protection against pertussis, infants under 6 months of age were considered not up-to-date (not UTD) even if they recieved up to two doses of the vaccine [52] (details in Supplementary Materials 2.9.3). We summarized demographic and clinical characteristics of cases by vaccination UTD status and age group.

We measured vaccine efficacy of DTaP vaccine series against pertussis severity in this population using methods from Halloran, et al. (2010). [13]. Severe cases were defined as having at least one of these conditions: hospitalization, seizure, encephalitis, pneumonia, and/or death. Vaccine efficacy for disease progression (VE_p) measures the effect of vaccination on diseases progression, here disease severity, conditional on having disease. It is a measure of the decreased severity of breakthrough disease in individuals who are UTD for the vaccine compared with disease in unvaccinated or undervaccinated individuals. VE_p was estimated by subtracting the relative risk of severe disease in UTD children, compared with unvaccinated or undervaccinated children, from 1; VE_p = 1-[(number of severe UTD cases/number of all UTD cases) / (number of severe non-UTD cases/number of all non-UTD cases)] [53]. We fit a generalised linear model with Poisson family and a log link to estimate the association between disease severity and being UTD for pertussis vaccine. The model was adjusted for age.

2.3.2 Pertussis temporal trends

We described pertussis trends in different age groups using annual age-specific incidence and age distribution of pertussis cases in 5-year intervals between 1999 and 2017.

2.3.3 Mapping the spatio-temporal trends in pertussis

Census-tract level pertussis incidence was summarized as age-standardized incidence ratios (SIRs) [54]. Pertussis SIR for each area i and time t was estimated as $\frac{Y_{it}}{E_i}$ where Y_{it} is the count of pertussis cases in census tract i at time t and $E_i = \sum_j p_j N_{ij}$ is the expected number of pertussis cases in census tract i with age-group specific reference rates p_j for j age-groups (Supplementary Materials 2.9.4). N_{ij} is the number of individuals in area i and age-group j. Census tract boundaries in King County changed between 2000 and 2010 (378 census tracts were redrawn into 397 census tracts). For ease of analysis, we

decided to use 2010 census tract boundaries in our study and used census tract population estimates from the 2010 US census. We assumed that the population per census tract does not change in this period, however this is a strong assumption. Census-tract level raw pertussis SIRs for each year from 1999 to 2017 (Supplementary Materials 2.9.4 Figs. 1(a) and 1(b)) show the variability in expected number of pertussis cases in each census tract summed over the entire time period. The number of expected cases are low in census tracts in the eastern part of the county that had large areas but were sparsely populated, resulting in spuriously high incidence ratios.

To identify spatial and temporal trends in pertussis disease within King County, we used flexible Bayesian spatio-temporal disease mapping models with structured and unstructured spatial and temporal random effects as well as a space-time interaction term to examine spatial and temporal dependence at the census tract level over time [41] [55] [56]. Assume that the number of pertussis cases in each census tract (i=1,...,397) and year (t=1,...,19) follows a Poisson distribution with relative risk λ_{it}.

$$Y_{it}|\lambda_{it} \sim Poisson(E_i\lambda_{it})$$

where Y_{it} is the observed number of cases in census tract i and time t, E_i is the expected number of cases in census tract i which is constant over time, and λ_{it} is the relative risk of pertussis in census tract i compared to the reference risk in each stratum and can be interpreted as the SIR_{it}.

The Bayesian spatio-temporal model with interaction is,

$$log(\lambda_{it}) = \alpha + \theta_i + \nu_i + \gamma_t + \phi_t + \delta_{it}$$

where α is the intercept interpreted as the overall county-wide pertussis log risk in a given census tract and time period with all random effects set to zero, θ_i is the spatially structured random effect, ν_i is the spatially unstructured random effect, γ_t is the temporally structured random effect, ϕ_t is the temporally unstructured random effect, and δ_{it} is the spatio-temporal interaction. The unstructured temporal term is modeled as $\phi_t \sim_{iid} N(0, \sigma^2_\phi)$, $t = 1, ..., T$ where $\sim$ iid is short for " independently and identically distributed". γ_t which is the smooth temporal term is given a second-order random walk prior, written as RW2. The unstructured spatial term ν_i is assigned a $N(0, \sigma^2_\nu)$ prior, and the structured spatial term θ_i is assigned an intrinsic conditional autoregressive (ICAR) prior [55].

As described in [41], interaction between space and time can be modeled in four ways: interaction between unstructured spatial and unstructured temporal effects (Type I), structured temporal and unstructured spatial effects (Type II), unstructured temporal and structured spatial effects (Type III), and structured spatial and temporal effects (Type IV). The details of the process of model selection and associated results are in Supplementary Materials 2.9.4. We found little evidence of a temporal structure but some evidence of spatial dependence in the maps (Supplementary Material 2.9.4 Fig. 2). This suggests that the Type III interaction term should be considered for this model. This was further examined by fitting all four types of interaction models and comparing the sum of the log conditional predictive ordinate (CPO) values. Models with large CPO values are preferred. The log CPO values for models with the four types of interaction terms were -21400.36, -15658.2, -10316.41, -41557.01, which suggests that Type III interaction is the most appropriate for this analysis. The models are implemented in INLA [41].

2.3.4 Census tract level spatial dependence in pertussis incidence using Moran's I

Moran's I statistic was used to test for census tract level spatial autocorrelation in pertussis incidence [57]. Moran's I values lie between -1 and 1 and are close to zero if there is no spatial autocorrelation. A positive value suggests that areas with high or low incidence tend to cluster together. A negative value implies that areas with high or low incidence are more spatially dispersed than expected under complete spatial randomness. Because pertussis is a rare disease, Poisson models were fit to estimate census-tract level pertussis incidence separately for each year with adjustment for latitude and longitude of census tract centroids to remove large scale spatial trends. Spatial autocorrelation over the entire study area was estimated using Moran's I applied to the residuals of the Poisson models.

2.3.5 Short-term spatial dependence between pertussis cases using the τ statistic

Spatial dependence between individual cases occurring within a month of each other was estimated using the tau (τ) statistic [40] [58]. This statistic estimates the probability of two cases occurring within distances *d1* and *d2* and within 30 days of each other relative

to the independent probabilities of observing two cases within distances *d1* and *d2* over the entire time period and of observing two cases within 30 days of each other over the whole study area. We assumed that cases occurring within 30 days of each other are likely to be related because the generation time of pertussis is 21-30 days. In this method, underlying spatial differences between census tracts such as population density impact both the numerator and denominator in the same way, so they do not bias the measure of spatial dependency [40]. The τ statistic is calculated as:

$$\tau(d_1, d_2) = \frac{Pr(b \in \Omega_a(d_1, d_2, t_1=0, t_2=30))}{Pr(b \in \Omega_a(d_1, d_2, .))Pr(b \in \Omega_a(., t_1=0, t_2=30))}$$

where $\Omega_a(d_1, d_2, t_1 = 0, t_2 = 30)$ is the set of cases that occur within 30 days between distance d_1 and d_2 of a case a, $\Omega_a(d_1, d_2, .)$ is the set of cases within distance d_1 and d_2 of case a over the entire time series, and $\Omega_a(., t_1 = 0, t_2 = 30)$ is the set of cases that occur within 30 days around case a over the entire area. The value of d_2 in our model ranges from 0.02 km to 1 km, and value of d_1 is kept constant at 0.01 km. The 95% confidence intervals of the τ statistic are based on 500 bootstrapped values. We estimated $\tau(d_1, d_2)$ using cases in time periods 1999-2003, 2004-2008, 2009-2012, 2013-2017 to assess whether spatial dependence changed over time. This analysis is implemented within R package `IDSpatialStats` [59].

2.3.6 Correlation in epidemics between school districts

To assess spatial synchrony and time-lagged cross-correlation in pertussis outbreaks, we computed pairwise spatial correlation and time-lagged cross-correlation between pertussis time series in King County school districts. For this analysis, we chose a spatial scale of school districts instead of census tracts to because weekly pertussis case counts at the census tract level were sparse. Distance between pairs of school districts was calculated as the Euclidean distance between their centroids. Pearson's correlation [60] in pertussis time series between two school districts is given as,

$$\rho_{i,j} = \frac{Z_i - \mu_i}{\sigma_i} \frac{Z_j - \mu_j}{\sigma_j}$$

where $\rho_{i,j}$ is the correlation coefficient between Z_i and Z_j, which are weekly pertussis time series in school-districts i and j, μ_i is the marginal mean of weekly pertussis cases

in school district i and σ_i is the marginal standard deviation.

Spatial dependence as a function of continuous distance was estimated non-parametrically using a spline correlogram. The `ncf` package in R [61] was used to obtain the spline correlogram and bootstrapped 95% CI [62]

2.3.7 Identification of space-time clusters of pertussis and their overlap with spatial clusters of vaccine exemption in King County

The tests of spatial dependency performed above examine evidence for clustering but tell us nothing about the location of clusters within King County. The Kulldorff scan statistic can be used to identify clusters of pertussis cases in a space-time setting [43] [63]. The scan statistic was performed at the census tract-level. We repeated this analysis to find census-tract level clusters of non-medical vaccine exemptions. Subsequently, we investigated whether pertussis clusters overlapped with clusters of non-medical exemptions (NMEs) in King County. We now provide further details.

Space-time clusters of pertussis in King County

A discrete space-time Poisson model was used to identify the locations and timings of high-risk pertussis clusters occurring between 1999 and 2017. In this model, the number of cases in each census tract is assumed to be Poisson distributed, and the model was adjusted for age (grouped as <1 year of age, 1-4 years, 5-9 years, 10-14 years, 15-18 years, 19-64 years, and ≥ 64 years) [64]. The age adjusted expected number of cases in a given census tract was calculated as:

$$E_i = \sum_j p_j N_{ij}$$

where p_j is the age-group specific pertussis reference rates for j age-groups calculated as the (number of pertussis cases in an age group)/(total individuals in that age group at risk) in King County and N_{ij} is the number of individuals in area i and age-group j.

Spatial clusters of non-medical vaccine exemptions in King County

We used data on number of school children with an NME from immunization requirement and number of children attending grades K-12 in King County schools in 2015-2016 school year to estimate vaccine exemption rates. These data are reported annually by King County schools and are available freely [51]. As vaccine exemption rates did not change appreciably between school years 2011-2012 (earliest data available) and 2016-2017, we selected the 2015-16 shool year in this analysis. In 2015-16, there were 669 schools in King County with enrollment ≥ 10. These schools were geocoded using their street addresses to obtain the census tracts in which they are located. Students enrolled and students with NMEs were aggregated to the census tract. We used a spatial Poisson model for identifying NME clusters because overall NME rates in King County did not change much between 2012 and 2017 and we assume census tract level NME rates remained constant over time [51].

The SatScan space-time algorithm moves a cylindrical window with a circular base over the entire study area evaluating thousands of overlapping scanning windows, where the height of each cylinder represents a time interval. A purely spatial scan imposes just a circular window on the map [64]. The size of the circular base in both algorithms and the cylindrical height in the space-time algorithm can vary in size, producing potential clusters that vary in area and time spans. We set the upper limit of percentage of population at risk at 20%. We set the algorithm to detect clustering of cases that occur within a minimum temporal window of 1 month and a maximum of 50% of the study time period (default setting). No geographic overlap was allowed between clusters. The number of observed cases, the number of expected cases, and the Poisson generalized likelihood ratio (GLR) are estimated for each cylinder (space time pertussis clusters) and each window (spatial NME clusters) [64]. The maximum GLR from the observed data is compared to the maximum GLRs from 999 random Monte Carlo simulations under the null hypothesis of no clustering. A p-value indicates statistical significance of each cluster and significance is assessed at the 0.05 level [64]. This method identifies the most likely cluster and secondary clusters. According to the SatScan software, p-values of secondary clusters should be interpreted as the ability of the secondary cluster to reject the null hypothesis on its own strength, whether or not the more likely clusters are true clusters or not [64]. P-values of secondary clusters can be problematic to interpret because the distribution under the null of GLR for secondary clusters is not known. So secondary clusters should be viewed more cautiously because the null distribution

is not correct. Analysis was performed using SatScan version 9.6.1 software [65] and cluster maps were exported and displayed using RStudio [66]. All spatial and statistical analyses were performed in R statistical software and RStudio [66] [67]

2.4 Ethics Statement

This study was reviewed and approved by the Washington State Institutional Review Board, the PHSKC Research Administrative Review Committee, and the University of Washington Institutional Review Board.

2.5 Results

Of the 3,904 resident cases diagnosed in King County between 1999 and 2017, 3,592 (92%) were confirmed, 115 (2.9%) were probable, and 197 (5.1%) were suspected cases based on CDC pertussis cases definitions [45]. Of the 3,904 cases, 3,332 (85%) were successfully geocoded to their street address with a score of 85 or higher. Five hundred and sixty-three (14.4%) cases with missing or incorrect addresses were geocoded to the centroid of their zip code of residence, with random jitter incorporated to avoid over-plotting in one region. There may be several census tracts within a zipcode, so there is some error associated with geocoding these 563 cases to their correct census tracts (Supplementary Material 2.9.2). In all, 3,895 pertussis cases were included in the analysis (8 cases that could not be geocoded and 1 case with missing age were removed).

Of the 3,895 cases included in the analysis, 2,357 (60.5%) were UTD for pertussis vaccine series at the time of diagnosis. Seven hundred and forty (19%) of the cases had 0 doses of DTaP vaccine and 2,843 (73%) of the total cases were <19 years of age. Four hundred and seven (10.9%) of cases reported philosophical or religious reasons for not vaccinating. Infants and children under 4 years had lower vaccine UTD rates compared to other age groups (Supplementary Material 2.9.7 Table 1). Among cases who were hospitalized, 86% were not UTD for the vaccine series and spent a median of 1 day longer being hospitalized. As both clinical symptoms and number of pertussis vaccine doses depend on age, we looked at differences in clinical symptoms by age groups. Severe symptoms were more common in infants compared to other age groups. Out of 437 infant cases, 322 (73.5%) had apnea, 166 (38%) had cyanosis, 163 (37%) had to be hospitalized, and 332

(76%) had post-tussive vomiting. Pneumonia was more frequent in both infants (4.8%) and adults older than 65 years (6.1%) compared to other age groups (Supplementary Material 2.9.7 Table 2).

Based on our definition of severe pertussis, 97 out of 3,458 (2.8%) cases aged 6 months and older had severe illness. There were 55 (2.3%) severe cases among 2,301 cases that were UTD and 41 (3.9%) severe cases among 1,053 cases that were not UTD for the vaccine series. Vaccine efficacy to reduce disease severity was estimated to be 40% (95% CI: 9.0%, 59%), adjusted for age.

Weekly time series of pertussis from 1999 to 2017 shows two large peaks in 1999 and 2012, with smaller peaks in 2004 and 2015 (Fig. 2.1(A)). In the period between 1999 and 2004, most pertussis cases occurred among infants and adolescents. This pattern changed between 2005 and 2011, where proportion of adolescent cases decreased, but proportion of infant cases remained high. In the most recent period between 2012 and 2017, the proportion of cases among school aged children and adolescents increased again (Fig. 2.1(C)). The introduction of the TDaP booster for adolescents in the US in 2005 could explain the decrease in incidence among adolescents between 2005 and 2011 (Fig. 2.1(B)). The increase in adolescent cases after 2012 might be because these birth cohorts were primed with the DTaP vaccine for their primary vaccine series, as opposed to the earlier birth cohorts that received the whole cell DTwP vaccine as primary series. Since 2013, ACIP has recommended a maternal TDaP dose for pregnant women in their third trimester [68] and uptake was high in Washington state (80% in 2016 and 83% in 2017) [69]. This might have resulted in reduced incidence among infants in recent years.

We also plotted the mean and standard deviation of pertussis incidence using a four-year rolling window to assess stationarity of time series [70] (Supplementary Material 2.9.5A). A four-year period was selected to reflect the inter-epidemic period of pertussis in post-vaccination era [71]. There is a decline in mean weekly cases over time up to 2012, and then a sharp increase caused by the large epidemic occurred in 2012, after which incidence declines again. Variability in proportion of cases reported during each month for each age group was plotted to explore age-specific seasonality of pertussis in King County [22] (Supplementary fig 2.9.5B). Proportion of cases in adolescents peaked in May and June. Infants and toddlers appear to have higher proportion of cases in the months July-September. However, the data were too sparse to draw any conclusions about transmission between age groups.

Fig. 2.2 displays the variation in spatial clustering of pertussis risk over time between 1999 and 2017. We plotted the smoothed posterior median relative risk estimates (the term $\alpha + \theta_i + \nu_i + \gamma_t + \phi_t + \delta_{it}$ from the Bayesian spatio-temporal model). This map is smoother compared to the map of raw SIRs plotted in Supplementary Material (2.9.4 Fig. 1(a)) with neighboring areas having similar SIRs, but the spatial distribution of risk can be observed. Years of moderate to large outbreaks (1999, 2003, 2005, 2012, 2015) are easily identified. Vashon Island, which lies to the west of the county, shows high relative risk for pertussis in most years. In 1999, high risk was concentrated in the north-west of the county. Incidence in the southern and south-eastern parts of the county started increasing thereafter, and in 2016 and 2017 incidence was concentrated in southern King County. There were differences in the pattern of spatial dependence in the outbreak and non-outbreak years. During outbreak years (1999, 2003, 2005, 2012, and 2015) pertussis activity was higher in north-western, south-western parts of King County, although in 2012, the entire county was affected. However, during periods of low incidence (2000, 2002, 2006, 2008, 2011, 2013-2014, 2016-2017) the south-eastern and eastern parts of the county were most affected. A map of standard deviations of the fitted posterior median relative risk estimates shows the uncertainty around model fits for each census tract over the years (Supplementary Material 2.9.6). There is high uncertainty around relative risk estimates for Vashon Island particularly during non-outbreak years. We did not find any temporal structure in annual pertussis dynamics likely because of the scale of aggregation. The generation time of pertussis is 3-4 weeks. Hence, aggregation of pertussis counts by month might have been more appropriate to study its space-time dynamics. However, given the already sparse case data and large number of census tracts, a model with monthly data would have been very complex (would need to estimate 397 x 138 random effects for interaction terms) [34].

To check if the spatial clustering visualized in the space-time maps is statistically significant, we estimated Moran's I statistic and associated p-values for each year (Table 2.1). We found statistically significant positive spatial autocorrelation at the census tract level in 1999 (p-value=0.05), 2012 (p-value=0.03), and 2015 (p-value=0.05), all of which were outbreak years. We did not find significant spatial autocorrelation in the rest of the years. This might explain the overall low contribution of the spatial ICAR random effect (20%) to total spatial variability in the Bayesian heirarchical model without interaction term.

While Moran's I statistic can estimate global clustering using data aggregated at the census tract level, spatial dependence in directly transmitted infections can occur at even smaller spatial scales. Using the τ statistic we measured spatial dependence between individual pertussis cases. We found that in the period 1999-2003, two pertussis cases were 4.03 times (95% CI: 3.86, 4.30) more likely to occur within a distance of 0.01 km (10 m) of each other in a 30 day window compared to the independent probabilities of occurring within 0.01 km of another case over a 5-year period and occurring within the same month across the entire county (Fig. 2.3). This relative risk fell to 1.70 (95% CI: 1.61, 1.74) for cases occuring within 0.1 km of each other and 1.28 (95% CI: 1.25, 1.30) for cases occuring within 0.2 km. In the period 2009-2012, pertussis cases were 1.22 times (95% CI: 1.16, 1.27) more likely to occur within 0.01 km in the same month and 1.04 times (95% CI: 1.03, 1.05) more likely to occur within 0.1 km distance of each other within the same month. Spatial dependence was observed over very short distances (<0.1 km) for all time periods.

Fig. 2.4 illustrates the results of spatial correlation in pertussis outbreaks between school districts in King County. Outbreaks were spatially synchronized up to at least 60 kms with regional average outbreak correlation of 0.15 (Fig 2.4a). Correlation decreased with distance, but overall correlation was low because there were fewer cases than locations for most weeks. Mean correlation was 0.29 (95% CI: 0.09, 0.32) for school districts <10 km apart, 0.16 (95% CI: 0.11, 0.26) for those between 10 and 20 km of each other, and 0.10 (95% CI: 0.01, 0.26) for school districts >50 km apart. One-week (Fig 2.4b), two-week (Fig 2.4c), three-week (Fig 2.4d), and four-week (Fig 2.4e) lagged spatial cross-correlation functions show that outbreaks persisted spatially for up to 4 weeks, however strength of correlation decreased with increasing time lags.

Using the discrete Poisson space-time model in the SatScan package, we identified six statistically significant clusters of census tracts (in green) (Table 2.2; Fig. 2.5). The clusters were located in Vashon Island, Seattle, Kent, Auburn, Northgate, and Lake city neighborhoods. One census-tract was a cluster by itself with 97 cases which occurred between January and April 1999. The largest cluster had 259 cases in it and occurred between March and December 2012. The time frame for clusters ranged from 2 months to 1 year.

We overlaid the space-time pertussis clusters on clusters of non-medical exemptions to visualize whether there is any overlap between the two (Fig. 2.5, Table 2.3). We hypoth-

esized that census tracts with higher exemptions would have higher risk of pertussis. NME clusters (in blue) were found in Vashon Island, Enumclaw (in south-east King County), north-western, and north-eastern regions. There is considerable overlap between pertussis and NME clusters. The odds of a census tract in a pertussis cluster also being in an NME cluster was 2.07 (95% CI 1.29, 3.35).

2.6 Discussion

This study described the spatial dependence in pertussis cases on small spatial scales and long-term spatio-temporal trends of pertussis within King County, WA. We found significant short-term spatial dependence in pertussis cases, suggesting pertussis transmission occurs at a very fine spatial scale. There was synchrony in epidemics in neighboring school districts suggesting short-term spatial dependence may have led to correlated outbreaks between school districts. Furthermore, we found risk of pertussis at the census tract level is associated with geographic hotspots of vaccine refusal.

We found pertussis vaccines to be protective against severe disease in pertussis cases in King County. This finding is consistent with other studies that showed that pertussis vaccines reduce disease severity [72] [73]. A follow-up study of 30,000 individuals in Senegal in 1993 found that pertussis vaccines were 48% effective (95% CI 39%– 55%) in reducing disease severity among pertussis cases [53]. Using their method, we estimated the effect of pertussis vaccine on reducing disease severity in this population to be 40% effective (95% CI: 9%, 59%). Our vaccine efficacy estimate is somewhat lower than the Senegal study possibly because they compared vaccinated children to unvaccinated children, while we compared individuals older than 6 months of age who were UTD to those who were not UTD for the pertussis vaccine series. Furthermore, they defined severe pertussis using a scale that scored various combinations of clinical signs and symptoms, while we used a different, simpler classification method.

Several states in the US and other countries such as UK and Australia experienced a large pertussis outbreak in 2012 and recorded thousands of cases [7] [28] [29] [74]. King County also experienced its largest epidemic since 1999 in 2012 with 895 cases. While we observed a substantial increase in pertussis incidence among adolescents, we did not find similar increases in incidence for adults >20 years of age as observed in Massachusetts [22] and UK and Wales [70]. A steady increase in pertussis incidence was

observed in Massachusetts between 1990 and 2005 and in UK between 2004 and 2012. In contrast, we observed a slight decrease in pertussis activity from 1999 to 2012, with a short upward trend brought on by the 2012 epidemic. This decrease could not be explained by changes in vaccine coverage which remained steady at around 75% during this period [46]. Furthermore, the structured temporal random effects from the Bayesian heirarchical model were also negligible, suggesting no temporal changes in annual pertussis incidence in King County between 1999 and 2017 (Supplementary 2.9.4 Fig. 2 and Fig. 3).

Most of the census tract level spatial dependence in the Bayesian space-time models was explained by the unstructured spatial random effect, suggesting that measuring spatial dependence at the scale of census tracts may not provide useful insights into the mechanism of pertussis transmission, especially in years when incidence is low. This was confirmed by Moran's I statistic which was only significant for outbreak years. Because we had individual-level point pattern data and pertussis is expected to transmit over short distances, the τ statistic is a useful method to measure short-term spatial dependency. It has been used previously to measure spatial dependency for dengue [75] [76], influenza [58], and HIV [77]. Significant spatial dependence up to 0.1 km meters is consistent with pertussis transmission within households and to immediate neighbors [78] [79].

Although we did not find significant spatial dependence between individual cases beyond 0.2 km, we found correlation in pertussis epidemics in neighboring public-school districts. Correlation in pertussis cases between pairs of school districts decreased with increasing distance between school districts. These patterns could be explained by movement of infected individuals between school districts. Infected children might travel to other school districts for sporting or social events, children might attend private schools in districts where they do not reside, and adults commute to metropolitan areas such as Seattle and surrounding cities for work [80]. Vaccinated individuals that may be asymptomatic or have reduced symptom severity might not change their behavior and come in contact with more individuals than a symptomatic case [16]. One case study found that casual contact from the community is responsible for a third of pertussis transmission events to young infants [81]. Spatial correlation between pertussis outbreaks persisted up to 4 weeks, which is not surprising given that the infectious period of pertussis is approximately 21 days. While a school district might be too large a spatial scale to estimate community transmission for pertussis as pertussis is mainly transmitted over short distances and there might be limited movement of children across school districts,

it was selected over census tracts because the time series data were very sparse at the census tract level.

A study of geographical analysis of pertussis and vaccination exemption clusters done at the national level identified King County as well as Seattle as hotspots of pertussis activity in the US [37]. Our study showed that there is spatial heterogeneity in pertussis incidence even within King County. Similar to other studies, we found an overlap between pertussis and non-medical exemptions clusters in school districts with high vaccine exemption rates such as Vashon Island (12.5% exempt), Auburn (5.7%), and Federal Way (5.6%), and Seattle (5.0%) school districts [82]. All space-time clusters were found in 1999 and 2012, again suggesting significant spatial dependence at census tract level was observed only during the large outbreak years. Spatial dependence at census tract level might be due to direct transmission as well as clustering of census tracts with high vaccine exemption. One limitation of this analysis is that NME clusters included only children attending K-12 grades in 2015-16 while the space-time pertussis clusters included pertussis cases of all ages from 1999 to 2017.

Underreporting of pertussis cases is a major limitation of this study. Pertussis is a nationally notifiable disease, but surveillance is passive. Reporting rate for pertussis also varies by age and likely by location and is estimated to be between 10-32% in the US [8]. There is evidence that asymptomatic cases can transmit infection [16], but do not get reported because they may not seek medical help. Active surveillance and intense contact tracing is needed to reduce this problem, but often these might happen only during ongoing epidemics. Clustering in this study was based on home location of cases. However, transmission can occur in other settings such as schools, day cares and work places, and clustering patterns might be different in these settings. Details on locations of school and work places of cases should be collected to analyze transmission in different settings.

In conclusion, even though pertussis incidence has increased among adolescents in King County, we did not find evidence for increase in overall incidence since 1999. Pertussis cases are highly spatially dependent at very small distances consistent with transmission to immediate neighbors. Human travel to neighboring school districts might explain correlated epidemics at school district level. This provides evidence for transmission of pertussis between communities. There is spatial heterogeneity in pertussis incidence within King County and detecting space-time clusters regularly can be beneficial to health departments to devise focused vaccination strategies [43].

2.7 Tables

Table 2.1: Moran's I statistic testing spatial autocorrelation in residuals of pertussis standardized incidence ratios

Year	**Moran's I**	**p-value**
1999	0.03	0.05
2000	0.04	0.06
2001	0.15	<0.01
2002	-0.04	0.9
2003	0.035	0.9
2004	-0.02	0.8
2005	-0.02	0.8
2006	0.0008	0.4
2007	0.02	0.8
2008	-0.002	0.5
2009	0.019	0.2
2010	-0.02	0.7
2011	-0.001	0.5
2012	0.05	0.03
2013	-0.014	0.6
2014	0.015	0.2
2015	0.04	0.05
2016	0.2	0.2
2017	-0.02	0.7

Poisson models were fit to estimate census-tract level pertussis incidence separately for each year adjusting for latitude and longitude of census tract centroids. Spatial autocorrelation was estimated using the residuals of these Poisson models.

Table 2.2: Pertussis space-time clusters within King County, Washington from 1999 to 2017 using the Kulldorff scan statistic

Cluster*	**Cluster centroid**		**Cluster radius**	**Start date**	**End date**	**p-value°**	**Observed number of cases**	**Expected number of cases**	**Relative risk±**	**Cluster description**
	Latitude	Longitude								
1	47.63	122.03	0 km°	1/1/99	4/30/99	<0.001	97	0.19	516.7	Sammamish area within Lake Washington school district
2	47.18	121.01	28.5 km	3/1/12	12/31/12	<0.001	259	33.84	8.13	Federal Way, Auburn
3	47.69	122.21	9.7 km	1/1/12	8/31/12	<0.001	203	26.57	7.64	Ballard, Phinney Ridge, University District, Maple Leaf, Northgate, Lake city, North-east Seattle
4	47.45	122.47	8.7 km	1/1/99	12/31/99	<0.001	48	1.41	34.5	Vashon Island
5	47.52	122.36	6.3 km	5/1/12	6/30/12	< 0.001	40	2.6	15.42	West Seattle, Beacon Hill
6	47.57	122.15	10.7	1/2/99	4/30/99	< 0.001	63	9.8	6 .44	Mouth Baker, Bellevue

∗ Cluster 1 is the most likely cluster and the rest are secondary clusters that do not overlap geographically with the most likely cluster

∘ P-value calculated by SatScan using Standard Monte Carlo method under the null hypothesis of complete spatial randomness. P-values of secondary clusters should be interpreted as the ability of the secondary cluster to reject the null hypothesis on its own strength, whether or not the more likely clusters are true clusters.

± Relative risk of pertussis inside the cluster versus outside the cluster.

° Radius is 0 km because only one census tract is included in this cluster.

Table 2.3: Non-medical vaccine exemption spatial clusters for school year 2015-2016 within King County, Washington using the Kulldorff scan statistic

Cluster*	**Cluster centroid**		**Cluster radius**	**Observed no.of exempted children**	**Expected no.of exempted children**	**Relative risk±**	**p-value°**	**Cluster description**
	Latitude	Longitude						
1	47.7	122.34	16.4 km	2057	1445.5	1.42	<0.001	Maple Leaf, Northgate, Lake city
2	47.45	122.47	0 km [o]	161	35.37	4.55	<0.001	Vashon Island
3	47.24	121.63	33.4 km	468	249.92	1.87	<0.001	Enumclaw
4	47.45	122.19	0 km	27	2.53	10.7	<0.001	Renton
5	47.50	122.38	2.2 km	70	26.86	2.61	<0.001	Burien
6	47.70	121.83	10.1 km	129	71.35	1.81	<0.001	Duvall, Carnation
7	47.30	122.31	0 km	38	12.1	3.16	<0.001	Federal Way

∗ Cluster 1 is the most likely cluster and the rest are secondary clusters that do not overlap geographically with the most likely cluster

◦ P-value calculated by SatScan using Standard Monte Carlo method under the null hypothesis of complete spatial randomness. P-values of secondary clusters should be interpreted as the ability of the secondary cluster to reject the null hypothesis on its own strength, whether or not the more likely clusters are true clusters.

± Relative risk of non-medical exemptions inside the cluster versus outside the cluster.It is calculated as observed number of exempted children/ expected number of exempted children

[o] Cluster radius is 0 km because cluster includes only one census tract.

2.8 Figures

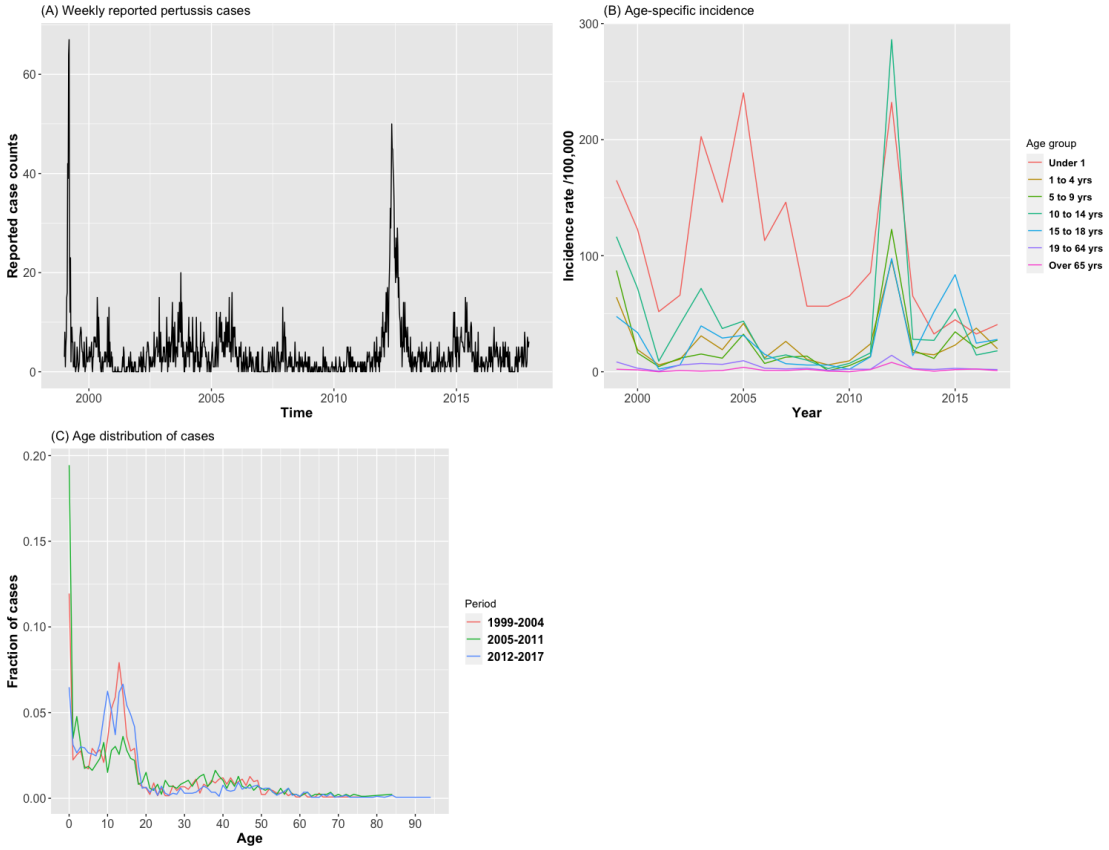


Figure 2.1: Temporal trends in pertussis in King County: 1999-2017

Fig 1(A) Weekly reports of pertussis cases between 1999 and 2017 in King County, WA; (B) Incidence of pertussis cases by age group plotted against time; (C) Age distribution of pertussis cases over three periods (1999-2004, 2005-2011, 2012-2017)

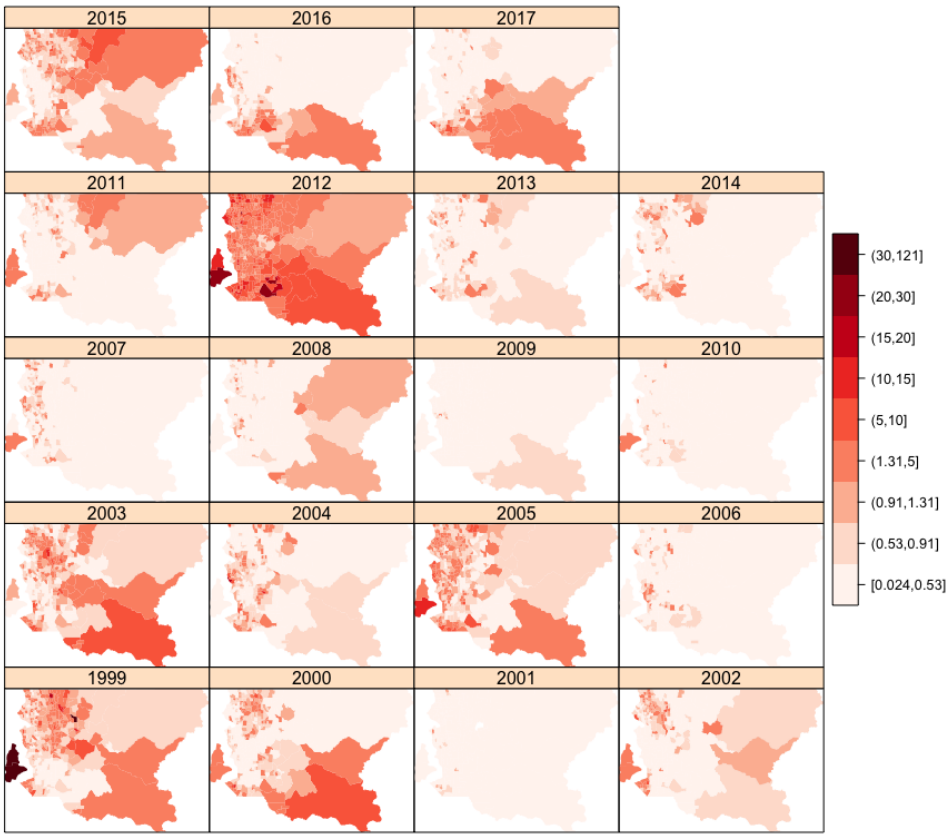


Figure 2.2: Posterior median estimates of pertussis relative risk from Bayesian spatio-temporal models with Type III interaction in King County, WA from 1999 to 2017

The pertussis risk estimates are smoothed in space and time over the 19 year period. This map shows variation in spatial clustering in pertussis risk in each year. The model is smoothed in space using intrinsic conditional autoregressive (ICAR) model and smoothing in time using random walk of second order (RW2).

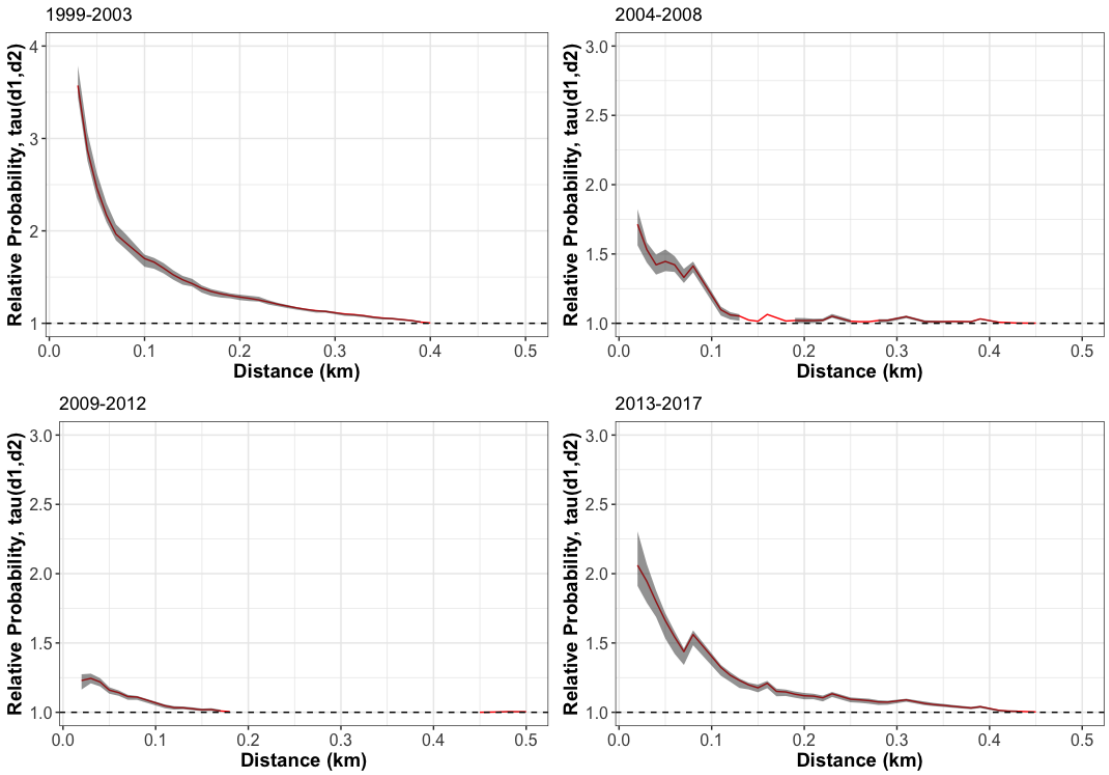


Figure 2.3: Spatial dependence between individual pertussis cases

τ statistic results for pertussis cases in five- year windows (except 2009-2012). The solid red line represents relative probability of cases occurring within the same month separated by distance $d1 - d2$. The size of the spatial window ranges from 0.01 km to 0.5 km and $d1$ is always equal to 0.01 km (10 m). The shaded area represents bootstrapped 95% confidence intervals for τ generated by resampling cases with replacement 500 times .

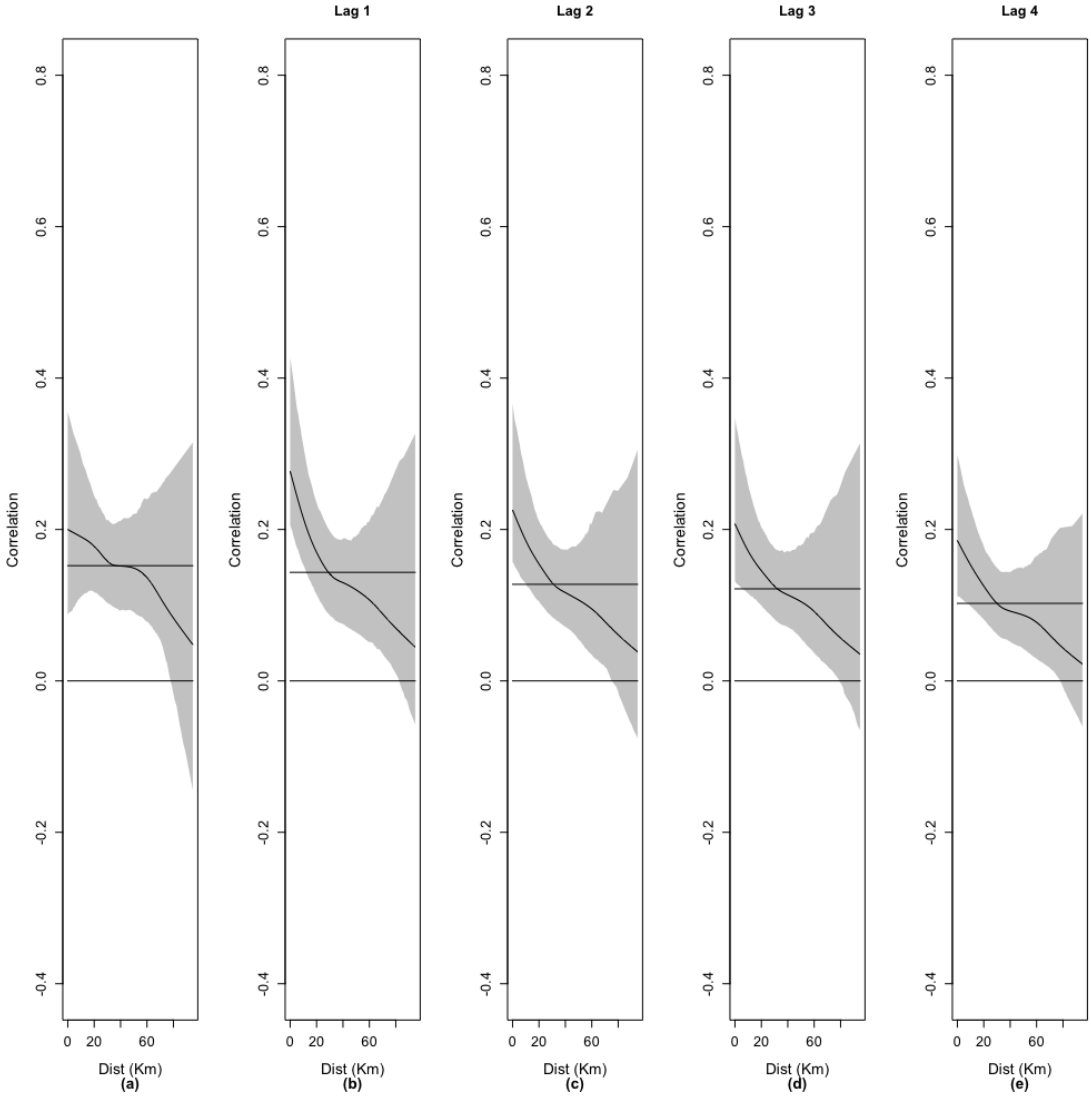


Figure 2.4: Spatial synchrony in pertussis time series between pairs of school districts in King County

The (a) non-parametric spatial covariance function of pertussis time series plotted against distance between school districts, and cross-correlation between pertussis time series lagged by (b) 1 week, (c) 2 weeks, (d) 3 weeks, and (e) 4 weeks plotted against distance between school districts. The grey shaded regions represent 95% bootstrapped confidence intervals. Distance between a given pair of school districts is calculated as Euclidean distance between their centroids. Only school districts with more than 50 cases were used for this analysis (one school district with only 14 cases between 1999 and 2017 omitted).

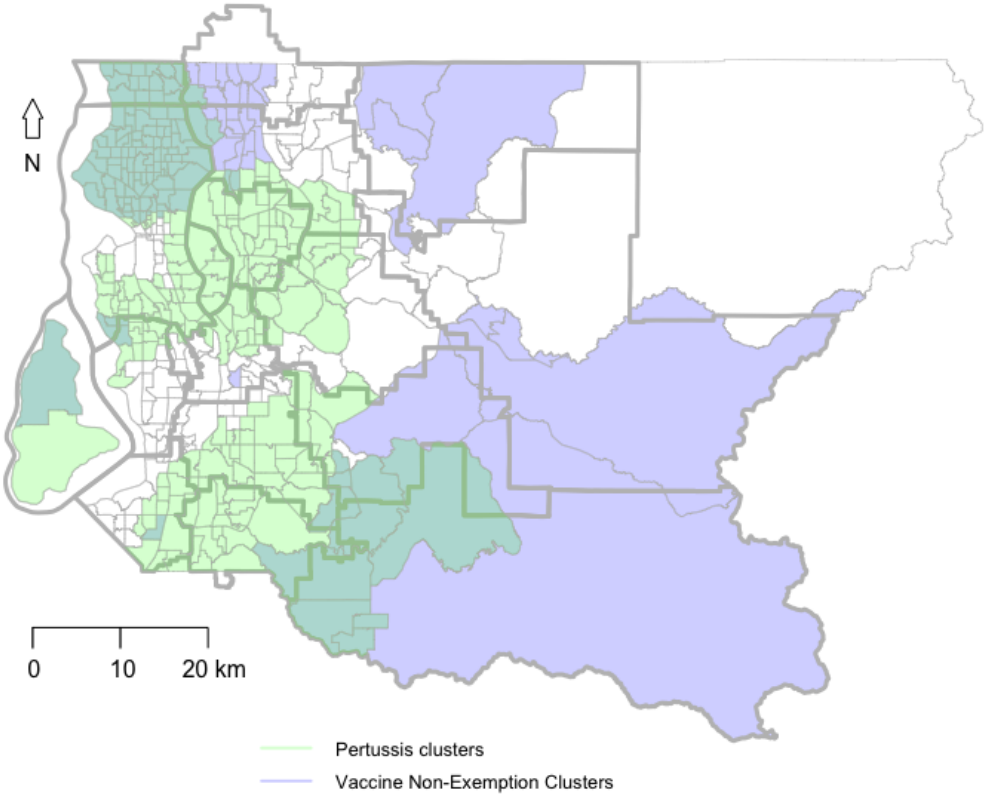


Figure 2.5: Overlap between space-time clusters of pertussis and spatial clusters of non-medical exemptions in King County, WA

Space-time clusters of census tracts with high pertussis activity within each school district are shown in green while spatial clusters of non-medical exemptions (NMEs) are in purple. Areas of overlap are darkened. Light grey depicts boundaries of census tracts and dark grey depicts boundaries of school districts.

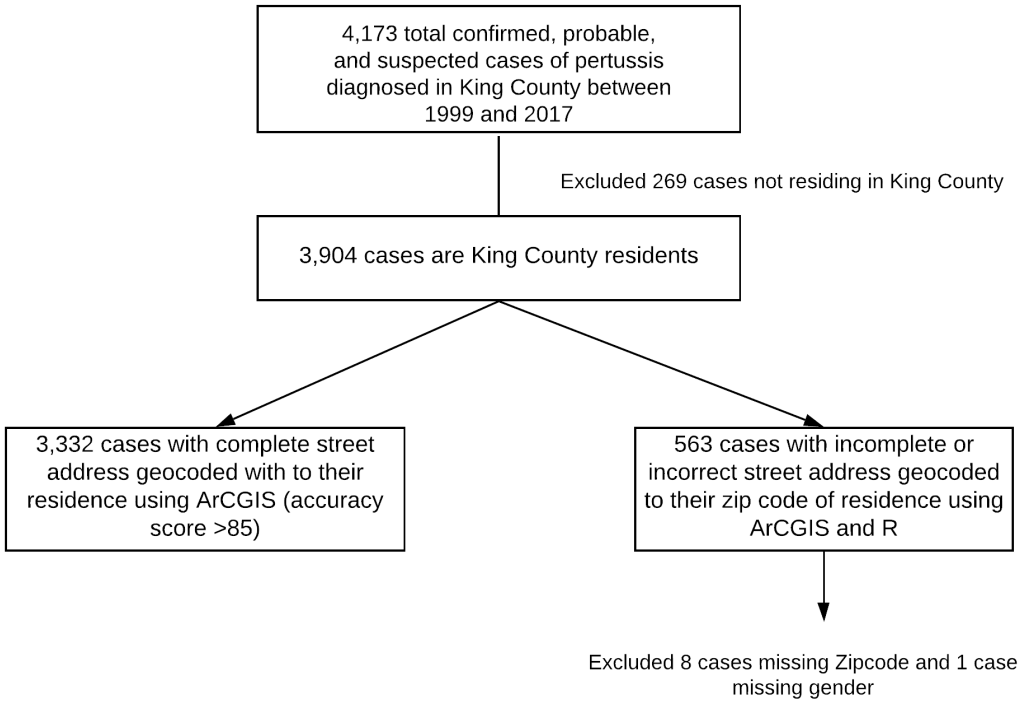
2.9 Supplementary Materials

2.9.1 Case Classification of Pertussis

The clinical case definition of pertussis is a cough illness lasting $\geq$ 2 weeks with at least one of the following: paroxysms of coughing or inspiratory "whoop" or post-tussive vomiting, or apnea (with or without cyanosis) for infants aged <1 year only. Cases were further classified as suspected, probable, or confirmed based on the CDC recommendations [45]:

Suspected	**Probable**	**Confirmed**
A case that is polymerase chain reaction (PCR) positive but does not meet the clinical case definition *or* An epidemiologically-linked case with a 2-week cough but without any other case-defining symptoms. (Though not included in the CDC pertussis case classifications, WA-DOH collects data on PCR-positive suspect cases and requests that local health jurisdictions (LHJ) complete case reports on these cases) [2]	A case that meets clinical definition and is not laboratory-confirmed with culture or PCR and is not epidemiologically-linked directly to a confirmed case.	Acute cough illness of any duration with isolation of B. pertussis from a clinical specimen. *or* A case that meets the clinical case definition and is confirmed by detection of *B. pertussis*-specific nucleic acid by PCR *or* A case that meets the clinical case definition and is epidemiologically-linked directly to a laboratory-confirmed case of pertussis.

2.9.2 Flowchart of the geocoding process using ArcGIS and R



2.9.3 Vaccination status of pertussis cases

Vaccination status of pertussis cases was determined using the case's age, the number of pertussis vaccines received before they became a case, and the date they received their last vaccine dose. Information on pertussis-containing vaccines that cases received was collected by the PHSKC surveillance staff as part of outbreak investigation using the WA state immunization registry, medical records, patient vaccination cards, patient self-report, or immunization registries of other states, if available. For this study, cases were classified as "up-to-date" (UTD) if they received the appropriate number of pertussis vaccine doses for their age, and "not up-to-date" (not UTD) otherwise.
By general consensus, three doses of DTaP are required to build full protection against pertussis diseases. Hence, infants under 6 months of age are considered not "not up-to-date" in this study because they have not yet received the full 3 dose series. Due to high frequency of missing data on pertussis vaccinations in adults, adults over 18 years of age were considered UTD if their vaccine information for missing. We assumed that older adults are likely to have been either exposed to pertussis or vaccinated during their lifetime, so they would have some degree of immunity to pertussis. Vaccination status of cases 18 years and younger with missing data on number and timing of vaccine doses (n=106) was categorized as "Unknown".

2.9.4 Bayesian hierarchical spatio-temporal model for pertussis

Let Y_{ij} and N_{ij} represent the number of pertussis cases and population at risk in stratum j, $j = 1, \ldots, J$, and area i, $i = 1, \ldots m$. We stratified pertussis counts by 7 age groups (under 1, 1 to 4, 5 to 9, 10 to 14, 15 to 18, 19 to 64, 65 and above) as different areas will contain different proportion of individuals in each stratum and not accounting for this will lead to biased conclusions. p_j is the risk of disease in age-group j which is calculated as (pertussis cases in an age group)/(total individuals of that age group at risk) in King County between 1999 and 2017. We assume that the population of census tracts does not change over time. Then, expected number of pertussis cases in area i is given by $E_i = \sum_j p_j N_{ij}$ and $Y_i = \sum_j Y_{ij}$.

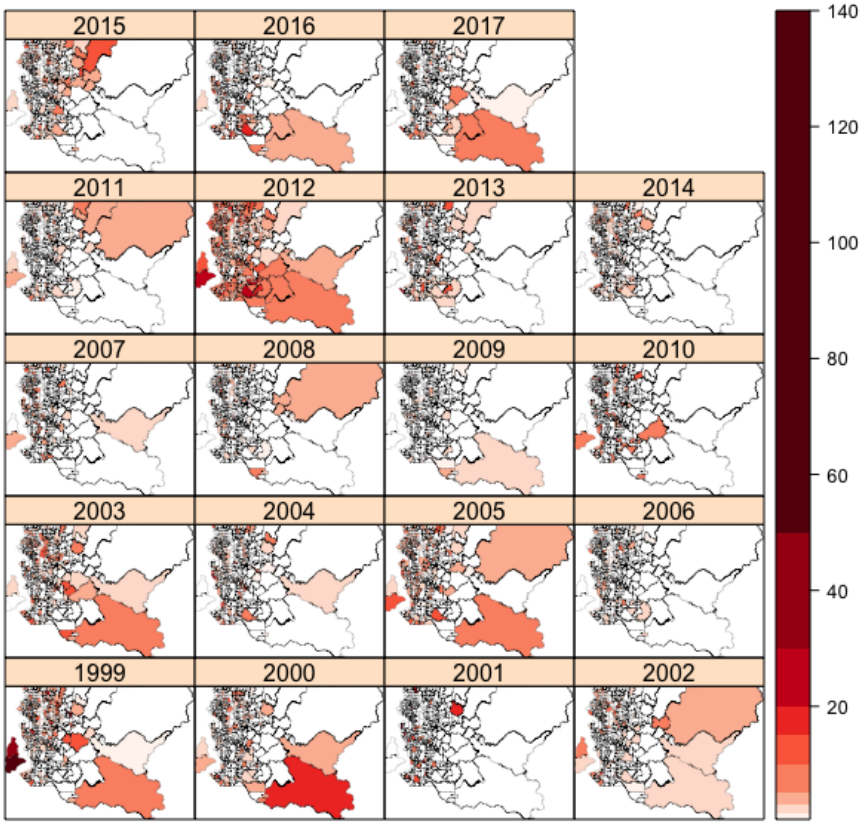
The number of pertussis cases in each census tract (i=1,...,397) and year (t=1,...,19) follows a Poisson distribution with relative risk λ_{it} such that

$$Y_{it}|\lambda_{it} \sim Poisson(E_i\lambda_{it})$$

The raw SIRs calculated as $\lambda_{it} = \frac{Y_{it}}{E_i}$ and the expected number of pertussis cases summed

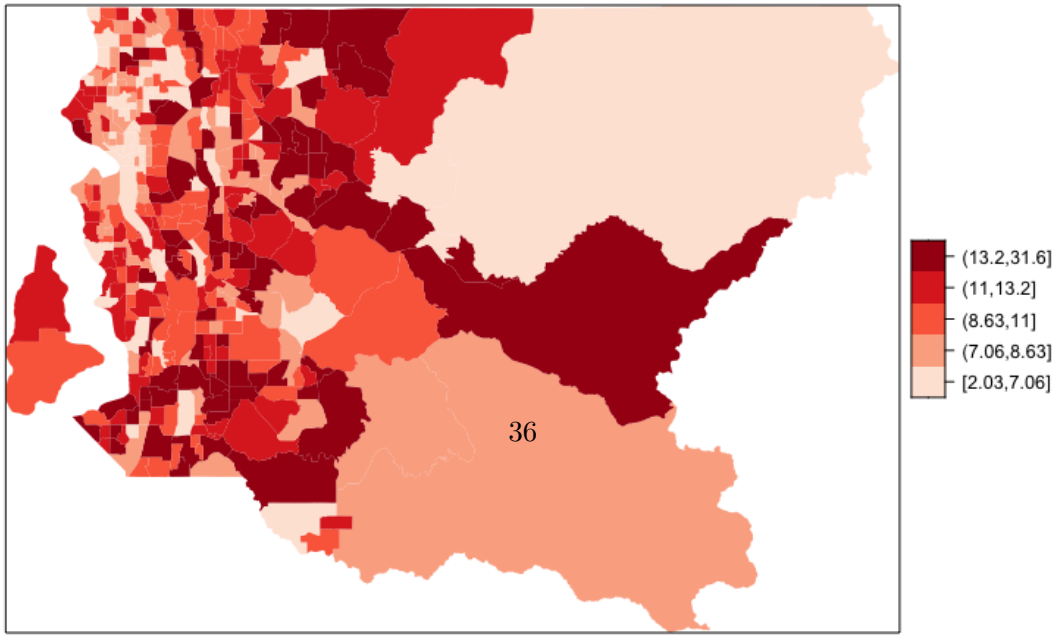
over the entire time period are plotted in Supplementary Fig 4a and 4b. Because pertussis is a rare disease and we have a fairly small spatial scale of census tracts, the data contains many zeros, which are plotted in white.

Figure 1 Spatial distribution of pertussis and expected counts in King County, WA



(a) Raw age-standardized incidence ratios (SIRs) of pertussis at census tract level in King County, WA from 1999 to 2017.

Large outbreaks occurred in 1999 and 2012, and smaller outbreaks in 2005 and 2015. Pertussis relative risk in Vashon Island was high in all outbreak years. Pertussis relative risk is high in the northern and southern parts of King County. Sampling variability could be an issue in the large census tracts to the east of the county where population size and consequently expected numbers are very low. There are fewer than 5 expected cases in the large census tracts to the east resulting in spuriously high SIRs.



(b) Expected pertussis cases at census tract level in King County, WA summed over the time period 1999 to 2017

To obtain Bayesian smoothed estimates of pertussis SIRs, we first fit a Bayesian spatio-temporal model without an interaction which can be written as,

$$log(\lambda_{it}) = \alpha + \mu_i + \nu_i + \gamma_t + \phi_t$$

where α is the intercept, μ_i is the spatially structured random effect, ν_i is the spatially unstructured random effect, γ_t is the temporally structured random effect, ϕ_t is the temporally unstructured random effect. The unstructured temporal term is modeled as $\phi_t \sim_{iid} N(0, \sigma_\phi^2)$, $t = 1, ..., T$ while γ_t which is the smooth temporal term is given a second-order random walk prior RW2. The unstructured spatial term ν_i is assigned a $N(0, \sigma_\nu^2)$ prior, and the structured spatial term μ_i is assigned an intrinsic conditional autoregressive (ICAR) prior.

We plotted structured and unstructured spatial and temporal effects from the model without interaction in Supplementary Fig 1.9.4.2. The unstructured temporal random effect is the residual temporal variability after accounting for structured temporal effect common to all areas. The structured random effect shows a small decrease in overall pertussis over 19 years. However, the range of the estimate of γ_t is narrow (-0.05- 0.08). This means that there is no appreciable overall change in pertussis incidence between 1999 and 2017 in King County. Plots of ICAR spatial risk $exp(\mu_i + \nu_i)$ show clustering in the northern and southern regions of the county, as well as elevated risk in Vashon Island. Spatially unstructured random effects range from -1.1 to 2.2, while ICAR spatial random effects range from -0.3 to 0.6. Only 20.2% of the total variation in the model is explained by the ICAR component. While the unstructured spatial random effects dominate, there is evidence of some clustering in the map.

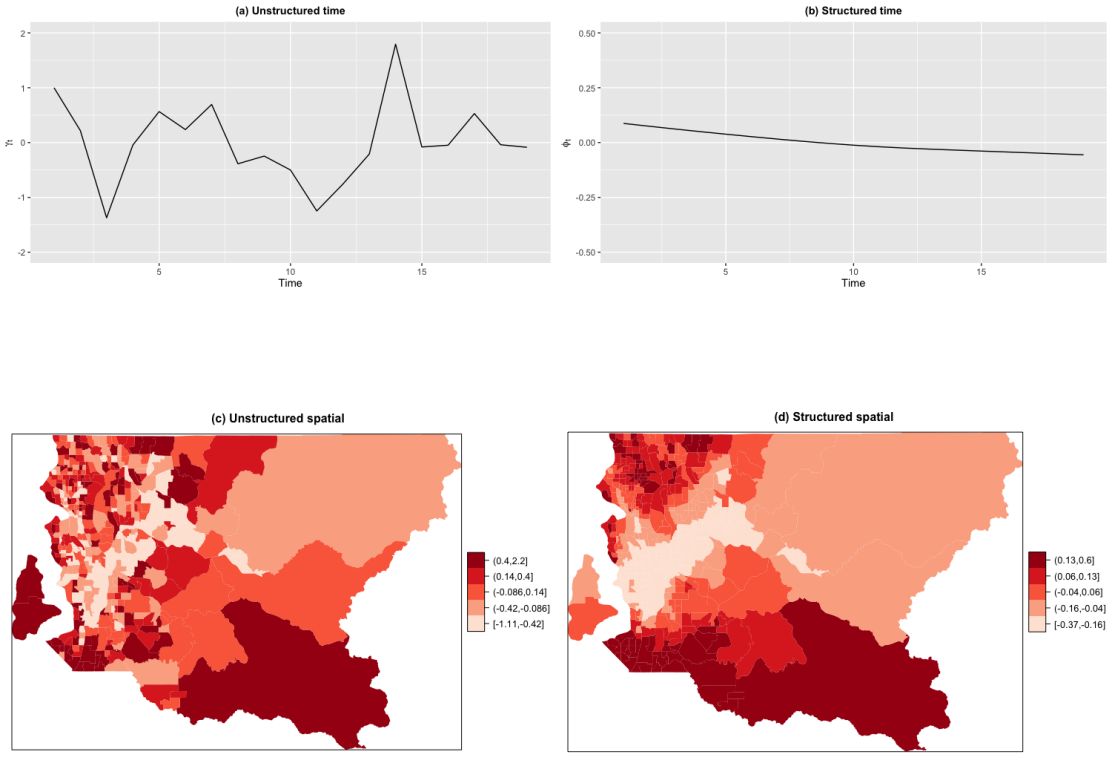


Fig 2 Estimated temporal components ϕ_t (a) and γ_t (b) , and spatial component ν_i (c) and μ_i (d) from the model without interaction for pertussis in King County, WA.

Spatio-temporal structure of the interaction can be of four types: Type I assumes that the two unstructured random effects ν_i and ϕ_t interact, thus there is no spatial and/or temporal structure on the interaction. This is not very realistic for modeling infectious disease spread. Type II assumes interaction between structured temporal effects γ_t and unstructured spatial effects ν_i. This assumes that there is an autoregressive structure on the time component in each census tract that is independent of other census tract. Type III combines the structured spatial component μ_i with the unstructured temporal component ϕ_t. This assumes that there is clustering in each year independent of other years. Finally the most complicated Type IV interactions assumes that the structured temporal and spatial random effects μ_i and γ_t interact such that temporal dependency in each area depends on the temporal patterns of the neighboring areas as well. The Bayesian model with an interaction term can be written as:

$$log(\lambda_{it}) = \alpha + \mu_i + \nu_i + \gamma_t + \phi_t + \delta_{it}$$

where δ_{it} is the interaction term. Because there was no temporal variation and some spatial clustering in random effects from the model without interaction terms, we added a Type III interaction term to the main model asuming that pertussis incidence clusters in space at a given time period independently of other time periods. The results from the Type III interaction are in Supplementary Fig 1.9.4.3 below. The temporal random effects look similar to the no-interaction model although the unstructured effect is now negligible. The ICAR spatial random effects are also negligible (range -0.002 to 0.003), as most of the spatial variation is explained by the unstructured spatial component.

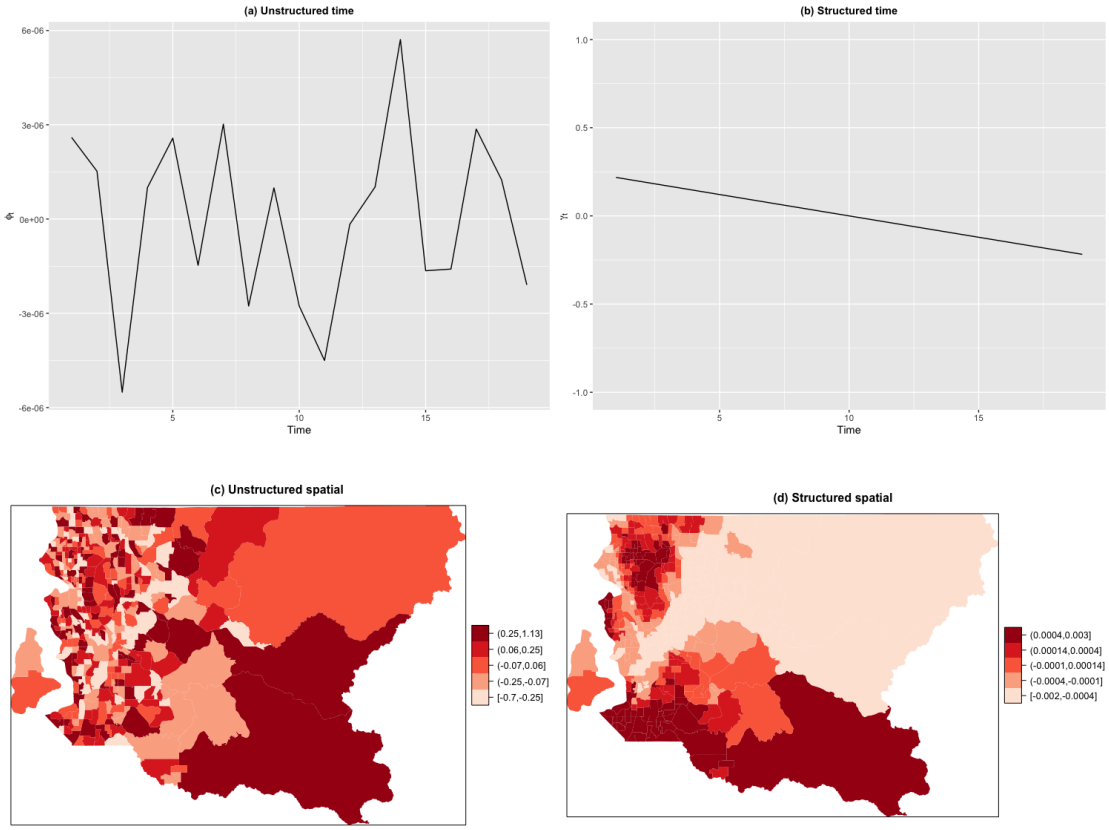


Figure 3 Estimated temporal components ϕ_t (a) and γ_t (b) , and spatial component ν_i (c) and μ_i (d) from model with Type III interaction for pertussis in King County, WA.

Posterior median estimates of the interaction term δ_{it} for all years are plotted in Supplementary Fig 1.9.4.4. This can be interpreted as the residual area-level relative risk of pertussis after accounting for unstructured and unstructured spatial and temporal random effects.

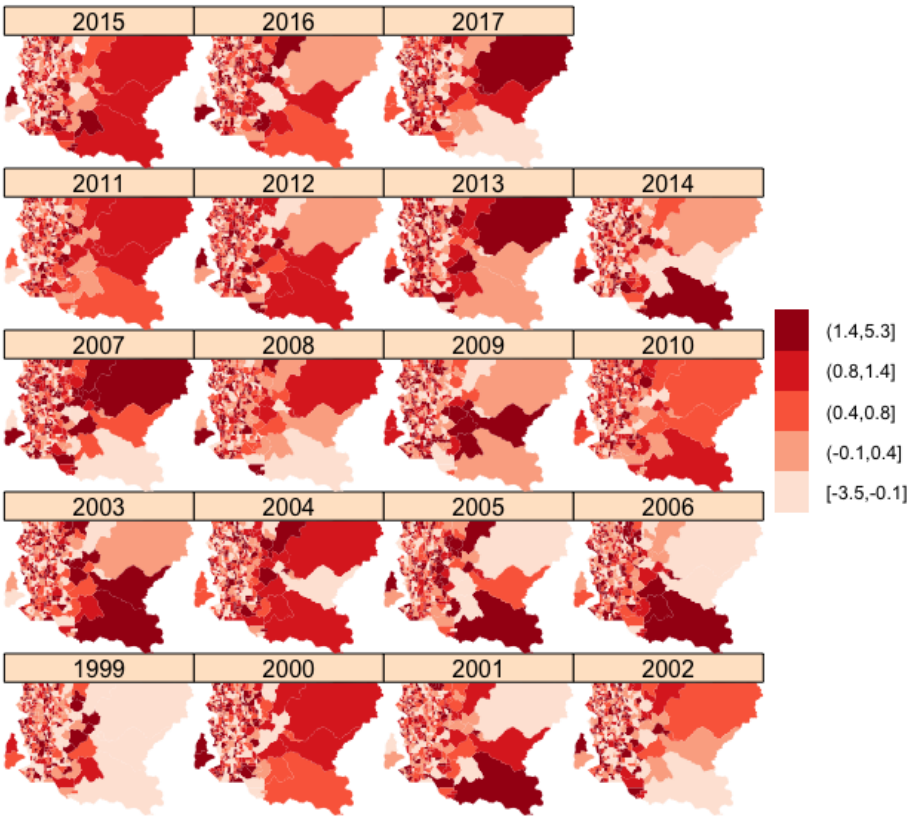
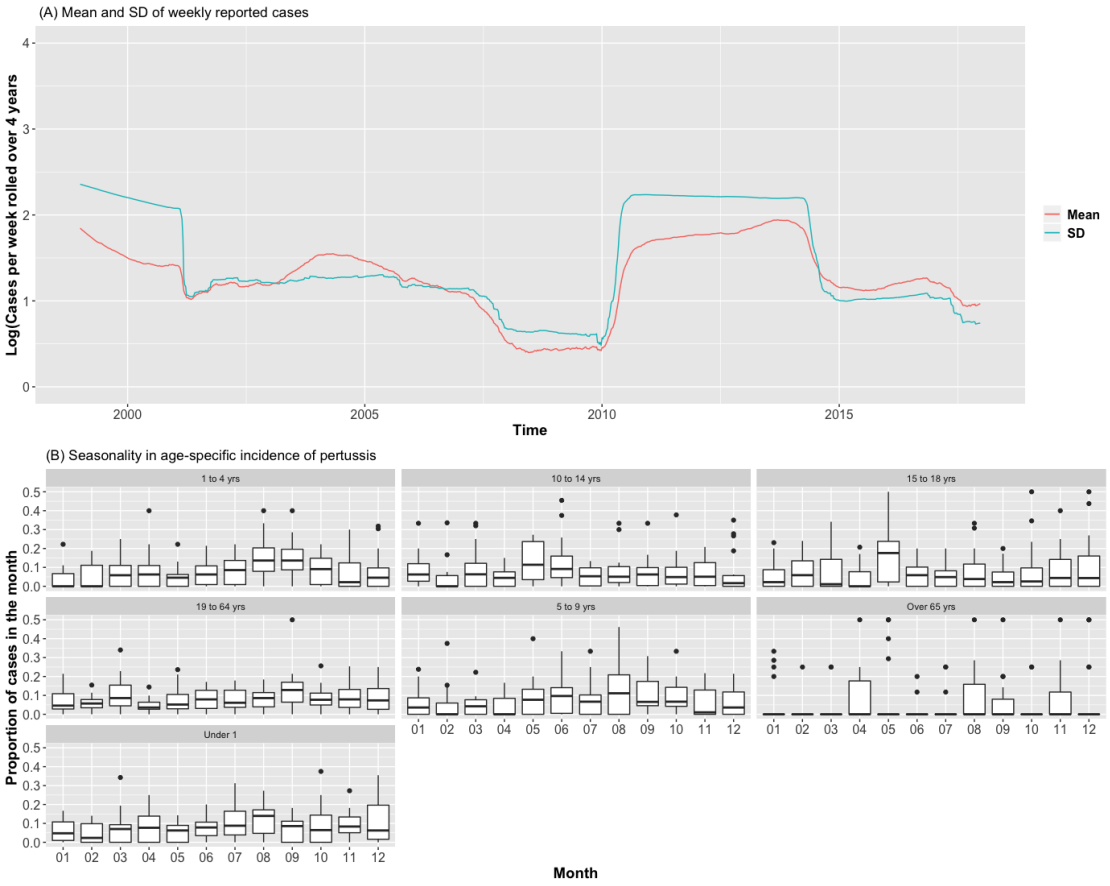


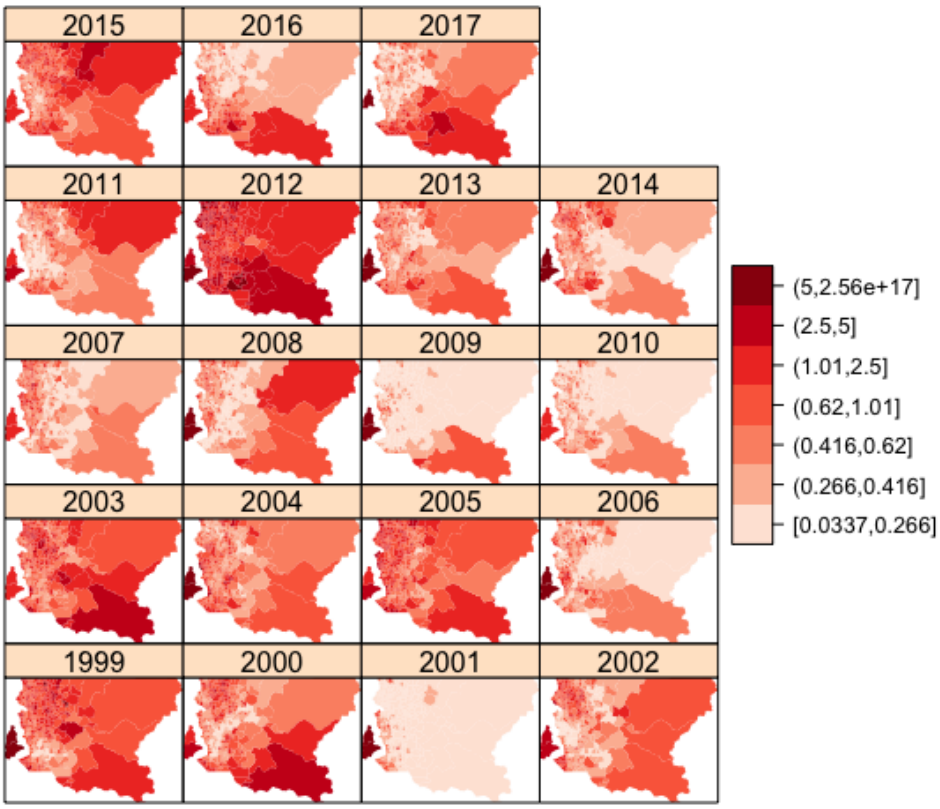
Figure 4 Posterior median estimates of the interaction term δ_{it} from the model with Type III interaction for pertussis in King County, WA.

2.9.5 Temporal trends in pertussis incidence



(A) Mean and Standard Deviation in the number of weekly cases calculated as a four-year rolling window against time; (B) Seasonality in age-specific pertussis incidence in King County, WA

2.9.6 Uncertainty in posterior median relative risk of pertussis



Standard deviations of fitted posterior mean relative risk estimates from Bayesian spatio-temporal models with Type III interaction term. Variance is high for relative risk estimates in Vashon Island especially in years of low incidence.

2.9.7 Supplementary Tables

Table1 : Pertussis case characteristics by vaccination status

Characteristics	**Up to date (N=2,357)**	**Not up to date (N=1,432)**	**Unknown (N=106)**	**Total (N=3,895)**
Age (in years), N(%)				
Under 1	43 (9.8)	394 (90.2)	0 (0.0)	437
1 to 4	228 (52.1)	191 (43.6)	19 (4.3)	438
5 to 9	337 (64.3)	161 (30.7)	26 (5.0)	524
10 to 14	693 (71.3)	245 (25.2)	34 (3.5)	972
15 to 18	322 (66.4)	136 (28.0)	27 (5.5)	485
19 to 64	696 (71.5)	277 (28.5)	0 (0.0)	973
65 and over	38 (57.6)	28 (42.4)	0 (0.0)	66
Sex, N (%)				
Male	1052 (58.0)	715 (39.5)	46 (2.5)	1813
Female	1305 (62.7)	717 (34.4)	60 (2.9)	2082
Race, N(%)				
White	1782 (61.7)	1024 (35.5)	81 (2.8)	2887
Black	94 (49.2)	94 (49.2)	3 (1.6)	191
Asian	137 (61.2)	84 (37.5)	3 (1.3)	224
Nat Am	17 (51.5)	16 (48.5)	0(0.0)	33
Other	112 (52.8)	98 (46.2)	2 (0.9)	212
NA/Unk	215 (61.8)	116 (33.3)	17 (4.9)	348
Ethnicity, N(%)				
Hispanic	217 (57.9)	152 (40.5)	6 (1.6)	375
Non-hispanic	1694 (58.2)	1141 (39.2)	77 (2.6)	2912
Unk	446 (73.4)	139 (22.9)	23 (3.7)	608
Hospitalized, N(%)				
Yes	27 (13.8)	168 (86.2)	0	195
No	2315 (62.9)	1260 (34.2)	104 (2.9)	3679
Unk/NA	15 (71.4)	4 (19.1)	2 (9.5)	21

Cough, N(%)				
Yes	2334 (60.7)	1412 (36.7)	99 (2.6)	3845
No	14 (48.3)	13 (44.8)	2 (6.9)	29
Unk/NA	9 (42.9)	7 (33.3)	5 (23.8)	21
Cold, N(%)				
Yes	1365 (61.7)	788 (35.6)	59 (2.6)	2212
No	876 (59.2)	570 (38.5)	33 (2.2)	1479
Unk/NA	116 (58.8)	74 (34.3)	14 (6.9)	204
Whoop, N(%)				
Yes	529 (58.3)	350 (38.5)	29 (3.2)	908
No	1757 (61.4)	1038 (36.3)	66 (2.3)	2861
Unk	71 (60.3)	44 (31.0)	11 (8.7)	126
Duration of cough (median days)	18	18	21	
Seizure, N(%)				
Yes	3 (27.3)	8 (72.7)	0 (0.0)	11
No	2315 (60.7)	1400(36.7)	97 (2.6)	3812
Unk	39 (58.3)	24 (29.2)	9 (12.5)	72
Encephalitis, N(%)				
Yes	6 (33.4)	11 (61.1)	1 (5.5)	18
No	2320 (60.7)	1402 (36.8)	97 (2.5)	3819
Unk	31 (56.9)	17 (29.3)	8 (13.8)	58
Apnea, N(%)				
Yes	953 (54.1)	766 (43.5)	41(2.4)	1760
No	1335 (66.2)	629 (31.2)	53 (2.6)	2017
Unk	69 (58.5)	37 (31.4)	12 (10.1)	118
Cyanosis, N(%)				
Yes	96 (31.7)	200 (66.0)	7 (2.3)	303
No	2170 (63.2)	1178 (34.3)	86 (2.5)	3434
Unk	101 (63.9)	44 (27.9)	13 (8.2)	158

Duration of hospitalization (Median days)	3	4	–	
Post-tussive vomiting, N(%)				
Yes	1129 (58.0)	764 (39.3)	53 (2.7)	1946
No	1187 (63.4)	644 (34.4)	42(2.2)	1873
Unk	42 (55.3)	23 (30.2)	11 (14.5)	76
X-ray positive for pneumonia, N(%)				
Positive	28 (45.2)	34 (54.8)	0 (0.0)	62
Negative	189 (52.8)	13 (3.6)	6 (1.6)	358
Not done	2131 (61.7)	1225 (35.5)	99 (2.8)	3455
Unk	10 (50.0)	9 (45.0)	1 (5.0)	20

% represent row percentages for categorical variables

Table 2: Pertussis case symptoms by age group (in years)

	Under 1 (N=437)	**1 to 4 (N=438)**	**5 to 9 (N=524)**	**10 to 14 (N=972)**	**15 to 18 (N=485)**	**19 to 64 (N=973)**	**Over 64 (N=66)**	**Total (N=3895)**
Cough, N(%)								
Yes	431 (98.6)	434 (99.1)	518 (98.9)	959 (98.6)	483 (99.6)	958 (98.5)	62 (94)	3845
No	5 (0.1)	3 (0.01)	3 (0.01)	6 (0.01)	0 (0.0)	10 (1.0)	2 (3.0)	29
Unk	1	1	3	7	2	5	2	21
Whoop, N(%)								
Yes	114 (26.1)	106 (24.2)	107 (20.4)	206 (21.2)	110 (22.7)	249 (25.6)	16 (24.2)	908
No	307 (70.2)	318 (72.6)	402 (76.7)	734 (75.6)	363 (74.8)	693 (71.2)	44 (66.6)	2861
Unk	16	14	15	32	12	31	6	126
Cold, N(%)								
Yes	287 (65.6)	235 (53.6)	281 (53.6)	564 (58.0)	255 (52.6)	553 (56.9)	37 (56.1)	2212
No	132 (30.2)	179 (40.9)	217 (41.4)	365 (37.6)	197 (40.6)	363 (37.3)	26 (39.4)	1479
Unk	18	24	26	43	33	57	3	204
Paroxysm, N(%)								
Yes	407 (93.1)	388 (88.6)	439 (83.8)	855 (88.0)	444 (91.5)	882 (90.7)	59 (89.4)	3474
No	26 (5.9)	41 (9.4)	77 (14.7)	98 (10.0)	34 (7.0)	78 (8.0)	3 (4.5)	357
Unk	4	9	8	19	7	13	4	64
Apnea, N(%)								
Yes	322 (73.5)	212 (48.4)	181 (34.5)	352 (36.2)	199 (41.0)	463 (47.6)	31 (46.9)	1760
No	110 (25.1)	210 (47.9)	331 (63.2)	575 (59.2)	276 (56.9)	484 (49.7)	31 (46.9)	2017
Unk/NA	15	16	12	45	10	26	4	118
Seizure, N(%)								
Yes	3 (0.01)	1 (0.0)	0(0.0)	2 (0.0)	1(0.0)	4 (0.0)	0 (0.0)	11

No	1429 (97.9)	430 (98.2)	512 (97.7)	948 (97.5)	476 (98.1)	954 (97.1)	63 (95.4)	3812
Unk	75	7	12	22	8	15	3	72
Hospitalization, N(%)								
Yes	163 (37.2)	7 (1.6)	1 (0.0)	4 (0.0)	1 (0.0)	14 (1.4)	5 (7.5)	195
No	272 (62.1)	431 (98.4)	520 (99.2)	961 (98.9)	483 (99.6)	951 (97.7)	61 (92.4)	3679
Unk 2	0	3	7	1	8	0	21	
Post-tussive vomiting, N(%)								
Yes	332 (75.8)	261 (59.6)	222 (42.4)	443 (45.6)	212 (43.7)	456 (46.9)	20 (30.3)	1946
No	98 (22.4)	168 (38.3)	292 (55.7)	508 (52.3)	262 (54.0)	503 (51.7)	42 (63.6)	1873
Unk	7	9	10	21	11	14	4	76
X-ray positive for pneumonia, N(%)								
Positive	21 (4.8)	8 (1.8)	6 (1.1)	8 (0.01)	0 (0.0)	15 (1.5)	4 (6.1)	62
Negative	111 (25.3)	26 (5.9)	28 (5.3)	52 (5.3)	26 (5.4)	106 (10.9)	9 (9.3)	358
Unk	8	0	1	3	3	4	1	20
Not done	297 (67.8)	404 (92.2)	489 (93.3)	909 (93.5)	456 (94.0)	848 (87.1)	52 (78.8)	3455

% represent column percentages for categorical variables

Chapter 3

Association between pertussis vaccine coverage and other sociodemographic factors and pertussis incidence using surveillance data

Abstract

Mass vaccination with an effective vaccine has been successful in driving down pertussis mortality and morbidity globally. However, despite high vaccination coverage, countries such as Australia, USA, and UK have experienced increase in pertussis activity over the last few decades. This may be due to local pockets of low vaccination coverage that result in persistence of pertussis in the population and occasionally lead to large outbreaks. The objective of this study was to characterize the association between pertussis vaccine coverage and sociodemographic factors and pertussis incidence at the school district level in King County, Washington. We used monthly pertussis incidence data reported to the Public Health Seattle and King County to get school district level pertussis incidence. We obtained immunization data from the Washington State Immunization Information System to estimate school-district level vaccination coverage as proportion of 19-35 month old children fully vaccinated with ≥ 4 doses of the Diphtheria-Tetanus-acellular-Pertussis vaccine in a school district. We used the ecological vaccine

model and the epidemic-endemic model to explore this association and compared and contrasted the two approaches. Even though the effect of vaccination is modeled differently in the two approaches, both models showed that vaccination coverage is associated with pertussis incidence. Using the ecological vaccine model, we estimated the vaccine effectiveness of 4 doses of Diphtheria-Tetanus-acellular-Pertussis vaccine to be 83% (95% credible interval: 63%, 95%). In the epidemic-endemic model, under-vaccination was statistically significantly associated with epidemic risk of pertussis (adjusted Relative Risk, aRR: 2.76; 95% confidence interval: 1.44, 16.6). Household size and median income were statistically significantly associated with endemic pertussis risk. However, the epidemic-endemic model suffers from ecological bias, whereas the ecological vaccine model provides less biased and more interpretable estimates of epidemiological parameters of pertussis, such as DTaP vaccine effectiveness and effective reproduction number of pertussis for each school district.

3.1 Introduction

Pertussis is a highly transmissible infectious disease caused by bacterium *Bordetella pertussis* [27]. There are currently two highly efficacious vaccines used worldwide- the Diphtheria-Tetanus-whole-cell-Pertussis (DTwP) and Diphtheria-Tetanus-acellular-Pertussis (DTaP) vaccines. A meta-analysis of recently published studies found that overall vaccine efficacy of DTaP vaccines is 84% (95% CI: 81%, 87%) while that of DTwP is 95%(88%, 97%) [83]. Global average vaccination coverage for the three-dose primary DTP vaccine series was 90% in 2019 in 125 countries as reported by the WHO [84]. However, despite high vaccination coverage with an effective vaccine, we have not yet been able to eradicate pertussis. In fact, several countries have experienced a resurgence in pertussis.

Aggregating estimates of vaccination coverage at the national or state level may hide local pockets of low vaccination. Sub-optimal local vaccination coverage could result in accumulation of susceptibles over time and an introduction of pertussis case in these areas would then result in an outbreak. Pockets of low vaccination coverage especially in areas of high population density could result in pertussis persistence in these areas [85]. It is of public health interest to investigate whether local cold-spots of vaccination coverage are driving pertussis incidence.

Previous studies on effects of pertussis vaccination on pertussis incidence have shown

heterogeneities in DTaP vaccination coverage at the local level and a significant association between area-level vaccination coverage and sociodemographic factors and pertussis incidence [86] [87] [88]. However, these studies were ecological analyses and failed to account for the dependent nature and the complex, non-linear dynamics of pertussis [89]. Analysis of aggregated infectious disease data using ecological regression methods results in ecological bias where inference at the group level may not apply at the individual level [90] [91]. Disease surveillance systems at the local, state, and national levels can readily provide incidence data for a large population over long periods of time. Methods have been developed to analyse the association between area-level vaccination coverage and infectious disease incidence that address some of these issues using surveillance data marked in space and time [92] [93].

The goal of this analysis was to explore the association between sub-county level pertussis vaccination coverage and socio-demographic factors and pertussis incidence using surveillance, vaccination, and demographic data aggregated over school districts within King County. We used the ecological vaccine model [92] and the epidemic-endemic model [93] to address this question. Both approaches use multivariate time series of infectious disease data and can model occasional large outbreaks beyond regular endemic behavior. Additionally, vaccination coverage can be included in these models as a covariate to study its effect on pertussis outbreaks. Additionally, the ecological vaccine model may reduce ecological bias when using aggregated data.

3.2 Methods

3.2.1 Data

Pertussis case data was obtained from the Public Health Seattle and King County (PHSKC) Department of Communicable Diseases and Immunizations. Pertussis is a nationally notifiable disease and cases are reported to PHSKC by primary care providers and diagnostic laboratories within 24 hours of detection. Suspected, probable, and confirmed pertussis cases of all ages diagnosed in King County between 1st January, 2010, and 31st December, 2017 were included in this analysis [94]. During this period, there was one large pertussis outbreak in King County in 2012 with 895 cases. Each case reported to PHSKC had a home address and date of diagnosis associated with it, alongside other demographic and epidemiologic information such as site of exposure, clinical symptoms,

laboratory test results, vaccination status, number of Diptheria-Tetanus-acellular Pertussis (DTaP) doses, date of last DTaP dose, hospitalization status, antibiotics administered, vaccine exemption status, and reason for exemption. Although we had access to individual level data on pertussis cases, data is often aggregated to some spatial level (for example, health district, school district, or county) to protect patient confidentiality. We aggregated pertussis cases at the school district level in this study (N=18). Enumclaw, Tukwila, and Vashon Island school districts are the three least populated school districts, while Seattle school district is the most populated (Table 3.1). Because pertussis is a rare disease, many school districts recorded zero cases during several months. Tukwila school district recorded only 18 cases over the span of 10 years. A map of King County school districts is in Supplementary Material 3.7.2 Fig 1.

3.2.2 Geocoding and aggregating pertussis cases at school-district level

We used ArcGIS 10.1 for geocoding cases' home addresses [47]. When address was missing or incorrect, we used the zipcode for geocoding. ArcGIS creates a new point shapefile of the geocoded home locations which can be spatially overlaid on a shapefile of King County school districts to aggregate pertussis cases to school district level. School district level population estimates and other demographic factors such as race, education, median income, household size, proportion of population foreign born, proportion of non-citizens, and proportion of population speaking languages other than English at home for the census year 2010 were obtained from the National Historical Geographical Information System (Supplementary Material 3.7.1 Table 1) [49].

3.2.3 Estimating vaccine coverage at school district level

We obtained DTaP immunization records for children 0-9 years old born between 1st January, 2008, and 31st December, 2017, living or born in King County, WA, from the Washington State Immunization Information System (WA-IIS). WA-IIS is a lifetime registry that tracks immunization records for people of all ages in Washington State [95]. Healthcare providers such as primary care physicians, hospitals, and healthcare plans voluntarily report patient immunizations to WA-IIS. Additionally, birth certificates of children born in King County are loaded into the registry periodically. The study cohort was restricted to children born in King County after 2008 to ensure data completeness and accuracy (as per the recommendations of Dr. Patricia DeHart at the WA Depart-

ment of Health). In recent years, WA-IIS has captured nearly all children born or living in King County through birth certificate data collected twice a month. Ninety-nine percent of children aged 4 months - 5 years have 2 or more immunizations recorded in the WA-IIS [95].

Using WA-IIS data, we created a retrospective cohort of 323,250 children 0 to 9 years of age. Vaccination information including vaccine name and date of receipt, vaccine manufacturer, and healthcare facility of administration was available for all pediatric vaccines recommended from birth through 9 years of age for each child. Demographic information included date of birth, sex, race, ethnicity, current residential address and county, and insurance information. Home addresses (or zip codes when home addresses were incorrect or not available) of WA-IIS participants were geocoded to their census tract of residence by WA Department of Health staff. We spatially overlaid the shapefile of geocoded home addresses of WA-IIS participants onto a shapefile of King County school districts to obtain school district of residence for each child. DTaP doses are recommended at ages 2, 4, and 6 months of age, with booster doses at ages 15-18 months and 4-6 years [68]. We used participants' date of birth and dates of DTaP vaccination to calculate age-appropriate vaccination status for each DTaP dose for each child. DTaP vaccine coverage at county and state level is reported routinely by the WA Department of Health as proportion of children 19-35 months old with ≥ 4 DTaP doses. For consistency, we estimated annual DTaP vaccine coverage at the school district level as proportion of 19-35 month old children living in a school district with ≥ 4 DTaP doses in a given year from 2011 to 2017 (the 2008 cohort turned 35 months old in 2011). We assumed that vaccine coverage in 2010 was the same as in 2011. We obtained school district level sociodemographic factors, namely, the proportion of population in each school district that are foreign born, Caucasian, non-citizens, speak a language other than English at home, proportion of population over 16 years old that have not completed high school, proportion of households with more than 4 people living in them, and median income, from the 2010 US Census data [96] .

3.2.4 Notation

Here we present some notation and assumptions common to the ecological vaccine model and epidemic-endemic model. Let Y_{it} and N_{it} be the number of cases and the total population in school-district i at time t. Let x_{it} be the time-varying vaccination coverage

estimated as proportion of 19-35 month old children vaccinated against pertussis with $\geq$ 4 DTaP doses. Let λ^*_{it} be the force of infection, i.e., risk of infection at time t for an individual who was susceptible at time $t-1$. For our analysis, we assume a time step of four weeks or a month which is the approximate generation time for pertussis. Assuming a constant hazard rate between time steps, the probability that a susceptible individual at time $t-1$ gets infected at time t is given by λ^*_{it}. Assuming that time until infection is independent for all susceptible individuals [13], the number of new cases in area i at time t can be modeled as:

$$Y_{it}|\lambda^*_{it} \sim Binomial(S_{i,t-1}, 1 - e^{-\lambda^*_{it}}) \tag{3.1}$$

Assuming λ^*_{it} is small, $1 - e^{-\lambda^*_{it}} \approx \lambda^*_{it}$. When the number of infections is small and the population is large, a Poisson distribution approximates the Binomial distribution. Thus, eq. (2.1) can be written as $Y_{it}|\lambda^*_{it} \sim Poisson(S_{i,t-1}\lambda^*_{it})$. Let $\mu_{it} = S_{i,t-1}\lambda^*_{it}$, then a general form of eq. (2.1) is,

$$Y_{it}|\mu_{it} \sim Poisson(\mu_{it}) \tag{3.2}$$

Both models assume that the number of infections is negligible compared to the number of susceptibles. Under this assumption, the number of susceptibles at time t, S_{it}, can be approximated by the initial number of susceptibles at time 0. Because this is a partially vaccinated population, the initial number of susceptibles is given by $(1 - x_{it})^{\beta_v} N$ for the epidemic-endemic model and $(1 - \phi x_{it})N$ for the ecological model. Thus, depletion of susceptibles is not explicitly modeled in either models.

3.2.5 The epidemic-endemic model

The epidemic-endemic model is motivated by the Poisson branching process with immigration. Total pertussis incidence μ_{it} is split into two components: the endemic component with rate ν_{it} and the epidemic component with rate $\lambda_{it}Y_{t-1}$, such that $\mu_{it} = \nu_{it} + \lambda_{it}Y_{t-1}$ [97]. The epidemic component can be further decomposed into the autoregressive and neighborhood components. In a model with spatial data, the autoregressive component models cases arising from infected individuals from the same area, while the neighborhood component captures cases arising from infected individuals in neighboring areas. The endemic component captures the remaining cases not explained by these two components. In this analysis, the endemic component models cases that were introduced from outside the county. Each component can be modeled

with a log linear model with covariates and fixed or random effects [93]. The model can be fit easily in the R `surveillance` package [98] and likelihood estimation is done using the quasi-Newton algorithm. When random effects are included, penalized and marginal log-likelihoods are maximized alternately until convergence [93].

The endemic component included the population as an offset modeled as the fraction of the population that live in school district i, $\left(\frac{N_i}{N}\right)$, denoted by e_i. To account for the temporal variance in incidence, the endemic component included an overall linear trend and a sinusoidal wave of frequency $\omega_t = \frac{2\pi t}{13}$. The endemic component is written as :

$$\log(\nu_{it}) = \alpha_{EN} + \beta_t(t) + \gamma \sin(\omega_t) + \delta \cos(\omega_t), \tag{3.3}$$

where ν_{it} is the endemic risk of pertussis and α_{EN} is the endemic intercept, assumed to be constant over the region. β_t is the parameter associated with temporal trend.

We included vaccination coverage in the epidemic component of the model as log proportion of children 19-35 months old that had fewer than 4 doses of DTaP vaccine because we are interested in the effect of pertussis vaccination on size and occurrence of pertussis epidemics. Effect of vaccination coverage on disease incidence has been modeled similarly for measles using the epidemic-endemic model before [99]. The epidemic component is written as:

$$\log(\lambda_{it}) = (\alpha_{AR} + \beta_v log(1 - x_{it})), \tag{3.4}$$

where λ_{it} is the epidemic pertussis risk, α_{AR} is the epidemic intercept, x_{it} is the vaccination coverage in school district i and at time t, and β_v is the parameter associated with vaccination coverage. The proportion of 19-35 month old children under-vaccinated or susceptible in the school-districts is $1 - x_{it}$. The intercept α_{AR} is assumed to be constant over all areas. In this simple epidemic-endemic model, we do not include a neighborhood component to be able to compare the α_{AR} and α_{EN} estimates with estimates from the ecological vaccine model. Overall, this model can be written as:

$$\mu_{it} = e_i \nu_{it} + \lambda_{it} Y_{i,t-1} \tag{3.5}$$

where μ_{it} is the total pertussis risk, e_i is the population fraction in school district i used as a multiplicative offset in the endemic component, and $Y_{i,t-1}$ is the observed number of cases in school district i at time $t - 1$.

We fit a separate epidemic-endemic model with socio-demographic covariates in the endemic component and the epidemic component split into autoregressive and neighborhood components to measure spatio-temporal dependence. School districts that shared a boundary were defined as neighbors. Weights w_{ji} represent the flow of cases from school district j to school district i when $j \neq i$. The model assumes that the epidemic can only arrive from adjacent areas and the neighborhood risk was scaled with population fraction e_i to reflect that areas with larger populations are more likely to import cases from neighborhoods. The neighborhood component can be written as

$$\zeta_{it} = \alpha_{NE} + \zeta_{it} e_i^{\beta_{Pop}}, \tag{3.6}$$

where ζ_{it} is the neighborhood associated risk of pertussis, α_{NE} is the intercept associated with the neighborhood component, and β_{Pop} is the parameter associated with population fraction.

Let β_z is a vector of parameters associated with sociodemographic covariates namely the proportion of population in each school district that are foreign born, Caucasian, non-citizens, speak a language other than English at home, proportion of population over 16 years old that have not completed high school, proportion of households with more than 4 people living in them, and median income. Combining equations (3.3), (3.4), and (3.6), we get:

$$\begin{aligned}
\log(\nu_{it}) &= \alpha_{EN} + \beta_t(t) + \gamma \sin(\omega_t) + \delta \cos(\omega_t) + \beta_z(z_i) \\
\log(\lambda_{it}) &= \alpha_{AR} + \beta_v log(1 - x_{it}) \\
\mu_{it} &= e_i \nu_{it} + \lambda_{it} Y_{i,t-1} + \zeta_{it} e_i^{\beta_{Pop}} \sum_{j \neq i} w_{ji} Y_{j,t-1}
\end{aligned} \tag{3.7}$$

where $Y_{j,t-1}$ is the number of cases in school-district j at time $t-1$.

3.2.6 The ecological vaccine model

The ecological vaccine model was developed with inference as the primary goal, and its detailed derivation can be found elsewhere [92]. In short, the ecological vaccine model was developed by starting with individual-level vaccine models and aggregating them to area-level to reduce ecological bias. The authors found under certain assumptions that individual-level vaccine models for both the all-or-none (or primary vaccine failure) and leaky modes of vaccine action can be aggregated to give the same ecological vaccine

model. The impact of vaccination in this model is defined as the ability of the vaccine to reduce susceptibility against infection. This ecological vaccine model is easily fit using the Bayesian framework in R package `rstan` [100]and provides estimates of epidemiologically relevant parameters [92].

In a partially vaccinated population, let ϕ be the reduction in a vaccine recipient's risk of infection, which can be interpreted as the vaccine effectiveness of DTaP vaccine [92]. Given that x_{it} is the vaccine coverage in school district i at time t, the number of susceptibles in school district i can be written as $(1 - \phi x_{it})N_i$. Using the Bayesian framework, we can incorporate prior knowledge about pertussis vaccine effectiveness into the ecological vaccine model. Randomized controlled trials have found the vaccine efficacy of the DTaP vaccine to be $\sim 85\%$ (95% CI: 81%, 87%) [83]. We fit the ecological vaccine model as:

$$Y_{i,t+1}|\mu_{it}, \phi \sim Poisson\left(N_i(1 - \phi x_{it})\left(\lambda_i \frac{Y_{it}}{N_i} + \nu_{it}\right)\right), \tag{3.8}$$

$\log \lambda_i = \alpha_{AR} + a_i$
$\log \nu_{it} = \alpha_{EN} + b_i + \gamma \sin(\omega_t) + \delta \cos(\omega_t) - \log(N_i)$
$a_i \sim N(0, \sigma^2_{AR})$
$b_i \sim N(0, \sigma^2_{EN})$
$\phi \sim \text{Beta}(c, d)$

where μ_{it} is the total risk (epidemic plus endemic risk) and ν_{it} and λ_i are the endemic and epidemic pertussis risk components. The school district specific random effects a_i and b_i are assumed to be independent; $\omega_t = \frac{2\pi t}{13}$; a strong beta prior Beta(10, 2.5) was used for ϕ which places 90% of the mass between 0.66 and 0.99. We assumed normal priors with mean 0 and variance 5 for α_{AR} and α_{EN} and variance 10 for γ and δ. We assumed frequency-dependent transmission in the formulation of λ_i [101]. Hamiltonian Monte Carlo sampling via R package `rstan` was used to fit this model [100]. We adapted code published previously for an ecological vaccine model for measles data to include time-varying vaccination coverage [102]. As a sensitivity analysis, we fit the same model with non-informative priors on ϕ, where 90% of the mass is between 0.05 and 0.95, to check the influence of priors on the estimate of vaccine effectiveness.

School-district-specific, time-varying autoregressive parameters were calculated as:

$$(1 - \hat{\phi} x_{it}) \exp(\widehat{\alpha}_{AR} + \hat{a}_i) \tag{3.9}$$

Fitted values were calculated as:

$$\widehat{Y}_{it} = (1 - \hat{\phi} x_{it}) \left[\exp(\widehat{\alpha}_{AR} + \widehat{a}_i) \widehat{Y}_{i,t-1} + \left(\frac{N_i}{N} \right) \exp\left(\widehat{\alpha}_{EN} + \hat{b}_i + \hat{\gamma}(\omega_t) + \hat{\delta}(\omega_t) \right) \right], \tag{3.10}$$

where $Y_{i,t-1}$ was the observed number of cases in school i and month $t - 1$.

3.3 Results

3.3.1 Descriptive Analysis

Between 2010 and 2017, 1,881 pertussis cases were reported in the 18 school districts in King County, WA. There was a large epidemic in 2012 and a smaller one in 2015. The largest number of cases in a single month was 50 during the 2012 epidemic and occurred in Seattle school district, which also recorded the highest number of cases overall. Pertussis incidence per 100,000 by school district is in Fig 3.1a.

Data from WA-IIS was used to estimate vaccine coverage at the school-district level which is displayed in Fig 3.1b. WA-IIS captures almost all children born or living in King County and has fairly complete information on nearly all pediatric immunizations administered. Even though we used 19-35 month olds to estimate vaccination coverage with ≥ 4 DTaP doses, we assumed that the vaccine coverage for the entire population of King County is the same as the coverage estimated for this analysis. A similar assumption was made in other studies on effect of measles vaccination coverage on measles incidence in Germany [99] [92]. Vaccination coverage is higher in school districts in northern and north-eastern King County, compared to school districts in the south. Vashon Island (in black) has the lowest vaccine coverage of all school districts. Within each school district, vaccine coverage appeared to increase with time between 2010 and 2017 (Table 3.1).

Pearson correlation coefficients showing correlations between vaccine coverage and disease incidence in each year are in Table 3.2. With the exception of 2013 and 2015, vaccine coverage with ≥ 4 doses of DTaP was negatively correlated with pertussis incidence.

This association was strong and statistically significant in the years 2010 (R: -0.74, 95% CI: $-0.89, -0.42$), 2012 (R: -0.72, 95% CI: $-0.88, -0.37$) and 2016 (R : -0.54, 95% CI:$-0.80, -0.09$). Thus, there is some indication that areas with higher vaccine coverage showed lower disease incidence.

3.3.2 Ecological vaccine model

Summaries of posterior medians of fixed effects from the ecological vaccine model are in Table 3.3. Using a strong prior for the vaccine effectiveness ϕ, we estimated it to be 83% (95% CI: 63%, 95%). Thus, the vaccine effectiveness associated with receiving ≥ 4 doses of DTaP compared to receiving < 4 DTaP doses is statistically significant. Our results agree with estimates of efficacy of DTaP vaccine found in the literature ($\sim 85\%$) [83]. As a sensitivity analysis, we ran the same model using a uniform prior for ϕ and vaccine effect was estimated to be 79% (95% CI: 33%, 96%) (Supplementary Material 3.7.1 Table 2). The uniform prior resulted in slightly smaller estimates for ϕ, but confidence intervals were wider. This vaccine effectiveness estimate does not differentiate between primary vaccine failure and failure due to leakiness. With a uniform prior, the estimate of epidemic intercept α_{AR} in this model was also smaller with wider confidence intervals. Using a uniform prior did not have much of an effect on the endemic intercept α_{EN}.

We computed school-district specific time-varying autoregressive components and their 95% credible intervals from the ecological vaccine model as $(1 - \hat{\phi} x_{it}) \exp(\hat{\alpha}_{AR} + \hat{\alpha}_i)$. This parameter is equivalent to the effective reproductive number, R_{eff}, which is defined as the average number of new cases per infectious case in a partially vaccinated population [92] . All estimates are below 1 and vary slightly with time within school-districts. R_{eff} in school districts with fewer cases have wider credible intervals (Fig 3.2).

Fig 3.3 shows the observed number of cases, incidence per 100,000 people, and model fits for each school district obtained from the ecological vaccine model. District-specific estimates of R_{eff} and endemic risk ν_{it} are given in each panel. The model seems to fit the data well especially for areas with a large number of cases such as Federal Way, Kent, Lake Washington, and Seattle. The correlation coefficients between average vaccine coverage (measured as total children 19-35 months old with ≥ 4 doses of DTaP between 2010 and 2017 divided by total children 19-35 months old for each school district) and R_{eff} ($r = -0.05$; 95% CI: $-0.50, 0.42$) and average vaccine coverage and endemic risk

ν_{it} ($r = 0.18$; 95% CI: $-0.31, 0.59$) are small and not statistically significant. Population density (measured as persons per square mile) is positively correlated with both R_{eff} ($r = 0.24$; 95% CI: $-0.25, 0.63$) and endemic risk ($r = 0.58$; 95% CI: $0.16, 0.82$). Thus, R_{eff} in this study was not statistically significantly correlated with either vaccine coverage or population density. Endemic pertussis risk, however, was signficantly correlated with population density.

Supplementary Material 3.7.2 Fig 2(a) and 2(b) show area-specific random effects of the autoregressive and endemic components. Areas with a large number of cases have larger autoregressive random effects. However, no such structure was found in the endemic random effects. Endemic and autoregressive random effects were not significantly correlated (Supplementary Material 3.7.2 Fig. 3) which supports our decision to use independent random effects in our model.

3.3.3 Epidemic-endemic models

We will first discuss the results of the simple model with vaccination coverage in the epidemic component as the only covariate and no neighborhood component or other demographic covariates included. The estimate of the exponentiated autoregressive intercept from this model ($\exp(\alpha_{AR})$=0.07; 95% CI: 0.003, 1.43) is lower compared to autoregressive intercept estimate from the ecological vaccine model ($\exp(\alpha_{AR})$=1.1; 95% CI: 0.75, 1.52) and has very wide confidence intervals (Table 3.4). The endemic estimate from the epidemic-endemic model (2.23; 95% CI: 1.99, 2.47) is lower than that from the ecological vaccine model (3.16; 95% CI: 2.77, 3.50). There is no comparable estimate to vaccine effect ϕ in this model. Here, the effect of vaccine coverage is estimated as $2^{\exp(\beta_v)}$. It is interpreted as for each doubling of pertussis under-vaccination rate (or doubling of susceptible population), the epidemic risk of pertussis increases multiplicatively by 3.54 fold (95% CI: 1.65, 23.05). Thus, this model suggests that epidemic pertussis risk was statistically significantly associated with DTaP vaccination coverage. There is no strong temporal or seasonal trend in the data. A statistically significant overdispersion parameter suggests that using the negative binomial distribution was an appropriate choice for this model.

We fit a second epidemic-endemic model where we split the epidemic component into the neighborhood and autoregressive components and added sociodemographic factors

in the endemic component to estimate the effect of sociodemographic factors on endemic pertussis risk. For this model, for each doubling of under-vaccination rate (or doubling of susceptible population), the epidemic risk of pertussis increased multiplicatively by 2.76 fold higher (95% CI: 1.44, 16.6) (Table 3.4). This result is consistent with the under-vaccination coverage estimate from the simple model. Household size and median income were statistically significantly associated with endemic pertussis risk. For every unit increase in proportion of households with more than 4 individuals, the endemic risk of pertussis increased by 15% (95% CI: 8.3%, 20.9%), adjusting for other covariates. For every $10,000 increase in median income, endemic pertussis risk decreased by 45% (95% CI: 25%, 60%).

The epidemic-endemic model fits the data well (Fig 3.4). According to this model, 32.8% of the time-averaged mean pertussis risk is explained by the endemic component, 34% by the autoregressive component, and 33% by the neighborhood component. The model suggests that a large proportion of cases in Renton, Lake Washington, and Bellevue school districts come from neighboring areas. This might be because these school districts have many neighbors and are also highly populated. On the other hand, Vashon Island has most of its incidence explained by the autoregressive component (Supplementary Material 3.7.2 Fig. 1).

3.4 Discussion

Local variations in DTaP vaccination coverage could lead to areas of high susceptibility to pertussis which could be prone to occasional large outbreaks. Previous studies that examined the association between area-level vaccination coverage and pertussis incidence did not account for the dependency of outcomes in models of infectious diseases. In this analysis, we explored the relationship between local area-level vaccination coverage and pertussis epidemicity and endemicity using the ecological vaccine model and epidemic-endemic model approaches with monthly pertussis data at school-district level in King County, WA. We obtained estimates of epidemiologically interesting parameters such as vaccine effectiveness and effective reproductive number from the ecological vaccine model. While the vaccine effect was statistically significant and commensurate with what is known in the literature, the effective reproductive number was not statistically significantly correlated with vaccine coverage. From the epidemic-endemic model, we found that under-vaccination was significantly associated with pertussis epidemics and

household-size and median income were associated with endemic pertussis risk.

The strengths of the epidemic-endemic models over an ecological Poisson regression model for infectious disease data are that the epidemic-endemic models can deal with multivariate time-series data and cope with the occasional large outbreak, in addition to incorporating covariates [99]. Moreover, they are easily fit and allow modeling of neighborhood effects to study spatio-temporal dependence. Epidemic-endemic models with various complexities have been used to study measles [99], meningococcal disease [103] [104], psychiatric hospital admissions [103], and Norovirus infections [105]. However, the epidemic-endemic models were developed with disease prediction as their main goal and are less suitable for inference because they do not address ecological bias. It has been shown before that for estimating effects of vaccination on infectious disease incidence, epidemic-endemic models can give biased estimates for the epidemic and endemic risks [92]. Thus, effects of covariates obtained from these models should be interpreted carefully, explicitly stating ecological bias as a drawback.

In the epidemic-endemic models, vaccination coverage (and other covariates) can be included in either the epidemic or endemic components. In earlier studies, this decision was driven by which model formulation fit the data best and not with inference as the primary goal. In eq. (3.4), given how vaccination coverage is included in the form of $\log(1 - x_{it})$, the parameter associated with vaccination coverage is the flexibility parameter that improves model fit [92]. The interpretation of this parameter as expected multiplicative change in disease risk for every 2 fold increase in undervaccination or susceptibility is difficult to interpret and non-intuitive.

The ecological vaccine model also appropriately models aggregated infectious disease data. In addition, it gives less biased and easily interpretable estimates of epidemic and endemic risks under certain assumptions. One might consider using ecological vaccine models over epidemic-endemic models when inference is the goal of the analysis. The estimates of R_{eff} from our model are lower than 1. Considering there was only one large and one small outbreak of pertussis in the 8 year study period, a low estimate of the autoregressive component $\hat{\alpha}_{AR}$ averaged over the entire period for an endemic disease is not surprising. Another explanation for low R_{eff} estimates in this study might be that pertussis is severely underreported, with $\sim$ 10% of cases reported [106] [8], and reporting could vary by age and school district. Our incidence data was discretized in months while actual transmission is a continuous process. Discretization of time can re-

sult in biased estimates of reproductive numbers [107]. Some limitations of the ecological vaccine model are that the model needs to be developed further include neighborhood structure and variable infectious periods. For example, in this model we assumed that cases occurring in a given month depend on cases in the previous month. However, the transmission process is continuous and it might be more realistic to have a model with weekly cases where cases occurring in week t can explain cases occurring in week $(t + l)$ where $l \in (1, 2, 3, 4)$ [92].

In summary, we analyzed pertussis surveillance data marked in space and time using the ecological vaccine model and the epidemic-endemic model to estimate effects of pertussis vaccination on pertussis incidence. We found that area-level vaccination coverage is statistically significantly associated with pertussis epidemics. We estimated the vaccine effect of receiving $\geq$4 DTaP doses to be 83% (95% CI: 63%, 95%) using the ecological vaccine model, but this model does not give the mechanism of vaccine failure. Effects of association between pertussis vaccination coverage and pertussis incidence from epidemic-endemic model do not address ecological bias and may not apply at the individual-level. It is possible to get less biased estimates of vaccine effectiveness using the ecological vaccine model when area-level vaccine coverage and disease incidence are known.

3.5 Tables

Table 3.1: Pertussis cases and estimated vaccine coverage in King County, Washington, between 2010 and 2017

School District	**Pop**[†]	**Pertussis Cases**		**Estimated Vaccine Coverage** [±]						
		Max cases [*]	**Sum cases** [°]	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
Enumclaw	25179	6	156	51.70	54.31	58.89	55.63	58.05	69.16	70.62
Auburn	84053	16	990	67.62	68.38	69.07	66.33	66.77	70.82	68.70
Bellevue	124003	6	426	66.43	68.22	67.91	65.33	69.13	80.05	75.79
Federal Way	130706	20	1290	63.99	67.17	67.62	67.03	64.98	73.41	70.14
Highline	124481	13	468	70.91	69.04	73.23	72.67	73.44	80.52	76.72
Issaquah	98660	9	438	76.43	77.14	76.60	74.22	76.56	82.58	79.85
Kent	158233	7	1170	65.83	67.38	67.50	69.22	70.72	77.77	73.13
Lake Washington	177476	13	912	69.47	71.12	70.94	71.00	74.09	82.54	77.28
Riverview	19315	5	216	74.94	73.89	70.00	71.53	69.15	79.42	75.35
Mercer Island	22699	2	90	76.81	76.73	78.90	73.85	80.36	84.93	79.29
Northshore	122684	14	606	72.80	74.62	75.62	72.27	70.16	80.08	74.02
Renton	115511	9	438	70.93	71.71	73.19	72.14	72.03	77.37	74.63
Seattle	609471	50	2886	71.13	73.50	74.44	73.68	74.80	83.76	81.80
Shoreline	65547	13	432	73.21	75.88	75.78	74.63	72.00	86.06	85.08
Snoqualmie Valley	35054	2	132	76.61	77.29	75.14	76.16	80.11	86.01	82.12
Tukwila	18038	1	18	64.27	66.09	67.75	68.41	69.52	75.20	70.13
Tahoma	37376	21	498	70.15	75.17	73.00	70.17	69.82	80.98	78.74
Vashon Island	10624	10	144	50.98	48.32	62.60	45.08	49.09	65.26	55.13

[*] Max Cases are the maximum number of cases in any given month

[°] Sum Cases are the total number of cases between 2010 and 2011

[±] Vaccination coverage of 2010 was assumed to be the same as 2011

[†] Total Population in Census year 2010

Table 3.2: Estimated Pearson's Correlation coefficients and 95% confidence intervals (CI) between pertussis incidence and vaccine coverage by year

Year	**Pearson's correlation coefficient, r**	**95% CI**	**P-value**
2010	−0.74	−0.89, −0.42	< 0.001
2011	−0.27	−0.65, 0.22	0.2
2012	−0.72	−0.88, −0.37	< 0.001
2013	0.13	−0.35, 0.56	0.6
2014	−0.063	−0, 51, 0.41	0.8
2015	0.08	−0.39, 0.53	0.7
2016	−0.54	−0.80, −0.09	0.02
2017	−0.19	−0.60, 0.30	0.4

Table 3.3: Posterior median estimates and 95% credible intervals (CI) from ecological vaccine model with a strong prior on vaccine effect ϕ, $\phi \sim Beta(10, 2.5)$

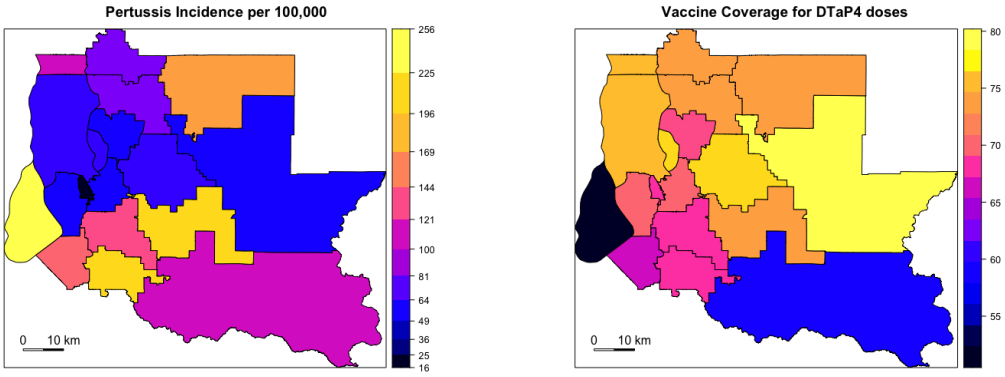
Parameter	**Posterior medians**	**95% CI**
α_{AR}	0.10	−0.28, 0.42
ϕ	0.83	0.63, 0.95
α_{EN}	3.16	2.77, 3.50
γ	−0.02	−0.13, 0.09
δ	−0.09	−0.21, 0.02
σ_{AR}	0.38	0.22, 0.66
σ_{EN}	0.45	0.30, 0.71
$\exp(\alpha_{AR})$	1.10	0.75, 1.52

Table 3.4: Results from Epidemic-Endemic models for effect of vaccination coverage and other sociodemographic factors on pertussis risk

Parameter	**Epidemic Component**		**Endemic Component**		**log (L)**	**p**	**AIC**
	β Coefficient	**95% CI**	**β Coefficient**	**95% CI**			
Vaccination coverage in epidemic component, no neighborhood component, no other covariates					-1939.89	7	3893.78
α_{AR}	-2.70	-5.78, 0.36					
log(Under-vaccination)	0.59	-0.32, 1.51					
α_{EN}			2.23	1.99, 2.47			
Temporal trend			0.002	-0.001, 0.006			
γ			-0.04	-0.20, 0.11			
δ			-0.02	-0.18,0.13			
Vaccination coverage in epidemic component,simple neighborhood structure, other covariates					-1859.43	16	3750.85
α_{AR}	-2.42	-5.8, 0.98					
log(Under-vaccination)	0.38	-0.63, 1.4					
α_{NE}	-0.74	-1.62, 0.14					
log(population fraction)	0.72	0.47, 0.97					
α_{EN}			7.69	-2.29, 17.66			
Temporal trend			0.005	0.0009, 0.01			
γ			-0.09	-0.31, 0.12			
δ			-0.09	-0.31, 0.11			
% Foreign born			-0.09	-0.34, 0.13			
% Caucasians			-0.03	-0.13, 0.007			
% Not Citizens			0.08	-0.15, 0.31			
% with Household size ≥ 4			0.14	0.08, 0.19			
% language other than English at home			-0.03	-0.38, 0.31			
% education less than high school			-0.2	-0.40, 0.009			
Median income (for every $10,000)			-0.61	-0.93, -0.29			

log(L): Log likelihood; p: number of parameters ; Akaike's Information Criterion is calculated as $(AIC) = 2log(L)+2p$; lower AIC values are preferred.

3.6 Figures



(a) (b)

Figure 3.1: Pertussis incidence per 100,000 and DTaP vaccine coverage for ≥ 4 doses in King County, Washington.

(a) Total pertussis incidence between 2010 and 2017 in plotted at school district level. Incidence is high in Vashon Island and school districts in southern King County; (b) Vaccine coverage averaged over 8 years from 2011 to 2017 is plotted by school district. Visually, it appears that school districts with low vaccine coverage have high pertussis incidence.

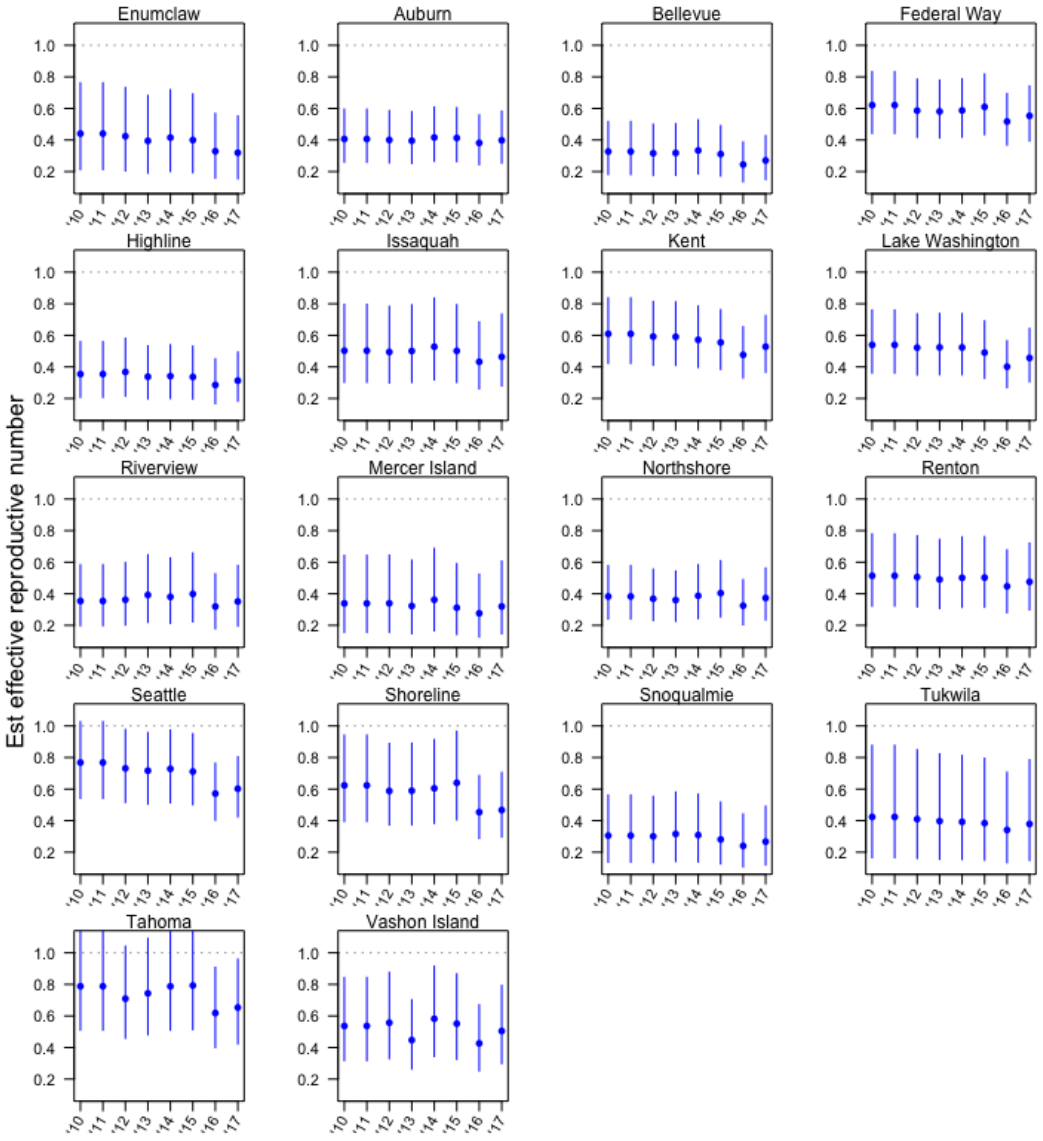


Figure 3.2: Estimated school-district-level autoregressive components (R_{eff}) and associated 95% Credible Intervals estimated using ecological vaccine model

R_{eff}, computed as $(1-\phi)x_{it}exp(\alpha_{AR}+\alpha_i)$ varies with time within each school district, reflecting changes in vaccine coverage over time within school districts. All estimates are below 1. Credible intervals for R_{eff} are larger for areas with fewer cases. Some school districts with high vaccine coverage have low estimates of R_{eff} (Bellevue, Mercer Island), but there is no clear effect of vaccine coverage on the autoregressive component. On the other hand, school districts with large population sizes had higher R_{eff} estimates(Seattle, Kent, Lake Washington, Federal Way).

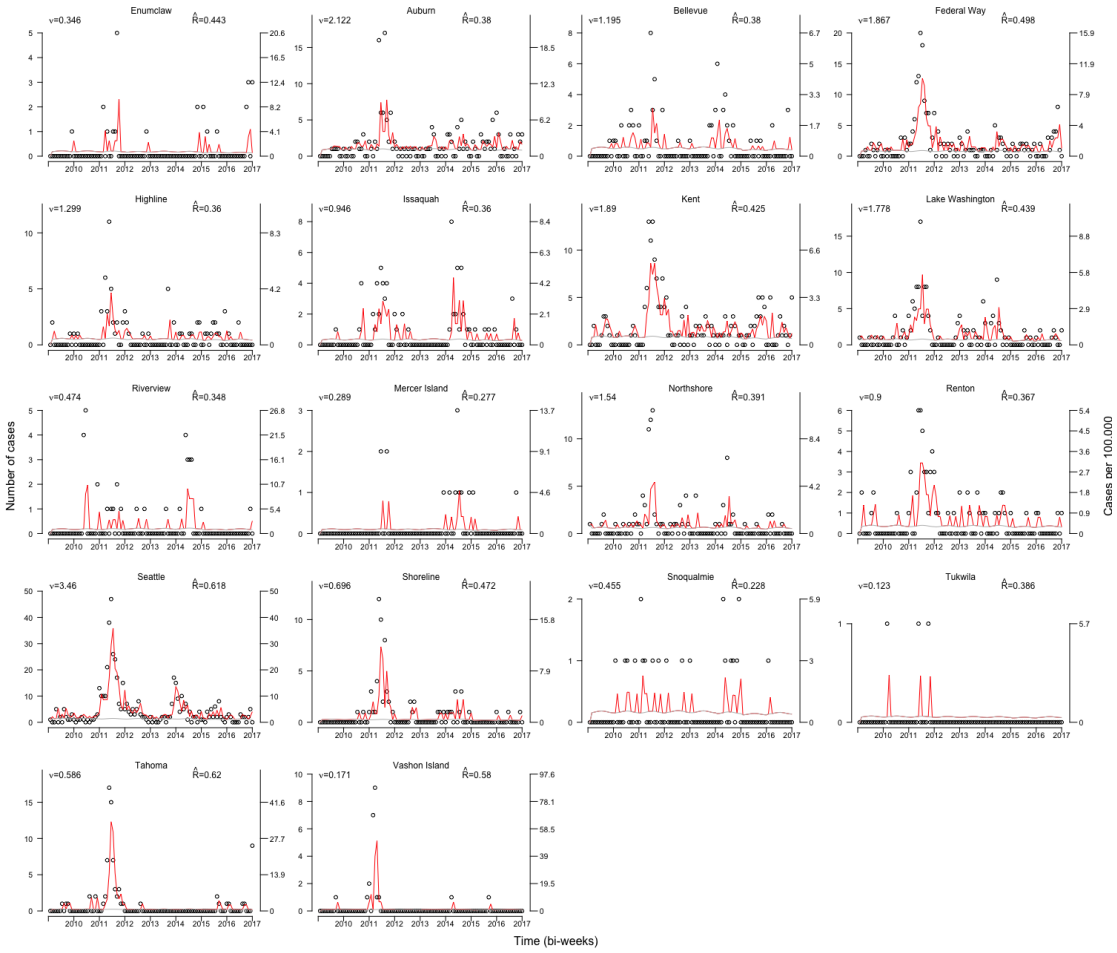


Figure 3.3: Ecological vaccine model fitted values by school district.

No. of pertussis cases (left axis) and incidence (right axis) by school district. Red lines are fitted epidemic component, grey lines show the endemic component, and black circles are absolute number of cases. The model fits well, especially in areas with high number of cases.

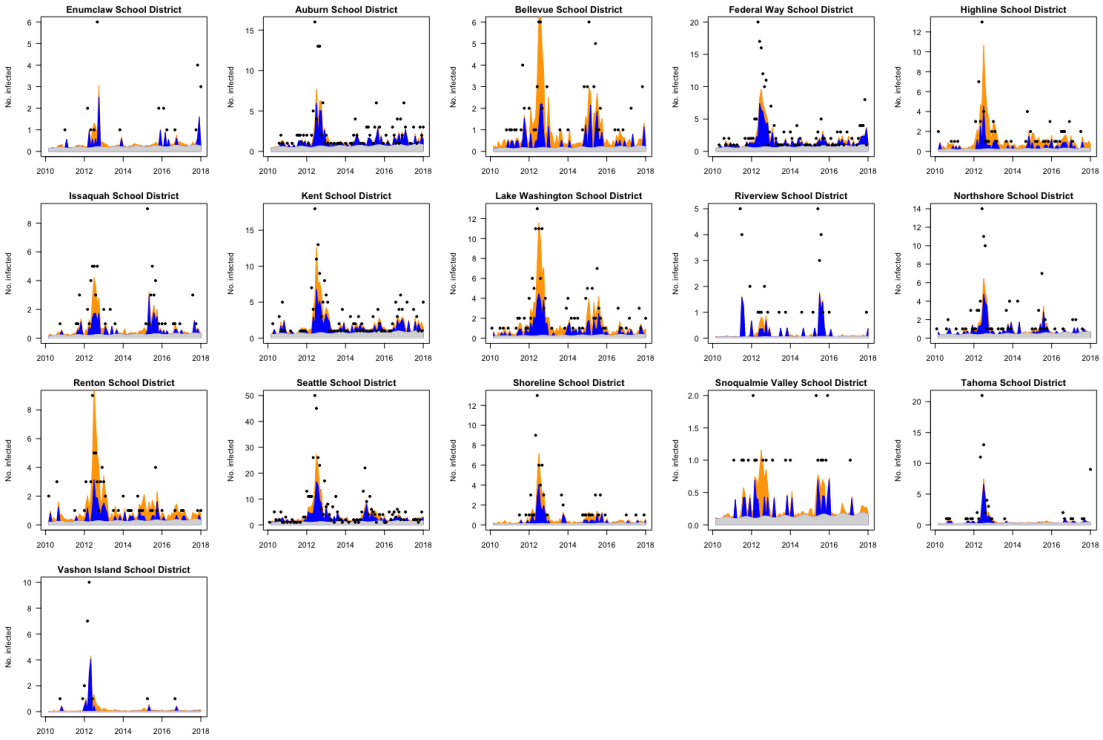


Figure 3.4: Epidemic-endemic model fitted values by school district.

Endemic component is denoted in grey, autoregressive in dark blue and neighborhood component in yellow. Observed number of cases are displayed as black dots except for time points with zero cases. The model fits the data reasonably well, especially for school disitrcts with higher number of cases

3.7 Supplementary Material

3.7.1 Supplementary Tables

Table 1: Demographic characteristics of King County school districts

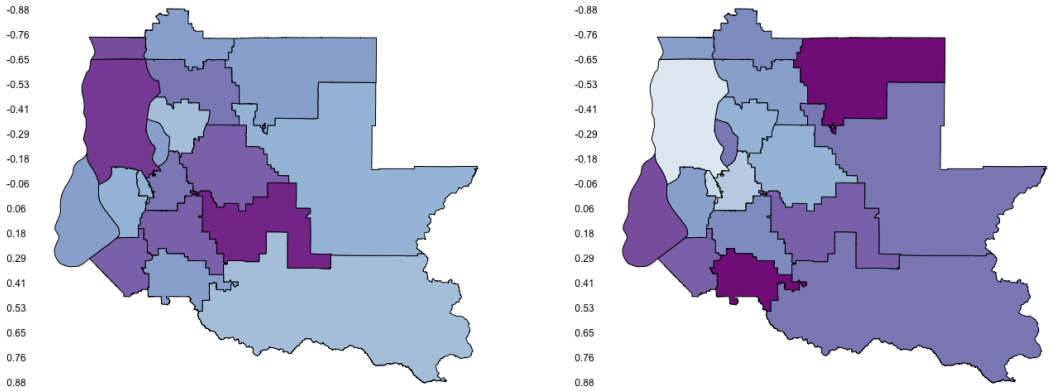
School district	**% Caucasian**	**% Foreign Born**	**% Non-citizens**	**% Households with ≥ 4 people**	**%Median Income**	**% with less than High School education**	**% Non-English speakers**
Enumclaw	91.7	3.3	2.0	23.2	68900	5.8	1.56
Auburn	70.9	17.7	10.9	27.2	55442	11.0	11.1
Bellevue	63.6	30.4	18.1	20.9	81656	3.7	14.4
Federal Way	60.4	21.3	11.5	27.3	61157	10.6	13.7
Highline	56.7	23.7	13.3	24.4	50946	16.2	17.3
Issaquah	73.9	19.6	10.6	28.3	101139	2.1	8.1
Kent	61.8	20.8	11.3	28.7	65808	10.7	14.4
Lake Washington	74.5	21.5	13.9	22.4	93294	3.2	7.7
Riverview	90.1	7.7	4.5	29.3	98248	5.1	4.1
Mercer Island	77.8	16.9	6.5	23.4	120994	0.7	4.2
Northshore	80.0	13.6	7.1	25.7	87473	3.5	6.0
Renton	53.2	24.9	12.3	23.1	62655	10.3	16.7
Seattle	69.4	16.9	8.1	13.2	60645	6.0	9.9
Shoreline	73.6	18.0	7.0	20.3	71020	5.8	9.2
Snoqualmie Valley	88.9	7.4	4.3	30.0	100038	3.9	3.2
Tukwila	43.4	35.5	23.7	25.8	41675	23.0	26.7
Tahoma	87.7	5.9	3.0	31.5	92596	4.3	3.2
Vashon Island	92.3	3.4	1.7	17.0	80000	2.6	0.7

Table 2: Posterior median estimates and 95% credible intervals (CI) from ecological vaccine model with a uniform prior on vaccine effect ϕ, $\phi \sim Beta(1, 1)$

Parameter	**Posterior medians**	**95% CI**	
α_{AR}	0.03	−0.56,	0.43
ϕ	0.79	0.33,	0.96
α_{EN}	3.09	2.47,	3.51
γ	−0.02	−0.13,	0.09
δ	−0.09	−0.21,	0.02
σ_{AR}	0.37	0.21,	0.66
σ_{EN}	0.46	0.30,	0.73
$\exp(\alpha_{AR})$	1.03	0.57,	1.54

3.7.2 Supplementary Figures

Fig 1: Map of school districts in King County, Washington



(a) (b)

Fig 2: Maps of random effects of autoregressive (a) and endemic components (b) from ecological vaccine model

There appears to be some spatial structure in autoregressive random effects but not in endemic random effects.

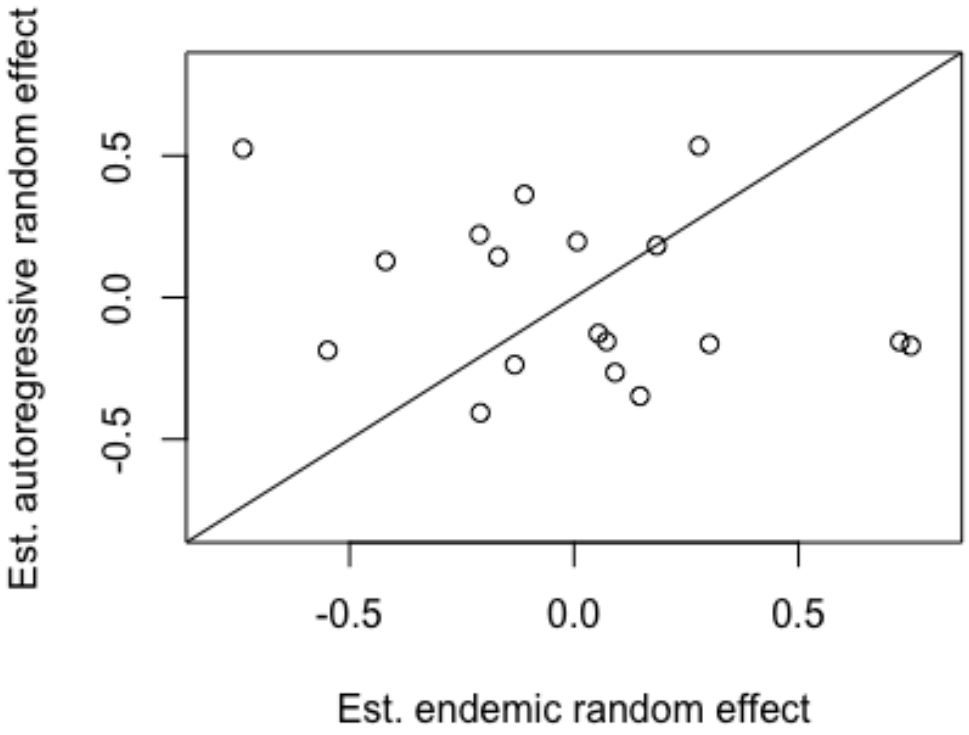


Fig 3: Correlation between random effects from ecological vaccine model

Chapter 4

Role of the Diphtheria-Tetanus-acellular Pertussis vaccine timing and number of doses on age-specific pertussis risk in infants and young children in King County, Washington

Abstract

Background: In most countries, the Diphtheria-Tetanus-acellular Pertussis (DTaP) vaccine is administered as a 3-dose infant series followed by additional booster doses in the first five years of life. Short-term immunity due to the DTaP vaccine can depend on the number, timing, and interval between doses. Not receiving doses in a timely manner might increase risk of pertussis.

Methods: Pertussis cases aged 0 to 9 years reported to the Public Health Seattle and King County between 1st January, 2008 and 31st December, 2017 were linked with immunization records for all children born and living in King County, Washington, during

the same period and registered in the Washington State Immunization Information System. Children who received fewer than the recommended doses at a given age were defined as "under-vaccinated". Children who received the correct number of doses at a given age but received them outside the recommended time frame were defined as "delayed". Association between delayed vaccination and under-vaccination and age-specific pertussis risk was estimated using log binomial models. We also estimated the association between delayed infant series initiation and pertussis risk in the first year of life. Additionally, we compared the effect of accelerated (3 doses at ages 2, 4, 6 months) and long (3 doses at ages 2, 4, and 12 months) vaccine schedules on pertussis risk in the first two years of life.

Results: Delays in vaccination and under-vaccination for the childhood DTaP series and boosters increased with age, but improved for successive birth cohorts. Vaccine delay among children who eventually received all doses was not longer than 5 weeks. Under-vaccination was significantly associated with increased risk of pertussis for the 3-dose primary series (adjusted Relative Risk (aRR): 4.8, 95% confidence interval (CI): 3.1, 7.6), the first booster (aRR: 3.2, 95% CI: 2.3, 4.5), and the second booster (aRR: 4.6, 95% CI: 2.6, 8.2). However, delay in vaccination among children who received the recommended number of vaccine doses was not statistically significantly associated with pertussis risk. Delay in series initiation among those who receieved 3 doses of vaccines by 7 months of age was not associated with increased pertussis risk in first year of life. There was no significant difference in pertussis risk between children who received accelerated vs. long vaccine schedules.

Conclusion: Under-vaccination is associated with increase in pertussis risk but short delays in vaccine receipt are less important as long as age-appropriate number of doses are administered. Even if there is a delay in series initiation, receiving the primary series by 7 months of age can protect against pertussis in the first year of life when infants are most at risk for serious disease.

4.1 Introduction

4.1.1 Pertussis vaccines in the US

Widespread rollout of the Diphtheria-Tetanus- whole-cell-Pertussis (DTwP) vaccines in 1940s resulted in a dramatic decrease in pediatric pertussis incidence until the 1980s

in the United States [4]. Due to concerns surrounding the safety and reactogenecity of DTwP vaccines, which led to decreased acceptance of the vaccine in many countries [108] [109] [110], efforts were made to develop a less reactogenic Diphtheria-Tetanus-acellular Pertussis (DTaP) vaccine, which was licensed in the US in 1996 [4] [5]. DTaP vaccines are safe and highly efficacious, and today, most developed countries recommend them for use in the infant primary series as part of their national immunization program [31]. Pre-school and adolescent booster doses are also included in immunization schedules, especially in high income countries, in response to the increase in age of infection and concerns about waning of immunity following vaccination with DTaP vaccines [111]. Despite high vaccination coverage nationally for both the primary series and boosters, pertussis incidence has been on the rise [11].

4.1.2 Acellular pertussis vaccine timeliness and effectiveness against pertussis disease

Short-term protection afforded by pertussis vaccines to age-appropriately immunized children depends on the number, timing, and interval between doses [112]. Because evidence suggests that DTaP immunity may wane over time [113] [114] [115], longer intervals between doses due to delays or missed immunizations could possibly increase the risk of pertussis infection even in partially vaccinated children. This could lead to sustained transmission of pertussis and periodic outbreaks [116] [117]. Observational studies in the US and Taiwan have suggested that under-vaccination or delay in vaccination results in an increase in risk of pertussis [118] [119] [120]. Current methods of estimating DTaP vaccination coverage at specific ages without estimating timeliness of each dose can mask delays in vaccination while showing high vaccination coverage at national levels [116] [117] [121].

4.1.3 Different vaccine schedules

Even though most countries administer the DTaP vaccine in the first year of life as a primary series, the number and timing of vaccine doses varies between countries. US and Australia administer the DTaP vaccine as a primary series at ages 2, 4, and 6 months (also called “3p” or “accelerated” schedule), followed by a second-year booster at 18 months and another pre-school booster at 4-6 years of age [31]. UK and Germany recommend a three dose schedule with even shorter intervals at ages 2, 3, and 4 months, followed by booster doses later [31]. Scandinavian countries such as Denmark

and Finland, on the other hand, follow a relaxed schedule for the primary series, administered at ages 3, 5, and 12 months (also called "2p +1" or "long" schedule), followed by a booster at 5 years. Countries such as Brazil and Thailand use whole-cell formulations of the vaccine (DTwP) in their primary schedules [31]. It is difficult to directly compare effectiveness of these various schedules in protecting against pertussis due to between-country differences in local pertussis epidemiology, case detection methods, case reporting fidelity, vaccine formulation used, and vaccine coverage achieved [122].

Our scientific goal is to examine the association between number and timeliness of vaccine doses and age-specific pertussis risk in this cohort. We first described the trends in timeliness of DTaP vaccination in a cohort of children born in King County, Washington, between 2008 and 2017 and registered in the Washington State Immunization Information System (WA-IIS). We then used log binomial models to compare pertussis risk between children who were age-appropriately vaccinated and those who were under-vaccinated children or recieved doses with a delay. Additionally, we compared pertussis risk between children who received their primary series at ages 2, 4, and 6 months ("3p schedule") vs. 3, 5, and 11/12 months ("2p +1" schedule).

4.2 Methods

4.2.1 Pertussis cases

Pertussis is a nationally notifiable disease. Pertussis cases aged 0-9 years reported between 1st January, 2008 and 31st December, 2017 were obtained from the Public Health Seattle and King County (PHSKC) surveillance database. The clinical case definition of pertussis used is a cough illness lasting ≥ 2 weeks with at least one of the following: paroxysms of coughing or inspiratory "whoop"or post-tussive vomiting, or apnea (with or without cyanosis) for infants up to 1 year of age. Cases are classified as suspected, probable, or confirmed based on Centers for Disease Control and Prevention (CDC) case definitions of pertussis [94]. Pertussis cases in this study include only ones that were symptomatic, sought medical help, and were reported to the PHSKC surveillance system. Surveillance data included information on cases' demographics (such as age, sex, race, ethnicity, home address), clinical symptoms, pertussis vaccination history, exposure history, transmission setting, pertussis treatment received, and epidemiologically-linked contacts. We geocoded home addresses of cases to their census tract of residence with ArcGIS 10.1 software [47].

4.2.2 Study Cohort

The study was conducted within a cohort of children registered in the Washington State Immunization Information System (WA-IIS), which is a lifetime registry that tracks immunization records for people of all ages in Washington State [95]. Healthcare providers such as primary care physicians, hospitals, and healthcare plans voluntarily report patient immunizations to WA-IIS. Additionally, birth certificates of children born in King County are loaded into the registry periodically. The study cohort was restricted to children born in King County after 2008 to ensure data completeness and accuracy (as per the recommendations of Dr. Patricia DeHart at the WA Department of Health). For recent years, WA-IIS has captured nearly all children born or living in King County through birth certificate data collected twice a month. Ninety-nine percent of children aged 4 months - 5 years have 2 or more immunizations recorded in the WA-IIS [95]. Using WA-IIS data, we created a retrospective cohort of 323,250 children 0 through 9 years of age, born or living in King County, WA, between 1st January, 2008 and 31st December, 2017, with a unique identifier for each child. Vaccination information including vaccine name and date of receipt, vaccine manufacturer, and healthcare facility of administration was requested for all pediatric vaccines recommended from birth through 9 years of age for each child. Demographic information included date of birth, sex, race, ethnicity, current residential address and county, and insurance information. Home addresses (or zip codes when home addresses were incorrect or not available) of WA-IIS participants were geocoded to their census tract of residence by WA Department of Health staff.

We calculated a census tract level neighborhood socioeconomic score (NSES) for each participant using the 2010 US census data [96]. Briefly, NSES score is an index score ranging from 0 to 100 created using measures of percent of adults 25 years and older with a high school degree (education), percent of civilian population 16 years and older with professional/managerial/executive occupations (employment), median household income (income), percent of families above the Federal Poverty Line (income), and percent of households with children under the age of 18 that are "female-headed" (no male present) [123]. The score was divided into quintiles, where the first quintile is the group with the lowest NSES (20th percentile or lower) and the fifth quintile is the group with the highest NSES (80th percentile or higher). NSES scores were estimated at the census tract level (N=397) because it is a geographic unit small enough to be a proxy of

neighborhood level exposure [124].

4.2.3 Linking surveillance and immunization datasets

Immunization records from the WA-IIS and surveillance data from PHSKC were merged based on a probabilistic matching algorithm that used participants' first name, last name, date of birth, sex, and city of residence [125]. Matching was performed using the `fastLink` package in R [126]. First and last names were linked using string distance matching and partial matches were allowed for these fields. All pairs that were matched with a posterior probability of >0.85 were retained for further review and false matches were removed manually. Unmatched records were also reviewed and matched manually based on additional data in the surveillance and immunization registry. After matching was completed, all personal identifiers were removed. A flowchart in the Supplementary Material (4.7.3 Fig 1) describes the data linkage process.

4.2.4 Censoring of registry participants

Follow-up time for participants began at birth. Follow-up continued until they were diagnosed with pertussis, died, moved to another county or changed healthcare providers (entered in the registry as 'moved or gone elsewhere' or MOGE), or until the end of follow-up period on 12/31/2017, whichever came first. Participants were flagged as "inactive" in the registry if they moved or died. When not flagged as inactive, participants were assumed be to be "active" members of the registry. As participants were not actively followed as in a prospective cohort study, vaccination dates for pediatric vaccines administered between birth through 9 years of age (except pediatric influenza vaccines) were considered proxies for continued enrollment in the WA-IIS and residence in King County. The flowchart in the Supplementary Material (4.7.3 Fig. 2) describes details of censoring. The following rules were used sequentially for censoring individuals in this study:

- Participants diagnosed with pertussis were censored on the date of diagnosis.

- If no immunization (*any pediatric immunization*) is ever recorded for a participant, then they are censored on their first birthday.

- For participants indicated as "active" in the WA-IIS and whose current residence is King County, follow-up ended at two years after the last recorded vaccine if last vaccine was recorded at age < 36 months, or at the end of study period if last vaccine was recorded at age ≥ 36 months.

- For active and inactive participants that no longer live in King County, follow-up ended on the date they moved out of King County or changed provider (recorded as Date MOGE or "Moved Or Gone Elsewhere" in the WA-IIS), if this date was recorded.

- For inactive participants missing MOGE date, follow up ended on the date their record was last updated for any reason (recorded as Date Last Update in the WA-IIS), assuming that they were King County residents at least until that date.

4.2.5 Timeliness of DTaP doses in the cohort

For each participant, age at vaccination (in days and months) was calculated using date of birth and date of vaccination. Days under-vaccinated for each DTaP dose were estimated using CDC Advisory Committee on Immunization Practices (ACIP) recommendations for minimum ages of vaccination and minimum acceptable intervals between doses (Table 4.1) [127] and a metric similar to one described by Luman, et al [116] [117]. ACIP recommendations in months and weeks were converted to days. ACIP defines a 4-day grace period before minimum age and recommended intervals during which vaccine doses may be considered valid. Any doses that were given before this grace period had to be repeated. For e.g. the minimum recommended age for DTaP dose 1 is 6 weeks, which corresponds to 38 days (42 days – 4 days grace period) for this analysis. A dose is not considered delayed as long as it administered within 30 days of the recommended age. Thus, children were considered to be vaccinated without delay if they received all 5 doses of DTaP within 4 days before minimum acceptable age through 30 days after recommended age ranges for each dose. Furthermore, if a child started the series late, but the rest of the doses in the series were administered at the correct intervals, then only the first dose was considered delayed and the rest of the doses in the series were

considered timely.

4.2.6 Definitions of vaccine timeliness: Under-vaccinated vs. delayed

Children who received fewer than the recommended doses at a given age (<3 doses by age 19 months, <4 doses by age 5 years, and <5 doses through age 9 years) were defined as "under-vaccinated". Children who received the recommended number of doses by a given age but received them outside the recommended window specified in Table 4.1 were defined as "delayed". For example, a child who received DTaP dose 3 at age 9 months instead of 6 months will be delayed but not under-vaccinated at age 19 months. Timeliness was calculated for each dose of DTaP. Definitions of undervaccination and delay for each DTap dose can be found in the Supplementary Material (4.7.1.)

Technically, under-vaccinated children can also be considered "delayed", and other studies that evaluated the effect of vaccine timeliness on pertussis risk indeed do not differentiate between the two [118] [119] [117] [128]. However, additional doses, even when delayed, can give more protection against pertussis, so it is worthwhile considering the two exposure groups separately.

4.2.7 Trends in vaccination timeliness

Trends in proportion of children delayed for DTaP doses for each birth cohort from 2008 through 2017 were analyzed using Poisson regression. Age-appropriate DTaP uptake over time was estimated by the Kaplan-Meier method with age as the timescale. Vaccination coverage at age a (in months) was estimated by $1-S(a)$, where the Kaplan-Meier survival function $S(a)$ is the cumulative probability of being vaccinated by age a [129].

4.2.8 Association between delayed vaccination and pertussis risk

The exposure in this study is under-vaccination or delay in vaccination, and the outcome is suspected, probable, or confirmed pertussis cases reported to PHSKC. DTaP doses are administered at age 2, 4, and 6 months (primary series), 15-18 months (second year booster), and 4-6 years (pre-school booster). In keeping with these recommendations, we measured the association of timeliness of DTaP primary series, second year booster dose, and pre-school booster dose with pertussis risk through ages 19 months, 5 years,

and 9 years, respectively. We measured the association between delay in series initiation and pertussis risk through 12 months of age. There is considerable variation in the schedules used for the primary vaccination series. In Scandinavian countries, a "2p+1" series (vaccine given at 2, 4 and 11/12 months), rather than a "3p" series (vaccine given at 2, 4, and 6 months) is used [31] [112]. Thus, we also compared pertussis risk between children who received a "3p" schedule vs "2p+1"schedule.

Log binomial models were used to calculate pertussis risk ratios and corresponding 95% CIs comparing children with delayed or missing vaccinations to those with timely vaccinations. Person-time at risk was calculated in months and used as an offset in the models. Additionally, we also measured the association between series initiation and series completeness using a Poisson regression model. Data management and manipulation was done using SAS 10.3 [130] and R [66]. Data analysis were done using R [66].

The log-binomial models are written as:

$$log(Y) = \alpha + \beta_1 * Delay.vacc + \beta_{2.1-2.4} * NSES + \beta_3 * age + log(\text{Time-at-risk}) + \epsilon \tag{4.1}$$

where Y = number of pertussis cases
α = intercept
$exp(\beta_1)$ = Change in risk of pertussis comparing children with delayed vaccinations and/or under-vaccination to timely and complete vaccination, adjusting for other variables in the model.
$exp(\beta_{2.1-2.4})$ = Change in risk of pertussis comparing children living in census tracts with NSES scores in quintiles 2, 3, 4, and 5 to children living in census tracts with NSES score in quintile 1, respectively
$exp(\beta_3)$ = Change in risk of pertussis per unit increase in age

Association between delay in primary series and pertussis risk through age 19 months

Per ACIP recommendations, a child should receive 3 doses of DTaP by 7 months of age. Hence, for this model, follow-up began at 7 months of age and continued up to 19 months of age or time of event/censoring. We excluded children who had fewer than 7 months of follow-up time since birth because they were not followed up long enough to

receive 3 doses of vaccine. Only cases diagnosed with pertussis between 7 months and 19 months of age were included in the case count. Children that became cases after age 19 months contributed time at risk in this model. The estimate of $\text{Exp}(\beta_1)$ gave the risk of pertussis up to age 19 months comparing children delayed or under-vaccinated for primary series to children with on-time primary series, adjusted for NSES and age. We re-ran the model restricting the cohort to children who received 3 or more DTaP doses by age 19 months, to assess association between pertussis risk and vaccine delay among those who were not under-vaccinated.

Association between delay in second year booster and pertussis risk through age 60 months

Per ACIP recommendations, a child should receive 4 doses of DTaP by 19 months of age. Hence, for this model, follow-up began at 19 months of age and continued up to 60 months of age or time of event/censoring. We excluded children who had fewer than 19 months of follow-up time since birth because they were not followed up long enough to receive the 4 doses of vaccine. Only cases diagnosed with pertussis between 19 months and 60 months of age were included in case count. Children that became cases after age 60 months contributed time at risk for this analysis. The estimate of $\text{Exp}(\beta_1)$ gave the risk of pertussis up to age 60 months comparing children that were delayed and /or under-vaccinated for second year booster to children with on-time second year booster, adjusted for NSES and age. We re-ran the model restricting the cohort to children who received 4 or more DTaP doses by age 60 months.

Association between delay in pre-school booster and pertussis risk through age 9 years

Per ACIP recommendations, 5 doses of DTaP are recommended by age 60 months (5 years). For this model, follow-up began at 60 months of age and continued until the end of study period (12/31/2017). We excluded children who had fewer than 60 months of follow-up time since birth because they were not followed up long enough to receive 5 doses of the vaccine. Only cases diagnosed with pertussis between 5 years and 9 years of age were included in case count. The estimate of $\text{Exp}(\beta_1)$ gave the risk of pertussis up to age 9 years comparing children who were delayed and/or under-vaccinated for the preschool booster to children with on-time preschool booster, adjusted for NSES and age. We re-ran the model restricting the cohort to children who received 5 or more DTaP doses by age 9 years.

Association between delay in series initiation and pertussis risk in the first year of life

For this model, follow-up began at 3 months of age (age at which DTaP dose 1 is recommended) and continued up to 12 months of age. We excluded children who had fewer than 3 months of follow-up time since birth because they were not followed up long enough to receive the first dose. Only cases diagnosed with pertussis between 3 months and 12 months of age were included in case count. Children that became cases after age 12 months contributed time at risk for this model. The estimate of $\text{Exp}(\beta_1)$ gave the risk of pertussis at age 12 months comparing children with delayed series initiation to children with timely series initiation, adjusted for NSES and age. We re-ran the model restricting the cohort to children who received 3 or more DTaP doses by age 12 months.

Pertussis risk up to age 24 months comparing children on 3p (accelerated) vs. 2p+1 (long) schedule

For this model, follow-up began at 3 months of age and continued up to 24 months of age. Only cases diagnosed with pertussis between 3 months and 24 months of age were included in case count. Children that became cases after age 24 months contributed exposed time at risk for this analysis. The estimate of $\text{Exp}(\beta_1)$ gave the risk of pertussis at age 24 months comparing children that received the DTaP 3p series to children with DTaP 2p+1 series, adjusted for NSES and age.

4.3 Results

A total of 438 cases of pertussis born between 1st January, 2008 and 31st December, 2017 were reported to the PHSKC surveillance department. Of these, 111 (26%) were younger than 6 months, 54 (10%) were 7-11 months old, 208 (47.5%) were 1-4 years old, and 65 (13.6%) were 5-9 years old. Forty-nine (11.2%) cases required hospitalization.

For the 323,250 children aged 0 to 9 years born or living in King County, WA, and registered in WA-IIS, complete information was available for first names, last names, and birth date. Only 36 participants were missing zip code of current residence. Missingness for demographic and immunization-related variables are described in the Supplementary Material (4.7.2 Table 1). About 7% (22,651) of the participants were marked as "Inactive" in the WA-IIS and <0.5% (1,493) had died. Of the 438 pertussis cases, 404 (92%) were successfully linked to the 323,250 children in the WA-IIS (Supplementary Material

4.7.3 Fig 1). We excluded 87 children from the WA-IIS with missing or unknown sex. The final cohort consisted of 323,163 WA-IIS participants aged 0-9 years born between 2008 and 2017, each with detailed vaccination information for DTaP and other pediatric vaccines as well as their pertussis case status (Supplementary Material 4.7.2 Table 2).

We assumed that if a child had no record of a DTaP dose in the WA-IIS, they did not receive it. Of the 316,404 children older than 3 months, 19,443 (6.3%) had no DTaP dose recorded in the WA-IIS and 116,193 (36.7%) were delayed for at least one DTaP dose (Table 4.2). There were no differences by sex between unvaccinated, delayed, and on-time, fully vaccinated children. Of those with delayed vaccinations, 12% were African American, compared to only 6.4% of those with timely vaccinations. Children living in census tracts with the lowest NSES scores were more likely to be delayed for DTaP vaccinations. A higher proportion of both unvaccinated (21.9% vs 18.1%) and delayed children (24.5% vs 16.2%) were found to be residing in census tracts with the lowest NSES quintile compared to the highest NSES quintile. Of the 404 pertussis cases, 116 (28.7%) were unvaccinated, 149 (36.9%) were delayed, and 139 (34.4%) received on-time and age-appropriate vaccinations.

4.3.1 Trends in vaccine timeliness

Table 4.3 describes timeliness of DTaP doses in this cohort. Delay in vaccination and under-vaccination increased with age. A total of 19,943 (6.3%), 28,096 (9%), and 34,435 (11.2%) of children eligible for DTaP doses 1, 2 and 3 respectively, did not receive it. Of the ones that did receive primary doses, 35,754 (11.3%) were delayed for dose 1 (mean delay 13.8 days), 50,727 (16.3%) were delayed for dose 2 (mean delay 16.1 days), and 67,895 (22.1%) were delayed for dose 3 (mean delay 37.3 days). Thus, vaccine delay among children who eventually received them was not longer than 5 weeks. Of the children eligible for DTaP boosters, 45,696 (16.5%) never received their second-year booster and 56,518 (20.4%) were delayed, whereas 20,409 (12%) never received their preschool booster and 50,388 (29.6%) were delayed. Again, mean delay for boosters among those who did receive them was not substantial likely due to wider windows of recommended ages. Trend analysis suggests while timeliness for each dose improved for successive birth cohorts, this change was not statistically significant (Fig. 4.1; Supplementary Material 4.7.2 Table 2).

Age-appropriate pertussis vaccination uptake is shown using inverse Kaplan Meier curves in Fig. 4.2. By 3 months of age, 86.2% (95% CI: 86.1%, 86.3%) of the cohort had received DTaP dose 1. By 5 months of age, 80% (95% CI: 79.9%, 80.2%) of the cohort had received DTaP dose 2. By 7 months of age, 73.9% (95% CI: 73.7%, 74.0%) had received DTaP dose 3, while $> 90\%$ coverage with three doses, which is nationally recommended, was achieved by age 14 months [131]. By 19 months of age, 67.9% (95% CI: 67.7%, 68.1%) had received DTaP4 and $> 85\%$ coverage for DTaP dose 4 was achieved by 48 months of age. By age 7 years, 86.2% (95% CI: 85.8%, 86.1%) of cohort had received DTaP dose 5, meeting the national vaccine coverage target for five doses of DTaP.

4.3.2 Association between under-vaccination and vaccine delay and age-specific pertussis risk

Table 4.4 shows the results of association between DTaP vaccine timing and number of doses and age-specific pertussis risk. Among children 7 months to 19 months of age, the risk of pertussis was 4.8 times higher (95% CI: 3.1, 7.6) for children delayed or under-vaccinated for the primary series, compared to those who completed the series on time. When restricted to children who had received 3 or more doses, this association was not statistically significant (RR: 0.86; 95% CI: 0.3, 2.2). This suggests that receiving the age-appropriate number of doses is crucial to protect against pertussis, even if the doses are administered with a few weeks' delay. For children 19 months to 60 months of age, the risk of pertussis was 3.2 times higher (95% CI: 2.3, 4.5) among children who were delayed or under-vaccinated for the second-year booster compared to those who were age-appropriately and timely vaccinated. When restricted to children who received 4 or more doses by 60 months of age, again there was no statistically significant association between delayed second-year booster and pertussis risk for (RR: 0.8; 95% CI: 0.5, 1.4). For children 5 to 9 years of age, risk of pertussis was 4.6 times higher (95% CI: 2.6, 8.2) among children delayed or under-vaccinated for preschool booster, compared to those age-appropriately vaccinated. Again, this association did not hold for children who had received 5 or more doses of DTaP by age 5 years.

Delay in initiation of primary DTaP series resulted in a 3.5 fold increase (95% CI: 2.3, 5.5) in pertussis risk in the first year of life. Again, risk was not elevated when the analysis was restricted to children who received 3 DTaP doses by age 7 months. This again suggests that initiating the series in a timely manner may not be as important as

receiving all recommended vaccine doses in the series. However, those who initiated the series late were 48% less likely (95% CI: 47%, 49%) to complete the DTaP series on time (Table 4.5). There was no statistically significant difference in pertussis risk in the first 24 months of life between children who received the accelerated 3p schedule compared to those that received the long 2p+1 schedule (Table 4.4).

4.4 Discussion

In this study, we measured the association between vaccine timeliness and number of doses on age-specific pertussis risk among infants and young children in King County, WA. Trends in vaccine delay have essentially remained unchanged since 2008 for this cohort. We found that receiving fewer than recommended doses by a given age increases risk of pertussis in children, despite high overall vaccine coverage. Even when administered with some delay, getting the primary series in the first year of life and the two booster doses in the second and fifth year effectively reduces pertussis risk in this age group.

Linking the WA State Immunization Information System (WA-IIS) with notifiable disease surveillance data collected by the local health department in King County (PHSKC) allowed us to create a population-based cohort with near-complete ascertainment of pertussis vaccination status for over 320,000 children aged 0-9 years born or living in King County between 2008 and 2017. WA-IIS has a high degree of internal and external validity, and the vaccination and demographic data elements are highly complete, making it a cost-effective and useful tool for answering our research question [132]. It is challenging to compare different DTaP vaccination schedules between countries because of differences in case ascertainment methods, the types of vaccines used, and the vaccine coverage achieved. This dataset allowed us to directly compare different vaccination schedules, such as "3-primary" vs "2+1 primary" doses in the first two years of life within the same population [122].

Other studies have found vaccine delay to be associated with increased pertussis risk but they did not differentiate between children who received fewer than the recommended number of doses and those who received all the doses, but with a delay [118] [119] [128]. It also appears that these studies did not measure the dose-specific effect of delay on pertussis for the appropriate age groups at risk, potentially resulting in misclassification

of person-time at risk. For example, Huang et al. [118] measured the effect of delay in any of the four of DTaP vaccine doses on pertussis incidence among all children 3 to 36 months old compared to no delay. But one would not expect delay in dose 4 to influence pertussis risk at age 12 months. Even though our conclusions are the same, person-time at risk is more carefully defined in our study and provides direct evidence in favor of the ACIP and World Health Organization (WHO) recommendations for vaccine schedules in the US.

The WHO recommends that the 3-dose infant primary series with DTaP vaccine should be completed by 6 months of age to maximize protection in infants who are at risk for severe disease and mortality [31]. UK switched from a 3-5-11 month primary series to a 2-3-4 month series and found significant reduction of cases in the 6 -12 month age group. A systematic review also found lower clinical protection and antibody titres with 2p+1 schedules compared to 3p schedules in the first year of life [133]. While we found evidence that receiving fewer than 3 doses by age 7 months increased risk of pertussis in the first year of life, we found no evidence that an accelerated 3p schedule is more effective than a longer 2p+1 schedule in protecting children during the first and second year of life in this cohort. This might be because <6% of the cohort received the 2p+1 schedule and only one case was reported in this group, resulting in effect estimates that were not statistically significant with wide confidence intervals. A systematic review of comparing effectiveness of different schedules in infant immunization against pertussis drew conclusions similar to ours [112]. Finland, Denmark, and Sweden, that use the 2p +1 schedule for the primary series and have high national DTaP vaccination coverage, have not experienced a resurgence in pertussis [31]. This suggests that a long primary schedule could be used for countries that are able to maintain high vaccination coverage for primary series.

Timing of initiation of the primary series did not influence pertussis risk in the first year of life, as long as three primary doses were administered in the first year. This result confirms the finding of another study in the Netherlands that found that initiating the primary schedule at 2 months instead of 3 months of age did not appreciably decrease pertussis incidence in this age group [134]. However, children who initiated the primary series late were less likely to complete it in the first year of life, as has also been seen in a study done in Australia [135]. There is evidence for incremental protection after each additional dose, so it is essential to be completely vaccinated with three primary doses for full protection against pertussis [114] [136]. Thus, clinicians should encourage

parents to initiate DTaP primary series at the earliest recommended age to ensure series completion and effective protection against pertussis.

Our study found evidence in favor of WHO's recommendation of a second year booster to be given at age 18 months. Pertussis incidence was higher among children aged 2-5 years who did not receive the second-year booster. Australia experienced a similar increase in pertussis incidence among 2-3 year olds when they discontinued the 18-month booster in 2003, which is also considered one of the main driving factors of pertussis resurgence in Australia [137]. Similarly, those who received a preschool booster dose at 4-6 years of age had lower risk of pertussis up to age 9 years, providing support for the recommendation of a preschool booster in the US. Studies have shown that the impact of the fifth dose might be short-lived, necessitating an adolescent dose at ages 11 or 12 years [113] [134] [138]. Modeling studies have shown that, at least in the US, school-age children are core transmission groups that help sustain pertussis transmission chains due to increased contact rates [22] [17]. Thus, a booster dose given to school age children between ages 4-6 years could be crucial to both protect them against pertussis as well as to reduce overall pertussis transmission.

Our study has limitations. Firstly, there could be misclassification of vaccination status if a child received a DTaP dose but it was not entered into the WA-IIS by the health-care provider. The WA-IIS also does not consistently record exactly when a participant moved out of King County, resulting in potential misclassification of person-time at risk. The WA-IIS does not capture important confounders of the association between delay in vaccination and pertussis risk such as household size, adolescent siblings, maternal education, and day care or school attendance. We also did not have information on reasons of vaccine delay (other studies have found vaccine hesitant parents to be a reason for vaccine delay) [128] [139]. Thus, we cannot explore the reasons of vaccine delay in this cohort. Different countries have different vaccination schedules for DTaP primary series and boosters, and some countries still use the whole-cell pertussis vaccine in their primary series. The results of this study are, thus, generalizable only to countries that administer acellular pertussis vaccines and use the same schedule as the US.

In summary, our study found that under-vaccination with DTaP vaccine increases pertussis risk for infants and young children, but a small delay in vaccine administration has no apparent impact on risk as long as the recommended number of doses are received in the first 6 months, second year and fifth year of life. Thus, parents should be

encouraged to follow recommended vaccine schedules, and in the event of a delay, the next dose in the series should be administered at the earliest possible opportunity. The results of this study indicate that the WHO and ACIP recommendations for DTaP series in the US are effective in reducing pertussis incidence among infants and young children.

4.5 Tables

Table 4.1: ACIP recommendations for DTaP vaccine schedule and interval for the US

DTaP dose	**Age of administration**	**Mininum age of administration**$^{*\circ}$	**Interval between doses**$^{\circ}$	**Age of delay**$^{\circ}$
Dose 1	2 months	38 days	90 days	
Dose 2	4 months	66 days	28 days	150 days
Dose 3	6 months	94 days	28 days	210 days
Dose 4	15-18 months	266 days	180 days	578 days
Dose 5$^{\pm}$	4-6 years	1456 days	180 days	2555 days

ACIP: Advisory Committee of Immunization Practices

* Grace period of 4 days included in the minimum age of administration
$^{\circ}$ Weeks and months are converted to days based on the method described in Luman, et al [117]
$^{\pm}$ No dose 5 required if dose 4 is given after age 4

Table 4.2: Characteristics of Washington State Immunization Information System participants 3 months or older born or living in King County, Washington between 2008-2017 by DTaP vaccination status

Characteristics	**Unvaccinated (N= 19,943)***	**Delayed (N=116,193)**$^{\circ}$	**Not delayed (N=180,268)**	**P-value**$^{\pm}$
Sex, N(%)				
Male	10165 (50.97)	59374 (51.1)	92486 (51.3)	0.43
Female	9778 (49.03)	56819 (48.9)	87782 (48.7)	
Race, N(%)				
Caucasian	9894 (49.61)	56517 (48.64)	94711 (52.54)	< 0.001
African-American	1641 (8.23)	14862 (12.79)	11590 (6.43)	
Asian	3080 (15.44)	15735 (13.54)	28720 (15.93)	
Native American	3 (0.02)	13(0.01)	20 (0.01)	
Other /Unknown	5325 (26.7)	29066 (25.02)	45227 (25.09)	
Ethnicity, N(%)				
Hispanic	2383 (11.95)	18707 (16.1)	22095 (12.26)	< 0.001
Non-Hispanic	16936 (84.92)	95188 (81.92)	152259 (84.46)	
Unknown	624 (3.13)	2298 (1.98)	5914 (3.28)	
NSES score, N(%)				
Q1 (lowest)	4368 (21.9)	28487 (24.52)	31997 (17.75)	< 0.001
Q2	3909 (19.6)	24727 (21.28)	32823 (18.21)	
Q3	4077 (20.44)	23271 (20.03)	36506 (20.25)	
Q4	3884 (19.48)	20589 (17.72)	37524 (20.82)	
Q5 (highest)	3610 (18.1)	18799 (16.18)	40817 (22.64)	
Pertussis cases	116 (28.71)	149 (36.88)	139 (34.41)	< 0.001
Median age of follow up (months)	11.8	61.9	49.2	< 0.001

Continued

Table 4.2 – *Continued*

Characteristics	**Unvaccinated (N= 19,943)***	**Delayed (N=116,193)°**	**Not delayed (N=180,268)**	**P-value±**
Insurance, N(%)				
CHIP	0 (0)	7 (0.01)	2 (0)	< 0.001
Medicaid	1643 (8.24)	44694 (38.47)	36417 (20.2)	
Native American or Alaskan	25 (0.13)	714 (0.61)	521 (0.29)	
Private	2666 (13.37)	46971 (40.42)	110218 (61.14)	
Underinsured	26 (0.13)	416 (0.36)	492 (0.27)	
Uninsured	223 (1.12)	1217 (1.05)	956 (0.53)	
Unknown	15360 (77.02)	22174 (19.08)	31662 (17.56)	
Birth cohort				
2008	2154 (6.06)	15386 (46.1)	17992 (53.9)	< 0.001
2009	1991 (5.74)	14472 (44.3)	18194 (55.7)	
2010	1813 (5.38)	13824 (43.32)	18085 (56.68)	
2011	1723 (5.18)	13149 (41.67)	18403 (58.33)	
2012	1468 (4.42)	12932 (40.72)	18825 (59.28)	
2013	1761 (5.49)	12599 (41.55)	17721 (58.45)	
2014	2192 (6.9)	11083 (37.45)	18515 (62.55)	
2015	2310 (7.53)	9959 (35.09)	18422 (64.91)	
2016	2492 (8.3)	8383 (30.44)	19160 (69.56)	
2017	2039 (9.53)	4406 (22.76)	14951 (77.24)	

NSES: Neighborhood Socio-Economic Status

* Unvaccinated children have 0 doses of DTaP vaccine

° At least one DTaP dose that is delayed

± Chi-squared P-value for categorical variables and Wilcox rank sum P-value for continuous variables

Table 4.3: Timeliness of DTaP vaccine by dose among Washington State Immunization Information System participants born or living in King County, WA between 2008-2017.

Vaccine dose	**Measured at age**	**Timeliness**					
		On-time, N(%)	Delayed, N(%)	Never, N(%)	Total	Mean age (months)	Mean delay (days)
Dose 1	3 months	260707 (82.4)	35754 (11.3)	19943 (6.3)	316404	3.46	13.8
Dose 2	5 months	232837 (74.71)	50727 (16.28)	28096 (9.01)	311660	5.53	16.1
Dose 3	7 months	204534 (66.65)	67895 (22.13)	34435 (11.22)	306864	8.24	37.3
Primary series	7 months	219799 (71.6)	52630 (17.2)	34435 (11.2)	306864	8.24	
Booster 1	19 months	175113 (63.14)	56518 (20.38)	45696 (16.48)	277327	20.42	34.6
Booster2	60 months	99614 (58.46)	50388 (29.57)	20409 (11.98)	170411	52.88	-123

Table 4.4: Estimated risk of pertussis comparing children that were undervaccinated or vaccinated with delay with those who were age appropriately vaccinated using Log binomial models

Model (Delay/Undervaccination vs.No delay)	**Cohort**	**Cohort size**	**Start follow-up time**	**End follow-up time**	**Total person-time at risk (person-months)**	**Total cases**	**Cases exposed**	**Cases unexposed**	**aRR**	**95% CI**
Primary series	7 months or older	298,166	7 months	19 months	3,621,122	85	54	31	4.81	(3.1, 7.6)
	7 months or older with $\geq$ 3 doses	257,913	7 months	19 months	3,163,925	36	5	31	0.86	(0.3, 2.2)
First booster	19 months or older	258,675	19 months	60 months	7,959,278	158	99	59	3.25	(2.3, 4.5)
	19 months or older with $\geq$ 4 doses	221,928	19 months	60 months	7,092,436	76	17	59	0.8	(0.5, 1.4)
Second booster	60 months or older	134,950	60 months	Age censored /end of study	4,010,335	55	38	17	4.6	(2.6, 8.2)
	60 months or older with $\geq$ 5 doses	111,387	60 months	Age censored/end of study	3,273,337	22	5	17	1.3	(0.5, 3.6)
Series initiation	3 months or older	301,494	3 months	12 months	3,586,294	92	38	54	3.56	(2.3, 5.5)
	3 months or older with 3 doses by age 7 months	295,325	3 months	12 months	3,423,860	41	0	41	NA	NA
3p vs 2p+1 schedule$^{\pm}$	3 months or older	301,494	3 months	24 months	5,178,190	157	1	60	3.9	(0.5, 28.8)

Exposure: Undervaccination and/or vaccination delay

* aRR: Risk ratios are adjusted for age and neighborhood SES score

$^{\pm}$ 3p: Accelerated 3-dose DTaP primary series to be administered at ages 2,4, and 6 months ; 2p+1: Long 3-dose DTaP primary series to be administered at ages 2,4, and 11/12 months continuous variables; 2p+1 is the exposed group

Table 4.5: Estimated risk of delay in DTaP series completion comparing children who initiated the series late vs those that initiated the series on time using Poisson regression

Model	**Cohort**	**Cohort size**	**Start follow-up time**	**End follow-up time**	**Total person-time at risk (person-months)**	**aRR***	**95% CI**
Effect of delay in starting series on series completion±	7 months or older	284,840	7 months	Age censored /end of study	15,266,023	0.52	(0.51, 0.53)

* aRR: Risk ratios are adjusted for age and neighborhood SES score

± Series completion defined as receiving age-appropriate number of vaccines

4.6 Figures

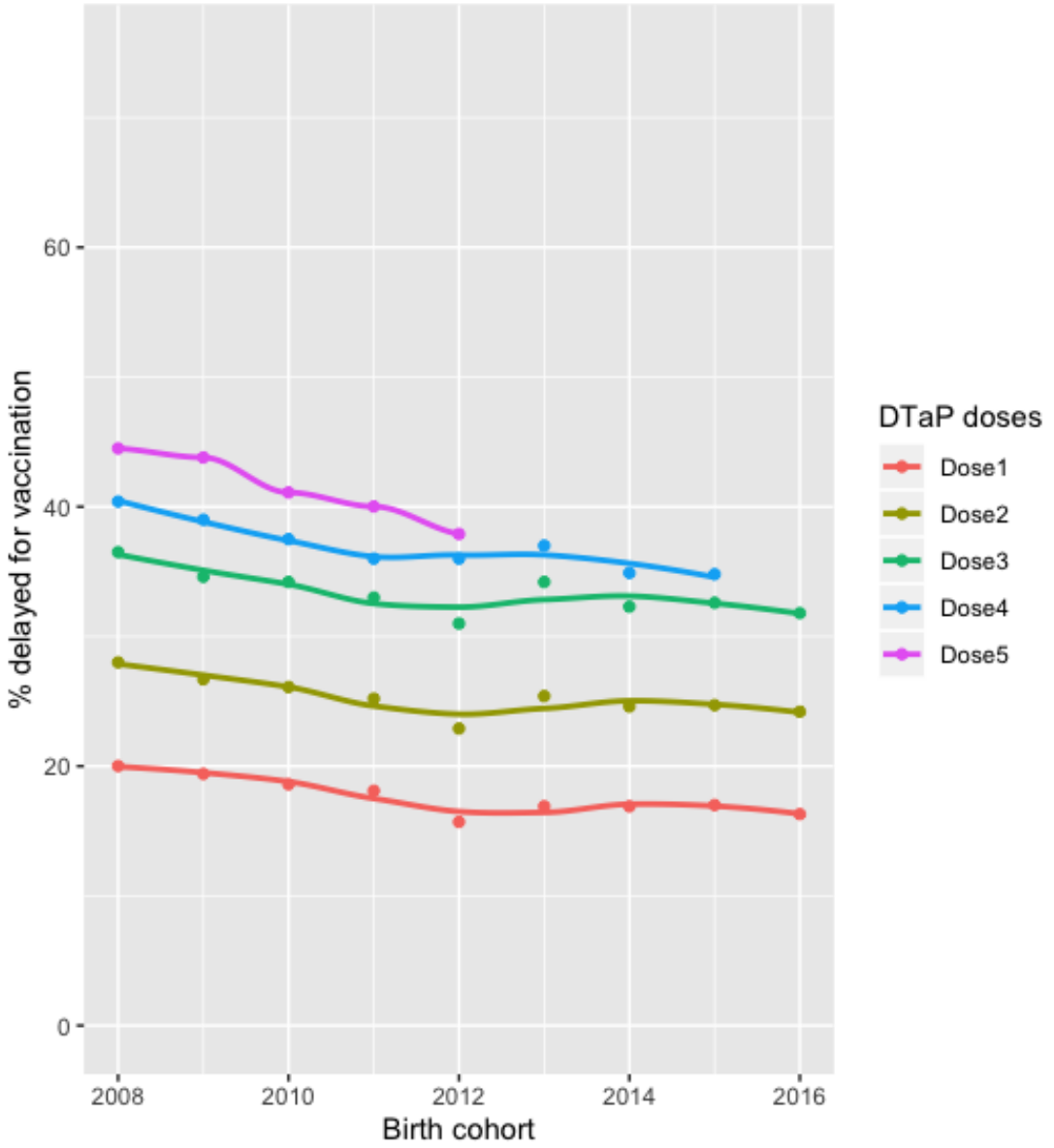


Figure 4.1: Trends in DTaP vaccination delay by birth cohort for 0-9 year olds in King County, WA

Trends in vaccination delay with DTaP doses by birth cohort. Poisson regression analysis suggested that trends were not statistically significant. Birth cohorts 2013-2016 did not receive DTaP5 and birth cohorts 2015-2016 did not recieve DTaP4. See Supplementary Table 2 for details.

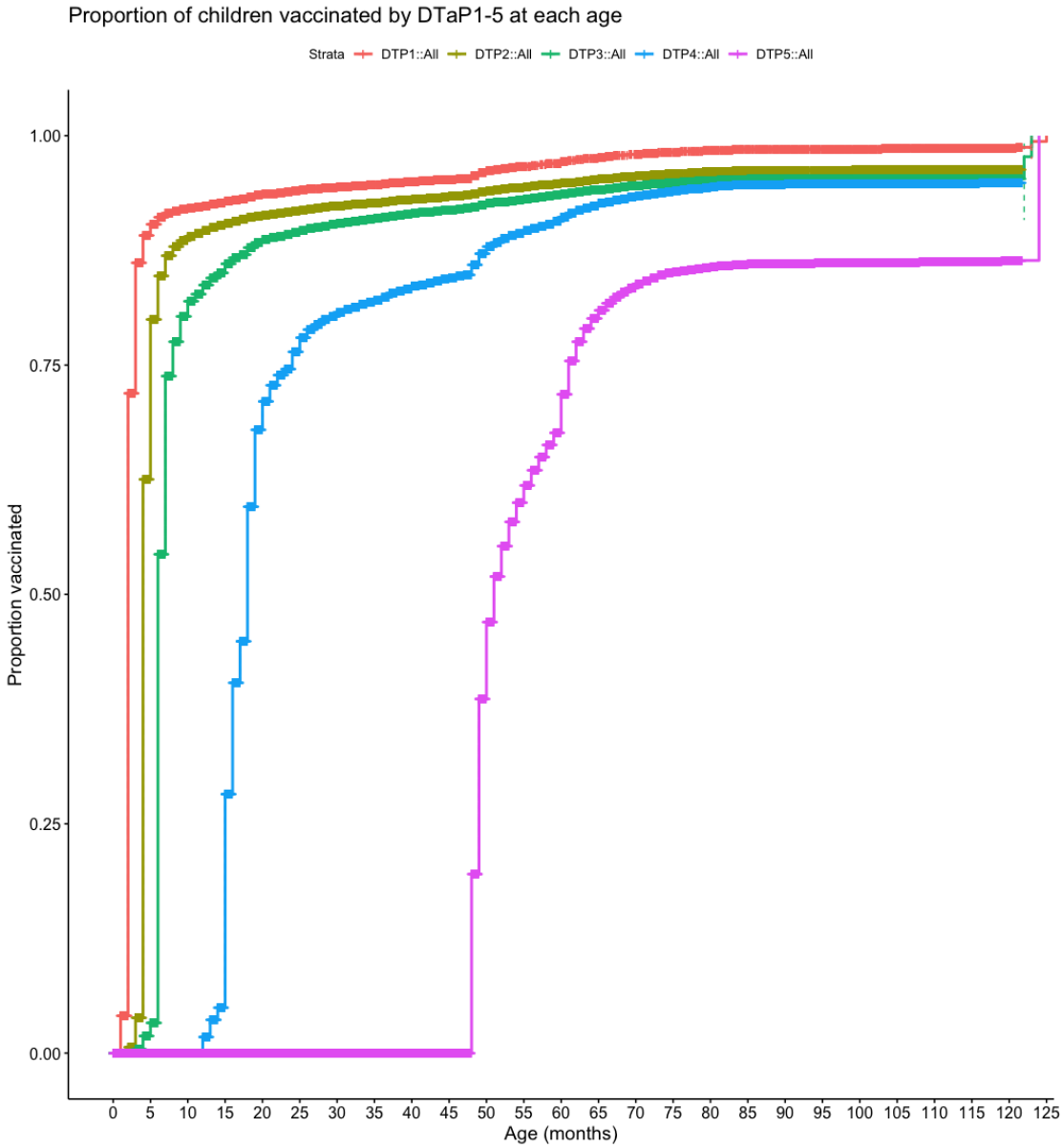


Figure 4.2: Age-appropriate DTaP vaccination using inverse Kaplan-Meier curves for 0-9 year olds in King County, WA

Proportion of 323,163 children aged 0 to 9 years vaccinated with DTaP1 (red), DTaP2 (yellow), DTaP3 (green), DTaP4 (blue), and DTaP5 (purple) at each age and 95% CIs are displayed using inverse Kaplan-Meier curve. Children in this analysis were born between January 2008 and December 2017. The 95% CIs are very narrow because the sample size is very large.

4.7 Supplementary Materials

4.7.1 Exposure definitions

On-time dose 1: Individuals 3 months or older who received dose 1 by 3 months of age, delayed otherwise.
On-time dose 2: Individuals 5 months or older who received dose 2 by 5 months of age, delayed otherwise.
On-time in dose 3: Individuals 7 months or older who received dose 3 by 7 months of age, delayed otherwise.
On-time in dose 4: Individuals 19 months or older who received dose 4 by 19 months of age, delayed otherwise.
On-time in dose 5: Individuals 5 years or older who received dose 5 by 7 years of age, delayed otherwise. If dose 4 was given after age 4, then dose 5 is not required and the child is considered "complete" for vaccine series.
On-time primary series: Individuals 7 months or older who received first 3 doses by 7 months of age, delayed otherwise. Doses 1 and 2 might be delayed, but child is considered to be on-time for primary series as long as all 3 doses are received by age 7 months.
On-time second year booster: Individuals 19 months or older who received 4 doses by 19 months of age, delayed otherwise. Doses 1, 2, or 3 might be delayed but child should get 4 doses by 19 months.
On-time preschool booster: Individuals 5 years or older who received 5 doses by 60 months of age, delayed otherwise. Doses 1, 2, or 3 might be delayed but child should get 5 doses by 60 months of age. Although dose 5 can be administered up to 7 years of age per ACIP, we decided to use a cut-off of age 5 years to receive dose 5 because most children begin pre-school at this age.
Delay in starting series: Individuals 3 months or older who had not started DTaP series by 3 months of age.

4.7.2 Supplementary Tables

Table 1 : Completeness of data elements from WA Immunization registry

Patient Characteristics, N=323,250	**Missing (%)**
First name	0 (0)
Last name	0 (0)
Gender	87 (0.03)
Birth date	0 (0)
Home address	0 (0)
Zip code of residence (CHS)	75,946 (23.5)
Zip code of current residence	36 (0.01)
Health care organization	10,187 (3.2)
Birth country	81,654 (25.3)
Birth state	81,710 (25.3)
Ethnicity	139,182 (43.1)
Race	119,493 (36.9)
Language	154,526 (47.8)
Insurance status	28,456 (8.8)
Vaccine Characteristics, N=7,444,325	**Missing (%)**
Administering facility	527,789 (7.1)
Manufacturer	4,622,248 (62.1)
Historically administered	1,344,164 (18.1)

Table 2 : Trends in vaccination delay for DTaP doses from 2008 to 2017 estimated by Poisson regression

DTaP Dose number	**Trend in delay, β (SE)**
1	-0.03 (0.03)
2	-0.01 (0.02)
3	-0.01 (0.02)
4	-0.02 (0.02)
5	-0.04 (0.05)

4.7.3 Supplementary figures

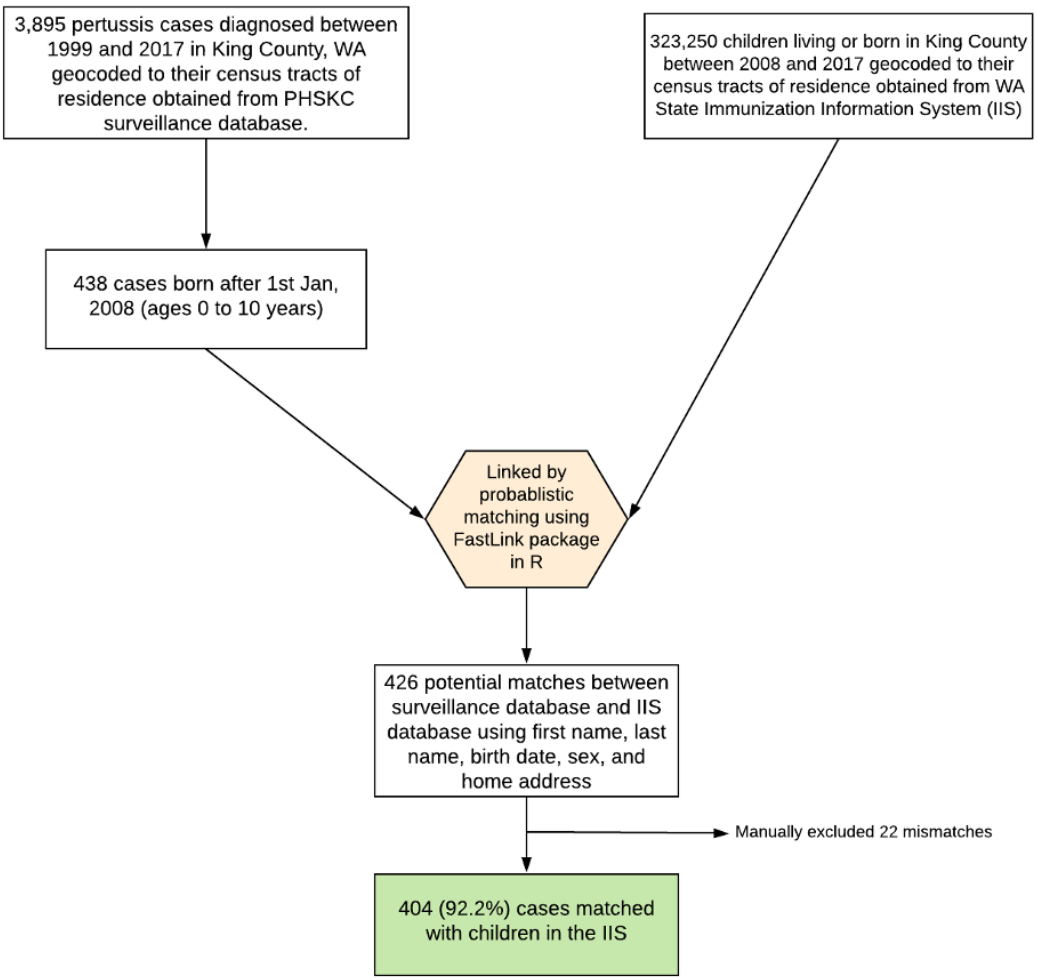


Fig 1: Flow chart describing probabilistic matching of surveillance database (pertussis cases) with immunization registry (immunizations)

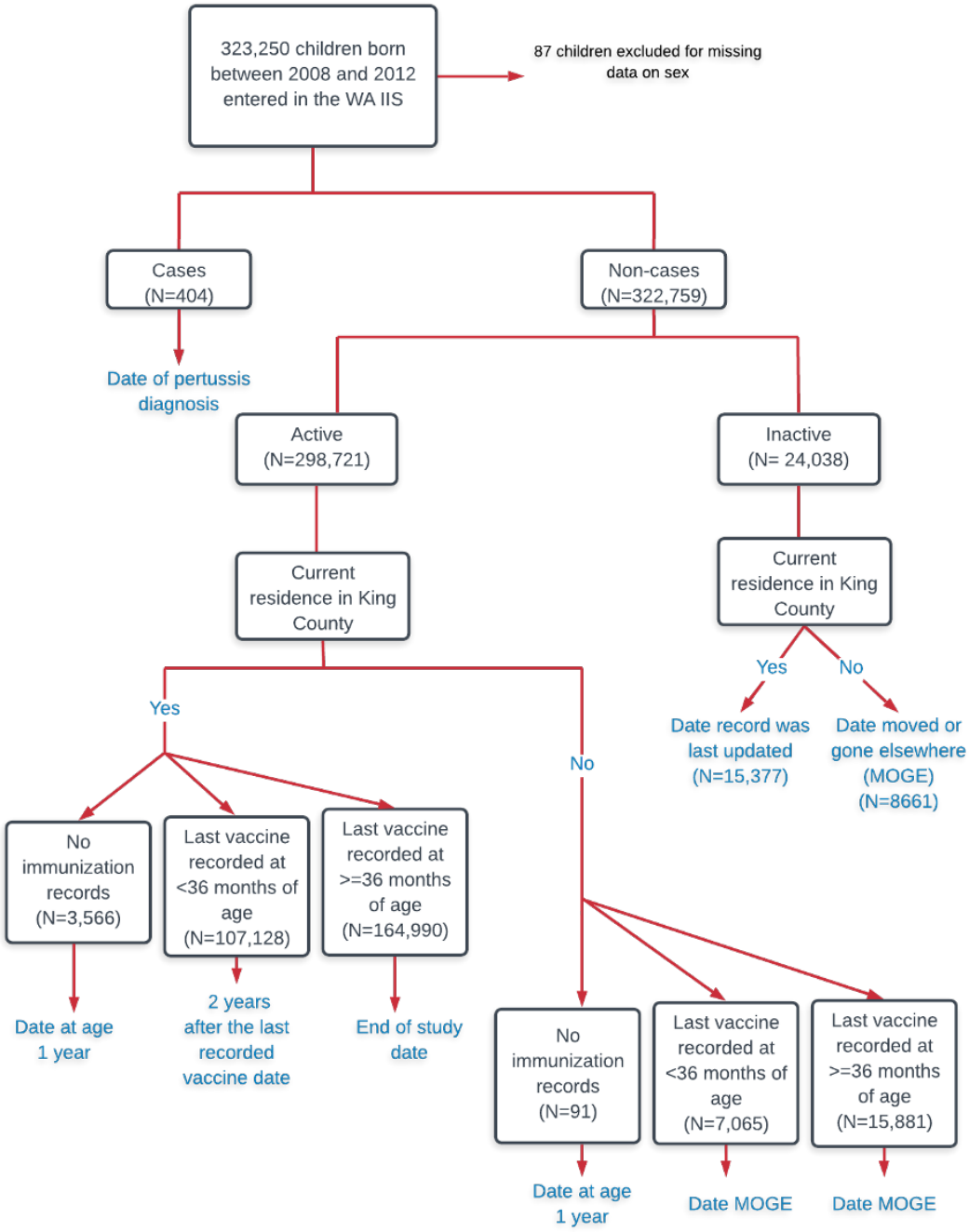


Fig 2: Flowchart describing censoring algorithm

All pediatric vaccines, except Influenza vaccines, administered to the children registered in the WA-IIS and recorded by the WA-IIS were used to determine censoring. Vaccines of all formulations (including unspecified formualtions) and combinations against Hepatitis A, Hepatitis B, Haemophilus Influenza B, Poliomyelitis, Diptheria-Tetanus-Pertussis, Menigococcus, Measles-Mumps-Rubella, Varicella, Pneumcoccus, Rotavirus were used. Historical shots for any of the recorded in the registry were also used.

Chapter 5

Estimating population-level effects of DTaP vaccine using routinely collected immunization data

Abstract

Background: The inability of the Diphtheria-Tetanus-acellular-Pertussis (DTaP) vaccine to reduce pertussis transmission from vaccinated persons and, thereby, to contribute to herd immunity has been hypothesized as one of the primary reasons for pertussis resurgence in countries with very high vaccine coverage. Measuring and reporting the different population-level effects of DTaP vaccination on pertussis disease in addition to direct effects can increase the public health utility of a vaccine.

Methods: We conducted a retrospective cohort study of children born between 1st January, 2008, and 31st December, 2017 (0-9 years of age) in King County, Washington, who were enrolled in the Washington State Immunization Information System (WA-IIS). DTaP vaccination data from WA-IIS was linked with pertussis case data from Public Health Seattle and King County. Census tract of residence for each case and non-case in the cohort was defined as their neighborhood, i.e., their "cluster". Cluster-level vaccination coverage was estimated as proportion of age-appropriately vaccinated children residing in the cluster. We estimated direct and population-level vaccine effectiveness for

the DTaP vaccine using the Cox proportional hazards model. Direct vaccine effectiveness was estimated by comparing pertussis risk in fully-vaccinated and under-vaccinated children. Different population-level vaccine effects were estimated by comparing pertussis risk in children living in census tracts in the highest vaccination coverage quartile to those living in census tracts in the lowest vaccination coverage quartile.

Results: Of the 310,398 children in the cohort older than 3 months of age, 73.4% were fully-vaccinated for their ages. Census tract level DTaP vaccination coverage ranged from 18% to 85%. For direct protection, estimated vaccine effectiveness was 76% (95% CI: 63% - 84%) in low vaccination coverage clusters and it decreased to 38% (95% CI: 13% - 68%) in high vaccination coverage clusters, after adjusting for confounders. The estimated indirect effect was 45.0% (95% CI: 1%, 70%), total effect was 93.9% (95% CI: 91%, 96%), and overall effect was 42.2% (95% CI: 19%, 60%).

Conclusion: Our findings suggest that DTaP vaccination provided direct as well as indirect protection in the highly immunized King County, WA. Routine DTaP vaccination programs may have the potential to provide not only protection for vaccinated individuals but also for the under-vaccinated individuals living in the same area.

5.1 Introduction

Vaccination can induce protection against disease at both the individual and population-level. Large-scale vaccination programs can result in reduction of disease transmission in the population and protection of even unvaccinated individuals against disease [13]. Estimating and reporting population-level effects of vaccines can help us understand the impact of a vaccine beyond its protective effects in vaccinated individuals and can make routine vaccination more cost-effective.

Despite showing high vaccine efficacy in randomized controlled trials against pertussis disease [83] [140], the ability of the Diphtheria-Tetanus-acellular-Pertussis (DTaP) vaccine to reduce infectiousness of breakthrough infections and transmission has been questioned. Recent studies in non-human primates showed that vaccination with acellular pertussis vaccine did not prevent colonization or transmission of (*B. pertussis*), suggesting that current acellular pertussis vaccines are ineffective against transmission [15]. Another study found evidence of asymptomatic transmission from vaccinated individu-

als [16] using wavelet analysis of long term pertussis incidence data and phylodynamic analysis methods. In contrast, some studies have inferred indirectly using population-level incidence data that increase in inter-epidemic periods and dramatic reductions in incidence and mortality among infants too young to be vaccinated after mass vaccination indicate presence of vaccine herd effects [10] [17]. A few studies have also provided direct evidence of population-level effects of pertussis vaccines in a field setting. A randomized controlled trial in Sweden estimated indirect effects of acellular pertussis vaccine as reduction in pertussis incidence among close contacts of DTaP vaccinated infants [18]. An observational study in Niakhar, Senegal, showed reduced pertussis transmission from vaccinated breakthrough cases compared to transmission from unvaccinated cases, concluding that pertussis vaccines likely provide indirect effects in a population [19]. Another prospective cohort study in Niakhar compared age-specific incidence during pertussis epidemics before and after large scale immunization and reported overall effects of the vaccine [141]. Amidst rising concerns about the inability of the current acellular pertussis vaccines to contribute to herd effects, there is a need for more recent evidence of indirect vaccine effects produced by DTaP vaccines. To our knowledge, this is the first study to measure different types of population-level effects for an acellular pertussis vaccine that was administered in a non-randomized fashion.

While methods for assessing population-level vaccine effects using randomized trials have been developed [142] [143] [13] [144] [145], randomized trials can be expensive and their results may not be generalizable to routine vaccination programs. For a vaccine that is already licensed and routinely used, such as the DTaP vaccine, a randomized controlled trial to measure vaccine indirect effects is neither feasible nor ethical. In such a situation, it might be possible and even preferrable to measure population-level vaccine effects using routinely collected data such as electronic medical records, insurance claims, and vaccination registries when available [146].

In this study, we estimated direct, indirect, total, and overall vaccine effects of the DTaP vaccine among vaccine-eligible children under 10 years of age living in King County, WA. We used a retrospective cohort study design and vaccination data wherein the DTaP vaccine is administered as part of a routine immunization program.

5.2 Methods

5.2.1 Study Cohort

The study was conducted within a cohort of children registered in the Washington State Immunization Information System (WA-IIS), which is a lifetime registry that tracks immunization records for people of all ages in Washington State [95]. Healthcare providers such as primary care physicians, hospitals, and healthcare plans voluntarily report immunizations to WA-IIS. In recent years, WA-IIS has captured nearly all children born or living in King County through birth certificate data collected twice a month. The study cohort was restricted to children born in King County after 2008 to ensure data completeness and accuracy (as per the recommendations of Dr. Patricia DeHart at the WA Department of Health). Ninety-nine percent of children aged 4 months - 5 years have 2 or more immunizations recorded in the WA-IIS [95].

We requested vaccination information including vaccine name, and date of receipt, vaccine manufacturer, healthcare facility of administration and medical insurance used for all pediatric vaccines recommended from birth through 9 years of age for each child. Demographic information included date of birth, sex, race, ethnicity, current residential address and county. Home addresses (or zip codes when home addresses were incorrect or not available) of WA-IIS participants were geocoded to their census tract of residence by WA Department of Health staff.

We calculated a census tract level neighborhood socioeconomic score (NSES) for each participant using the 2010 US census data [123]. Briefly, NSES score is an index score ranging from 0 to 100 created using measures of percent of adults 25 years and older with a high school degree (education), percent of civilian population 16 years and older with professional/managerial/executive occupations (employment), median household income (income), percent of families above the Federal Poverty Line (income), and percent of households with children under the age of 18 that are female-headed (no male present). The census-tract level score was divided into quintiles, where the first quintile is the group of census tracts with the lowest NSES (20th percentile or lower) and fifth quintile is the group with the highest NSES (80th percentile or higher). Distribution of census tract level NSES scores is in Supplementary Material (5.7.2 Fig 1). NSES scores were estimated at the census tract level because it is a geographic unit small enough to be a proxy of neighborhood level exposure [124].

5.2.2 Determination of DTaP vaccination status for the direct and population-level vaccine effects analyses

DTaP vaccine is routinely administered as a 5-dose series to children at ages 2, 4, 6 months, 15-18 months, and 4-6 years. For each participant, age at vaccination (in days and months) was calculated using date of birth and date of vaccination for each dose of pertussis. For the direct vaccine effect analysis, DTaP vaccination status was defined as "fully-vaccinated" (defined as receiving age-appropriate number of DTaP doses, i.e., at least 1 dose by age 3 months, 2 doses by age 5 months, 3 doses by age 7 months, 4 doses by age 19 months, and 5 doses by age 7 years) and "under-vaccinated" (had fewer than recommended doses given their age). For the population-level vaccine effects, DTaP vaccination status was defined as fully-vaccinated when children 7 months or older had received at least 3 DTaP doses, and under-vaccinated otherwise.

5.2.3 Pertussis case data

Pertussis is a nationally notifiable disease. Pertussis cases aged 2 months - 9 years reported between 1st January, 2008 and 31st December, 2017 were obtained from the Public Health Seattle and King County (PHSKC) surveillance database. The clinical case definition of pertussis used is a cough illness lasting $\geq$ 2 weeks with at least one of the following: paroxysms of coughing or inspiratory "whoop" or post-tussive vomiting, or apnea (with or without cyanosis) for infants up to 1 year of age [94]. Cases are classified as suspected, probable, or confirmed based on Centers for Disease Control and Prevention (CDC) case definitions of pertussis [94]. Pertussis cases in this study include only ones that were symptomatic, sought medical help, and were reported to the PHSKC surveillance system. Surveillance data included information on cases' demographics (such as age, sex, race, ethnicity, home address), clinical symptoms, pertussis vaccination history, exposure history, transmission setting, pertussis treatment received, and epidemiologically-linked contacts. We geocoded home addresses of cases to their census tract of residence with ArcGIS 10.1 software [47].

5.2.4 Linking immunization and surveillance data

Immunization records from the WA-IIS and surveillance data from PHSKC were merged based on a probabilistic matching algorithm that used participants' first name, last name, date of birth, sex, and city of residence [125]. Matching was performed using the

`fastLink` package in R [126]. First and last names were linked using string distance matching and partial matches were allowed for these fields. All pairs that were matched with a posterior probability of > 0.85 were retained for further review and false matches were removed manually. Unmatched recored were reviewed and matched manually based on additional data in the surveillance and immunization registry. After matching was completed, all personal identifiers were removed. Flowchart in Supplementary material (4.7.3 Fig. 1) describes the data linkage process.

5.2.5 Censoring of participants

Registry participants were followed up until they were diagnosed with pertussis, died, moved to another county or changed healthcare providers (entered in the registry as 'moved or gone elsewhere' or MOGE), or until the end of follow-up period on 12/31/2017, whichever came first. As participants were not followed over time like in a prospective cohort study, vaccination dates for pediatric vaccines administered between birth through 9 years of age (except pediatric influenza vaccines) were considered proxies for continued enrollment in the WA-IIS and residence in King County. Participants are flagged as "inactive" in the registry if they moved out of state or died to indicate that they are no longer active participants. However, this variable field is inconsistently recorded. For our anaylsis, we assume that a participant is "active" unless they are marked "inactive" or "dead" and the date they become "inactive" is recorded. Flowchart in Supplementary Material (4.7.3 Fig. 2) describes details of censoring. The following rules were used sequentially for censoring individuals in this study:

- Participants diagnosed with pertussis were censored on the date of diagnosis.

- If no immunization (any pediatric immunization) is ever recorded for a participant, then they are censored on their first birthday.

- For participants indicated as "active" in the WA-IIS and whose current residence is King County, follow-up ended at two years after the last recorded vaccine if last vaccine was recorded at age < 36 months, or at the end of study period if last vaccine was recorded at age ≥ 36 months.

- For active and inactive participants that no longer live in King County, follow-up ended on the date they moved out of King County or changed provider (recorded as Date MOGE or "Moved Or Gone Elsewhere" in the WA-IIS), if this date was recorded.

- For inactive participants missing MOGE date, follow up ended on the date their record was last updated for any reason (recorded as Date Last Update in the WA-IIS), assuming that they were King County residents at least until that date.

5.2.6 Vaccine coverage in clusters

To analyze population-level effects of DTaP vaccine, we defined the census tract in which the cases or non-cases reside as their "cluster". There are 397 census tracts or clusters in this study (which equals the total number of census tracts in King County). Census tracts were selected as clusters mainly because they are stable administrative units and information on socio-demographic variables that could be potential confounders is only available at census-tract level. Home addresses of the study population were geocoded to get their census tract of residence. Vaccine coverage was calculated for each cluster as number of children 3 months - 9 years of age who had age-appropriate number of DTaP vaccines divided by number of 0-9 year old children in the WA-IIS who were eligible for the DTaP vaccine as measured on 31st December, 2017. Vaccination clusters were divided into quartiles based on the proportion of children who were fully vaccinated against pertussis (i.e. vaccination coverage). Census tracts in the lowest quartile had vaccination coverage ranging from 18% to 69% while those in the highest quartile had vaccination coverage between 77% and 85%. We assumed that no inter-cluster transmission occurs, and that vaccination status of an individual in a cluster could affect others living that cluster, but not individuals living in other clusters [145]. Distribution of census tract level vaccine coverage into quartiles is shown in Supplementary Material (5.7.2 Fig. 1).

5.2.7 Cohort analysis

We used a retrospective cohort study design and survival analysis methods to estimate different vaccine effects in this study. Age is the time scale of analysis and participants were followed up until they were diagnosed with pertussis, died, moved to another county or changed healthcare providers (entered in the registry as 'moved or gone elsewhere' or

MOGE), or until the end of follow-up period on 31st December, 2017, whichever came first.

Direct effects: For direct effects estimation, follow-up time for participants began at age 3 months by which age the first DTaP dose is administered, i.e., zero time is age 3 months. Here, exposure is vaccination with DTaP (and unexposed means under-vaccinated). Vaccination status was included as a time-dependent exposure that indicated whether a person received dose 1 by age 3 months, dose 2 by age 5 months, dose 3 by age 7 months, dose 4 by age 19 months, and dose 5 by age 7 years. Vaccination status was estimated at each time interval and it was assumed to be independent of the vaccination status in the previous time interval. The hazard of pertussis at time t depends only on the value of the vaccination status at that time. Further, we assumed that the effect of DTaP doses is not cumulative. Children were considered unvaccinated for nth dose until they received the dose and contributed person-time as either exposed or unexposed to the vaccine during that time interval based on whether they received the dose. The effect of DTaP doses is modeled as a piecewise constant, with a constant effect within each of the five time periods.

Fig. 5.1 illustrates the time-varying nature of DTaP vaccination status in this study. Persons 1 and 2 were unexposed (unvaccinated) and exposed (vaccinated) for the entire duration of their follow-up period, respectively. Their follow-up ended when they were diagnosed with pertussis. Person 2 received their first DTaP dose by age 3 months, so they were exposed in the first time period between ages 3 and 5 months. They did not receive their second DTaP dose in time period 2, so they were unexposed/unvaccinated during this time. However, they received a dose during time period 3, so they contributed exposed person-time in time period. Person 3 was exposed during time periods 1 and 2, but did not receive any dose thereafter, so were unexposed for the rest of the time periods.

Population-level effects: We estimated population-level effects of being vaccinated with at least 3 doses of the DTaP vaccine in a cohort of children 7 months or older. Follow-up started at age 7 months, i.e. time zero was 7 months. Vaccination status for this analysis was defined as fully-vaccinated (received at least 3 DTaP doses by age 7 months) vs. under-vaccinated (received fewer than 3 doses by age 7 months), and was not time-dependent as in the direct effects analysis. We assessed direct, indirect, total, and overall vaccine effects from time zero to end of follow-up period using methods described in [13]. Our comparison group was under-vaccinated children instead of a placebo vaccine group.

Methods used here closely follow a cohort analysis for vaccine herd effects for cholera vaccine in Kolkata, India [147]. Indirect vaccine effect was estimated by comparing risk of pertussis in under-vaccinated children living in clusters in the lowest vaccination coverage quartile to clusters in the highest vaccination coverage quartile. Total vaccine effect was estimated by comparing risk of pertussis in fully-vaccinated children in the highest vaccination coverage quartile to under-vaccinated children in the lowest vaccination coverage quartile. Overall vaccine effect was estimated by comparing risk in all individuals living in clusters in highest vaccine coverage quartile to clusters in the low vaccine coverage quartile. Thus, for all population level vaccine effects, cluster-level vaccination coverage was the exposure variable. Table 5.1 defines the source populations and vaccine exposures for all vaccine effects.

We used the Cochrane-Armitage test to analyze changes in vaccine effectiveness trends by underlying vaccination coverage. Vaccine effect was calculated as (1-hazard ratio of pertussis comparing exposure categories) x 100% for both direct and population-level effects. We used the Cox proportional hazards model to estimate vaccine effects adjusting for the confounding factors, namely, birth cohort and neighborhood level SES quintiles. The model for estimating indirect effect was additionally adjusted for number of doses received, because the under-vaccinated group in this cohort consisted of children who could have received 0,1,or 2 vaccine doses. Adjusting the model by birth cohort accounted for calendar effects because baseline hazard of pertussis is expected to change over time. Because vaccine coverage was estimated using routine vaccination data and time of start of vaccination is variable for participants, coverage can be calculated in various ways. To check whether our results were influenced by the choice of the date when the coverage was estimated, we performed a sensitivity analysis by estimating vaccine coverage as proportion of children 3 months-7 years of age who had received age-appropriate DTaP vaccine doses as of 31st December, 2015.

5.2.8 Institutional Review

This study was reviewed by Washington State Institutional Review Board, University of Washington Institutional Review board, and the PHSKC Research Administration Review Committee.

5.3 Results

In the cohort used for estimating direct effects, there were 310,398 children older than 3 months of age, of whom 227,751 (73.4%) were fully vaccinated, 67,499 (21.7%) were partially vaccinated, and 15,148 (4.9%) had zero vaccines reported to the WA-IIS (Table 5.2). There were 298,093 children 7 months or older in the cohort used for estimating population-level effects, of which 270,361 (90.7%) had received 3 doses of DTaP vaccine by age 7 months.

Compared to children who were under-vaccinated, pertussis incidence was lowest among children who were age-appropriately fully-vaccinated for the DTaP series across all vaccine coverage groups. Among those who were partially vaccinated, incidence decreased with increasing vaccination coverage, and the Cochrane-Armitage trend test was statistically significant. The Cochran Armitage trend test for pertussis incidence by levels of vaccination coverage was not significant for the fully-vaccinated group and unvaccinated groups, although incidence decreased with increasing quartiles of vaccine coverage in the unvaccinated group (Table 5.3). The lack of statistical significance in incidence trend in the unvaccinated group might be due to fewer number of people and shorter follow-up time in this group.

Direct vaccine effectiveness (Table 5.4) remained stable among residents of census tracts in the lowest three quartiles of vaccine coverage (vaccine coverage range: 18% - 77%) but declined to 38% (95% CI: 13, 68) among individuals residing in census tracts in the highest quartile of vaccine coverage, adjusted for sociodemographic factors. For the entire population aged 0 to 9 years, direct vaccine effectiveness for DTaP vaccine comparing vaccinated time at risk to unvaccinated time at risk was 71% (95% CI: 64%, 77%).

To assess population-level vaccine protective effects (Table 5.5), we selected lowest quartile of vaccine coverage (69% or lower) as the reference category and compared pertussis incidence in this group with census tracts in the highest quartile of vaccine coverage (77% or higher). In the age and NSES adjusted models, the indirect effects for 3 or more DTaP doses were estimated to be 45% (95% CI: 1%, 70%). Total protection was 93.9% (95% CI: 91%, 96%) and overall protection was 42% (95% CI: 19%, 60%). Thus, we found evidence for population-level effects after 3 doses of DTaP vaccine among children 7 months of age and older. The results of sensitivity analysis using vaccination coverage estimates as of 31st December, 2015 were similar to our main results (Supple-

mentary Materials 5.7.1 Table 1).

5.4 Discussion

Estimating and reporting population-level effects of vaccines can help us understand the impact of a vaccine beyond its protective effects in vaccinated individuals and can improve the cost-effectiveness and public health utility of routine vaccination programs. This is particularly important for the DTaP vaccine where there is a growing concern that the vaccine does not provide herd effects. Our study describes how an observational study design can be used in a setting where a vaccine is administered as part of an ongoing, routine immunization program to assess population-level effects of the vaccine. We estimated indirect, total, and overall vaccine effects for the DTaP vaccine and found that three doses of the DTaP vaccine provides statistically significant indirect, total, and overall protection.

Few other studies have directly assessed population-level vaccine effects for pertussis. Trollfors et al. (1998) found evidence of indirect protection among parents of pertussis vaccinated children (but not in siblings) in a double-blind place-controlled randomized trial of an acellular pertussis vaccine [**?**]. However, results from a randomized controlled trial may not be generalizable to routine immunization programs [145]. Preziosi et al. (2002) estimated overall vaccine effectiveness for pertussis vaccine in a prospective cohort study, but could not measure indirect and total vaccine effects because they did not have vaccination status of study participants [141]. Preziosi and Halloran (2003) showed that pertussis vaccines reduce infectiousness among vaccinated individuals and concluded that the vaccine might be able to provide indirect effects, but only 6% of the cases in their cohort study had received an acellular pertussis vaccine [19]. Our analysis supports the findings of these studies and additionally delineates the different population-level vaccine effects in a population exclusively vaccinated with an acellular pertussis vaccine in a non-randomized manner.

As expected, pertussis incidence was highest among children with zero vaccine doses reported compared to children who were partially or fully vaccinated across all levels of vaccine coverage (Table 5.2). Surprisingly, we found statistically significant decrease in pertussis incidence with increasing cluster-level vaccine coverage for partially-vaccinated children, but not for unvaccinated children. We had expected a strong inverse rela-

tionship between pertussis incidence in unvaccinated children and cluster-level vaccine coverage because unvaccinated children living in areas of high vaccination would likely benefit the most from any vaccine herd effects. The reason for lack of statistical significance in this group could be a small number of cases and a much shorter follow-up time for children who had zero DTaP vaccine doses reported (median follow-up time of 513 days compared to 1700 days for fully- vaccinated and 1500 days for partially-vaccinated children, Table 5.2).

Direct vaccine effect in our study was $\sim 75\%$ in the bottom three quartiles of vaccine coverage but dropped in the highest vaccine coverage quartile. This finding is consistent with results from field trials of oral cholera vaccines that estimated variation in direct vaccine effects by area-level vaccine coverage [147] [148] [149]. This brings to light the implications of estimating direct vaccine effects in areas of high vaccine coverage for a vaccine that also provides indirect protection. Because disease transmission is generally low in areas of high vaccination coverage, the incidence rate in the under-vaccinated group is lower when vaccines also provide indirect effects, possibly distorting the vaccine effectiveness estimates in high coverage areas [150]. This result underscores the importance of considering underlying vaccination coverage in vaccine effectiveness studies for pertussis.

We defined cluster-level vaccine coverage as proportion of children 3 months to 9 years old who were age-appropriately vaccinated with DTaP doses as measured on 31st December, 2017. We chose to measure vaccine coverage this way was because the DTaP vaccine is part of a long-standing routine immunization program, and the start date of vaccination was different for different cohort members. Cluster level vaccine coverage can be measured in several different ways and is likely to be time-varying. For example, another way of estimating vaccine coverage could be proportion of 19-35 month old children with ≥ 4 DTaP doses in a given year or 4-6 year old children with ≥ 5 DTaP doses in a given year, or proportion of children with age-appropriate number of doses measured at the end of a different year instead of 2017. Differences in the way vaccine coverage is measured will affect estimates of population-level effects. To check this, we performed a sensitivity analysis with vaccination coverage as measured on 31st December, 2015. While the effect estimates were not different from the main analysis, not accounting for changing vaccination coverage is a limitation of this study.

This study is observational at two levels, both at the cluster-level vaccination coverage

and at the individual vaccination status. Thus, it is vulnerable to confounding bias due to systematic differences between the fully-vaccinated and under-vaccinated study participants. To control for confounding by measured covariates, we adjusted our Cox proportional hazards models for individual and cluster-level confounders such as birth cohort and neighborhood level SES. However, there still might be unmeasured confounders of the relationship between vaccination and pertussis incidence that we have not considered in this study, which could lead to potentially biased estimates of vaccine effects.

Other methods for confounder adjustment are available such as the Inverse Probability of Treatment Weighting (IPTW) using propensity scores [151]. Particularly in the context of vaccine effectiveness studies, Halloran and Hudgens (2008) have defined causal estimands for direct, indirect, total, and overall vaccine effects and developed unbiased estimators of the causal estimands for the four effects in a two-staged randomized design [145]. Perez-Heydrich et al. (2014) demonstrated the use IPTW methods using propensity scores to estimate the four effects from an individually-randomized cholera vaccination study [149]. However, these methods have not been developed for vaccines given as part of a routine immunization program in a non-randomized manner or for situations where vaccine coverage can be time-varying and vaccination effect is time-dependent. This can be a potential area for further research.

Census tracts were chosen as the unit of analysis in this study for two main reasons: Information on confounders was available only at the census tract level and census tracts are stable administrative units whose boundaries do not change over time. However, there are some drawbacks of using census tracts as clusters. Not all census tracts had one or more reported cases. The census tracts in this study are of varying sizes, depending on the number of people that reside in them, and some census tracts may be too large to detect any herd effects. In other studies, clusters were created using GIS-based methods [147] [150] [152] [149]. Sizes of GIS-based clusters can be optimized for number of cases in clusters and variation in vaccination coverage [147] [150]. An improvement in our analysis could be creating clusters of equal area around each case and non-case using GIS methods where the size of the cluster is optimized statistically. However, without detailed demographic data on confounding variables at the individual level, we may not be able to properly account for all known confounders in such an analysis. Because this is an observational study, adjusting for confounders is crucial. Despite not optimizing the size of clusters, we still found evidence for indirect vaccine effects (Table 5.5) but

the 95% confidence intervals were quite wide, likely because indirect effects were diluted due to the large size of some census tracts. We assumed no inter-cluster transmission occurs, but it is possible that children travel to other census tracts for schools or day care, especially in the metropolitan areas. This could further attenuate indirect effects.

For direct vaccine effects, our cohort consisted of children 3 months or older and vaccination status was time-dependent. However, we estimated population-level vaccine effects in a cohort 7 months or older who were either age-appropriately vaccinated with 3 doses of DTaP (exposed) or were under-vaccinated (unexposed), i.e., vaccination status was fixed at start of follow-up. This was because no methods have been developed yet for estimating herd effects of a vaccine for a time-varying exposure. We decided to follow-up children after age 7 months because they are supposed to be vaccinated with the 3-dose DTaP primary series by age 7 months.

Use of low vaccine coverage clusters instead of unvaccinated or placebo clusters probably underestimated the population-level effects in this study. Even in the lowest vaccine coverage quartile, mean census-tract level vaccination coverage was 65%, which is fairly high.

In summary, we estimated the direct, indirect, total, overall, vaccine effects of pertussis vaccination administered as part of a routine program in a highly vaccinated population. We found evidence of vaccine herd effects in children who had received at least 3 doses of the DTaP vaccine in an age-appropriate manner . Studies of DTaP vaccine effectiveness should consider estimating and reporting population-level vaccine effects to present a complete picture of the public health benefits of acellular vaccines in preventing pertussis.

5.5 Tables

Table 5.1: Different vaccine effects, study population, and exposure categories

Vaccine Effect	**Populations**	**Exposure Categories**
Direct Effect	Entire study population	Fully-vaccinated vs. Under-vaccinated children
Indirect Effect	Under-vaccinated children	Clusters in highest vs. lowest quartiles of vaccine coverage
Total Effect	Fully-vaccinated vs. under-vaccinated children	Clusters in highest vs. lowest quartiles of vaccine coverage
Overall Effect	Fully-vaccinated vs. under-vaccinated children	Clusters in highest vs. lowest quartiles of vaccine coverage

Table 5.2: Baseline characteristics of participants older than 3 months of age

Characteristics	**Fully-vaccinated (N=227,751)**	**Partially-vaccinated (N=67,499)**	**Unvaccinated (N=15,148)**
Age (Years)	5 (2.7)*	5 (2.8)	5 (2.9)
Females	111058 (48.7)	32974 (48.8)	7446 (49.1)
Race			
White	117926 (51.7)	33980 (50.0)	7284 (48.1)
Asian	34220 (15.0)	9173 (13.5)	2126 (14.0)
Black	21404 (9.4)	7795 (11.5)	1559 (10.3)
Native American	27 (0.001)	12 (0.001)	2 (0.001)
Other/NA/Unknown	54174 (23.8)	16539 (24.5)	4177 (27.5)
Birth Cohort			
2008	24381 (10.7)	8915 (13.2)	1850 (12.2)
2009	23914 (10.5)	8671 (12.8)	1673 (11.0)
2010	23891 (10.5)	7936 (11.7)	1527 (10.1)
2011	23698 (10.4)	7615 (11.3)	1449 (9.5)
2012	23698 (10.4)	7876 (11.6)	1117 (7.3)
2013	21399 (9.4)	8806 (13.0)	1343 (8.8)
2014	25794 (11.3)	3728 (5.5)	1589 (10.5)
2015	23766 (10.4)	4502 (6.6)	1592 (10.5)
2016	21920 (9.6)	5509 (8.1)	1683 (11.1)
2017	15290 (6.7)	3941 (5.8)	1325 (8.7)
NSES quintiles			
1	45434 (19.9)	14779 (21.8)	3277 (21.6)
2	43581 (19.1)	13708 (20.3)	2955 (19.5)
3	46210 (20.3)	13315 (19.7)	3028 (20.0)
4	45085 (19.8)	12793 (18.9)	3032 (20.0)
5	46749 (20.5)	12683 (18.8)	2796 (18.5)
Pertussis incidence per 10,000	6.37	13.48	73.28
Follow-up days, median (IQR)	1796 (1758)	1315 (1480)	513 (549)

* Standard Deviation for age, column percentages for the rest of the categorical variables, and interquartile r
follow-up days

Table 5.3: Trends in pertussis incidence by vaccination status and quantiles of vaccine coverage clusters

	Population	**Cases**	**Person-time (months)**	**Incidence rate per 100,000**
Unvaccinated				
Vaccine coverage				
$< 69.3\%$	5053	41	135805.54	3.02
69.3% to 73.9%	3920	28	107656.30	2.60
73.9% to 77.4%	3566	28	99973.05	2.80
$\geq 77.4\%$	2498	14	68096.16	2.06
Total	15037	111	411531.05	2.70
Cochrane-Armitage trend test (two-sided)				p=0.3
Fully-vaccinated				
$< 69.3\%$	50985	28	2883030.20	0.10
69.3% to 73.9%	55056	37	3245437.48	0.11
73.9% to 77.4%	59102	34	3573348.72	0.10
$\geq 77.4\%$	62463	46	3883412.56	0.12
Total	227606	145	13585228.95	0.11
Cochrane-Armitage trend test (two-sided)				p=0.3
Partially-vaccinated				
$< 69.3\%$	21472	39	1088518.52	0.36
69.3% to 73.9%	17652	23	885257.97	0.26
73.9% to 77.4%	15519	20	783841.90	0.26
$\geq 77.4\%$	12765	9	654884.56	0.14
Total	67408	91	3412502.95	0.27
Cochrane-Armitage trend test (two-sided)				p= 0.009

Table 5.4: Direct vaccine effects by level of vaccine coverage

Vaccine Coverage	**Population**	**Fully-vaccinated**		**Under-vaccinated**		**Crude** VE^a	**95% CI**	**Adjusted** $VE^{a,b}$	**95% CI**
		Cases	PT at risk	Cases	PT at risk				
$< 69.3\%$	77618	28	2333885	80	1609071	75	(62, 84)	76	(63, 84)
69.3% to 73.9%	76716	37	2671443	51	1407400	75	(59, 84)	73	(57, 83)
73.9% to 77.4%	78269	34	3029361	48	1267281	75	(61, 84)	75	(61, 84)
$\geq 77.4\%$	77795	46	3356042	23	1093739	48	(15, 69)	47	(13, 68)
Total	310398	145	11390731	202	5377491	72	(65, 77)	71	(64, 77)

[a] Cox Proportional Hazards models used to estimate crude and adjusted vaccine effects

[b] Cox PH models adjusted for birth cohort and neighborhood socioeconomic status

Table 5.5: Population-level vaccine effectiveness measures.

Measure of Protection	**Low vaccine coverage quartile**		**High vaccine coverage quartile**		**Crude VE^a**		**Adjusted VE^b**	
	Subjects	**Cases**	**Subjects**	**Cases**	**VE**	**95% CI**	**VE**	**95% CI**
Indirect	9361	55	4436	16	42	(-2, 67)	45.0	(1, 70)
Total	9361	55	69101	45	94	(91, 96)	93.9	(91, 96)
Overall	72496	95	73537	61	44	(22, 59)	42.2	(19, 60)

[a] Cox proportional hazards models used to estimate crude and adjusted vaccine effects

[b] Cox PH models adjusted for birth cohort and neighborhood socioeconomic status. Indirect effects model additionally adjusted for number of vaccine doses received.

5.6 Figures

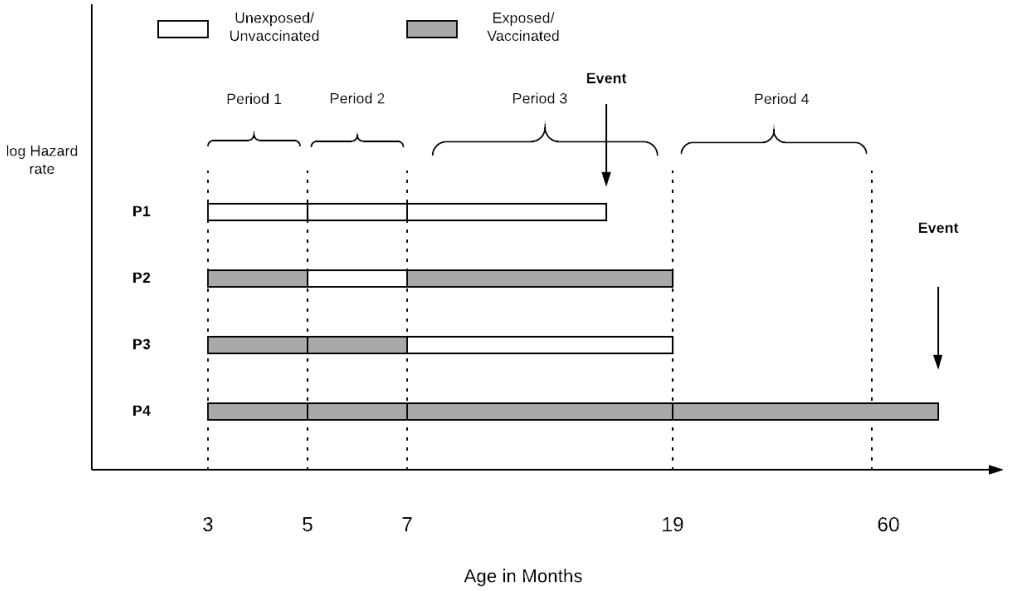


Figure 5.1: Time-dependent vaccination/ exposure status of the Diphtheria-Tetanus-acellular-Pertussis vaccine for estimating vaccine direct effects

5.7 Supplementary Material

5.7.1 Supplementary Tables

Table 1: Population-level vaccine effectiveness measures.

Measure of Protection	**Low vaccine coverage quartile**		**High vaccine coverage quartile**		**Crude ?VE^a**		**Adjusted VE^b**	
	Subjects	**Cases**	**Subjects**	**Cases**	**VE**	**95% CI**	**VE**	**95% CI**
Indirect	9494	50	4489	16	34	(-17, 62)	41	(-4, 67)
Total	9494	50	69690	38	94	(89, 96)	94.2	(90.7, 96.4)
Overall	74335	93	74179	54	48	(27, 63)	46	(21,63)

[a] Cox proportional hazards models used to estimate crude and adjusted vaccine effects

[b] Cox PH models adjusted for birth cohort, gender, neighborhood socioeconomic status, and insurance. Indirect effects model additionally adjusted for number of vaccine doses received.

5.7.2 Supplementary Figures

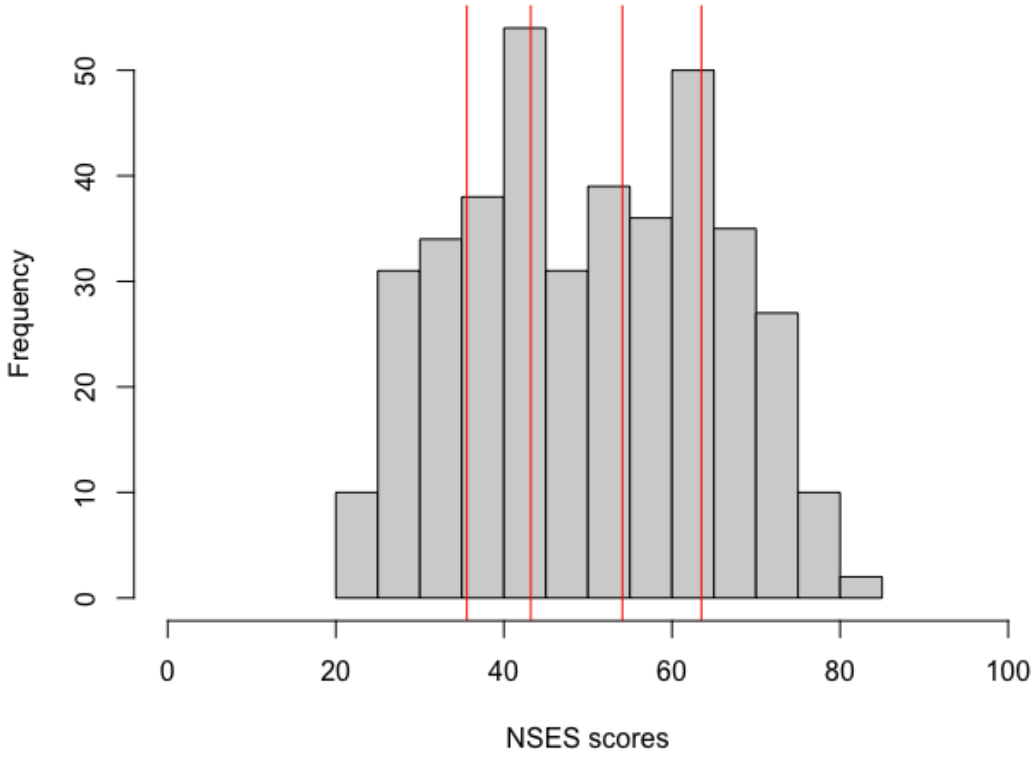


Figure 1: Distribution of neighborhood level socio-economic status (NSES) scores

NSES scores were calculated at census tract level. Data was divided into quintiles. Red lines indicate the quintile boundaries.

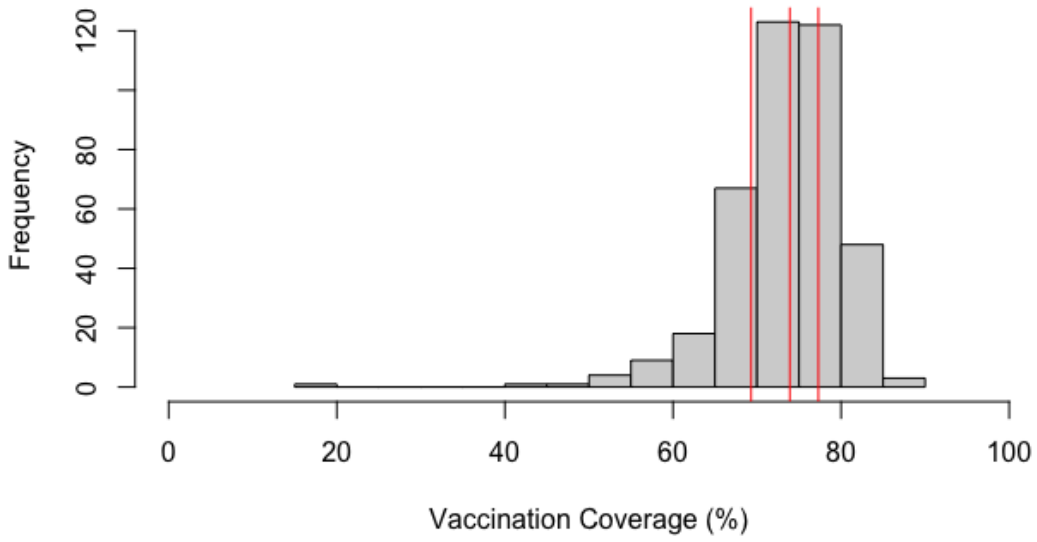


Figure 2: Distribution of census-tract level vaccine coverage

Data was divided into quartiles. Red lines indicate the quartile boundaries.

Chapter 6

Estimating waning effects of acellular pertussis vaccine following 5 doses of childhood vaccine series

Abstract

Background: Waning of immunity after vaccination with the Diphtheria-Tetanus-acellular-Pertussis vaccine (DTaP) vaccine has been proposed as one of the main reasons for pertussis resurgence in the US. In this study, we estimated waning in vaccine effectiveness after 5 doses of DTaP vaccine.

Methods: We conducted a retrospective cohort study among children 5-9 years old born or living in King County, Washington betweem 2008 and 2012 and who participated in the Washington State Immunization Information System. We estimated waning of vaccine effectiveness after 5 DTaP using Schoenfeld residuals obtained from fitting Cox proportional hazards models to the data.

Results: There were 55 pertussis cases in this cohort, of whom 22 (40%) were fully-vaccinated and 33 (60%) were under-vaccinated. Vaccine effectiveness remained high for up to 4 years after the fifth dose (VE= 0.83; 95% CI: 0.95, 0.39 in year 4). The association between scaled Schoenfeld residuals and time since vaccination was not sta-

tistically signifcant (p-value for hypothesis test for departure from proportional hazards assumption: 0.44).

Conclusion: We did not find evidence for waning of vaccine effectiveness after 5 doses of DTaP among 5 to 9 year old children in King County, WA.

6.1 Introduction

Despite availability of a highly effective vaccine since the 1940s, we continue to see pertussis outbreaks periodically even in countries with high vaccination coverage [5]. The WHO reported 151,074 pertussis cases globally in 2018 and 89,000 estimated deaths in 2008 [153]. A modeling study estimated an even larger death toll of 160,700 deaths among children under 5 years of age using data from 2014 [2]. Since the 1970s, US has experienced an increase in pertussis activity [7]. While incidence remains highest among infants, it has been steadily increasing among school children, adolescents, and adults [154] [155] [17]. In 2012, reported pertussis cases in the US peaked at 48,277, which was the largest number of cases reported since 1955 [32].

One popular driver of pertussis resurgence is hypothesized to be waning of immunity after vaccination with Diphtheria-Tetanus-acellular Pertussis vaccine (DTaP) [156] [157], which replaced the Diphtheria-Tetanus-whole-cell Pertussis (DTwP) vaccine in high-income countries after concerns about safety and reactogenecity of DTwP vaccines. However, estimates of duration of DTaP vaccine-derived immunity vary widely between studies. Several epidemiologic studies conducted in highly immunized populations concluded that immunity due to acellular pertussis vaccines wanes within a few years and have called for a vaccine with more long-lasting protection [158] [159] [160] [161] [113]. A systematic review of duration of immunity due to pertussis vaccines found that DTaP-induced immunity wanes after 4 – 12 years [24], while a meta-analysis concluded that only 10% of children vaccinated by DTaP would be protected over 8 years [25] . In contrast, modeling studies have found evidence of long-lasting, slowly waning immunity lasting up to 70 years using population-level longitudinal time series data [21] [162] [163].

We sought to estimate waning in DTaP vaccine effectiveness in a population of 5-9 year old children living in King County, Washington and registered in the Washington State Immunization Information System (WA-IIS) using survival analysis methods. Specifi-

cally, we used smoothed Schoenfeld residuals to non-parametrically estimate waning of vaccine effectiveness [164] [165].

6.2 Methods

6.2.1 Pertussis cases

Pertussis cases born between 1st January, 2008, and 31st December, 2012, (ages 5 to 9 years) and reported between 1st January, 2013, and 31st December, 2017, were obtained from the Public Health Seattle and King County (PHSKC) surveillance database. The clinical case definition of pertussis used is a cough illness lasting $\geq$ 2 weeks with at least one of the following: paroxysms of coughing or inspiratory "whoop" or post-tussive vomiting, or apnea (with or without cyanosis) for infants up to 1 year of age [45]. Cases are classified as suspected, probable, or confirmed based on Centers for Disease Control and Prevention (CDC) case definitions of pertussis [45]. Pertussis cases in this study include only ones that were symptomatic, sought medical help, and were ultimately reported to the PHSKC surveillance system. Surveillance data included information on cases' demographic information (such as age, sex, race, ethnicity, home address), clinical symptoms, pertussis vaccination history, exposure and transmission settings, pertussis treatment received, and epidemiologically-linked contacts.

6.2.2 Study cohort

The study was conducted within a cohort of children registered in the Washington State Immunization Information System (WA-IIS), which is a lifetime registry that tracks immunization records for people of all ages in Washington State [95]. Healthcare providers such as primary care physicians, hospitals, and healthcare plans voluntarily report patient immunizations to WA-IIS. Additionally, birth certificates of children born in King County are loaded into the registry periodically. For this analysis, study cohort was restricted to children born in King County from 1st January, 2008, and 31st December, 2012, (children between the ages of 5 and 9 years) followed up from 1st January, 2013, to 31st December, 2017. Thus, children born in 2008 were followed up for 5 years (from ages 5 to 9), children born in 2009 were followed up for 4 years (from ages 5 to 8), and so on. For recent years, WA-IIS has captured nearly all children born or living in King County through birth certificate data collected twice a month. Ninety-nine percent of

children aged 4 months- 5 years have 2 or more immunizations recorded in the WA-IIS [95]. Using WA-IIS data, we created a retrospective cohort of 134,950 children 5 to 9 years of age, born or living in King County, WA, between 1st January, 2008, and 31st January, 2012, with a unique identifier for each child. Vaccination information including vaccine name and date of receipt, vaccine manufacturer, and healthcare facility of administration was requested for all pediatric vaccines recommended from birth to 9 years of age for each child. Demographic information in the WA-IIS included date of birth, sex, race, ethnicity, current residential address and county, and insurance information.

Census tract level neighborhood socioeconomic score (NSES) for each participant was estimated using the 2010 US census data [96]. In short, NSES score is an index score ranging from 0 to 100 created using measures of percent of adults 25 years and older with a high school degree (education), percent of civilian population 16 years and older with professional/managerial/executive occupations (employment), median household income (income), percent of families above the Federal Poverty Line (income), and percent of households with children under the age of 18 that are "female-headed" (no male present) [123]. The score was divided into quintiles, where the first quintile is the group with the lowest NSES (20th percentile or lower) and fifth quintile is the group with the highest NSES (higher than 80th percentile).

6.2.3 Linking surveillance and immunization datasets

Immunization records from the WA-IIS and PHSKC surveillance database were linked based on a probabilistic matching algorithm that used participants' first name, last name, date of birth, sex, and city of residence [125]. Matching was performed using the `fastLink` package in R [126]. First and last names were linked using string distance matching and partial matches were allowed for these fields. All pairs that were matched with a posterior probability of > 0.85 were retained for further review and false matches were removed manually. Unmatched records were reviewed and matched manually based on additional data in the surveillance and immunization registry. After matching was completed, all personal identifiers were removed. Flowchart in Supplementary material (4.7.3 Fig. 1) describes the data linkage process.

6.2.4 Censoring of registry participants

Follow-up time for participants began at age 5 years. Follow-up continued until they were diagnosed with pertussis, died, moved to another county or changed healthcare providers (entered in the registry as 'moved or gone elsewhere' or MOGE), or until the end of follow-up period on 12/31/2017, whichever came first. As participants were not actively followed as in a prospective cohort study, vaccination dates for pediatric vaccines and registry record update dates were considered proxies for continued enrollment in the WA-IIS and residence in King County. Flowchart in Supplementary Material (4.7.3 Fig. 2) describes details of censoring.

6.2.5 Study design

We used a retrospective cohort study to estimate waning of DTaP vaccine effects. The CDC Advisory Committee for Immunization Practice (ACIP) recommendations for childhood DTaP doses in this age group were used to determine vaccination status [68]. Five doses of DTaP are recommended at ages 2, 4, and 6 months, 15-18 months, and 4-6 years. The main exposure of interest was DTaP vaccination categorized as fully-vaccinated (received 5+ doses by age 5 years) and under-vaccinated (received fewer than 5 doses by age 5 years). Participants may or may not have received all 5 doses in a timely manner but were considered fully-vaccinated as long as they received 5 or more doses by age 5 years. While ACIP recommendations allow dose 5 to be administered up to 6 years of age, we used a cut-off of age 5 years for dose 5 in this study because most children start pre-school by this age. We assumed that if a child had no record of a DTaP dose in the WA-IIS, they did not receive it. Outcome was all suspected, probable, and confirmed pertussis cases as described above. We decided to only include cases diagnosed 15 days after vaccination because it might take upto 15 days for the vaccine to show any protective effect [160]. Ultimately, no cases had to be excluded due to this criterion because all cases occurred at least 15 days after last DTaP vaccination.

6.2.6 Statistical analysis

We used survival analysis methods to estimate waning of vaccine effectiveness (VE) using Schoenfeld residuals from fitting Cox proportional hazards models. This method has been used to estimate waning effects of cholera vaccines [164] [26]. Vaccine efficacy for susceptibility to pertussis disease is denoted by VE_s and is the measure of how protective

the vaccine is against disease [13]. VE_s based on hazard rate ($VE_{s,\lambda}$) and VE_s based on cases per person-time at risk ($VE_{s,IR}$) are the preferred measures of VE in this case because they allow for different follow-up times [13].

We first constructed a diagnostic log-minus-log plot of Kaplan-Meier estimates of survival curves for the fully vaccinated and under-vaccinated groups to check the proportionality of hazards assumption. Vaccination status was defined as fully-vaccinated with 5 doses of DTaP vaccine at age 5 or under-vaccinated (fewer than 5 doses of DTaP vaccine at age 5). We fit an ordinary Cox proportional hazards (PH) model to the data to estimate the hazard ratio of pertussis comparing fully-vaccinated children to under-vaccinated children, adjusted for age and NSES index [166], and computed the scaled Schoenfeld residuals. Schoenfeld residuals are the scaled differences between actual and computed covariate values at each event time [167]. The scaled Schoenfeld residuals were then added to the model coefficient of vaccination status from the Cox PH model. The sum of model coefficient and Schoenfeld residuals was smoothed over time to obtain the time-varying regression coefficient, $\beta(t)$ for vaccination status. Smoothed vaccine efficacy estimate is given by $VE_{s,\lambda}(t) = 1 - HR(t) = 1 - e^{\beta(t)}$. Hypothesis test for departure from proportional hazards assumption was conducted, where the null hypothesis was $H_0 : \beta(t) = \beta$ for all times t [168]. The analysis was done using packages `survminer` [169], `survival` [170], and `kyotil` [171] in R Studio version 3.4.1 [66].

We fit the following Cox proportional hazards model:

$$h(t) = h_0(t) * \exp(\beta_1 * Vacc.status + \beta_{4.1-4.4} * SES + \beta_5 * Age) \tag{6.1}$$

$h(t)$: Hazard of pertussis at time t
$h_0(t)$: Baseline hazard function at time t, i.e. the value of hazard when all covariates are set to zero
$\exp(\beta_1)$: Hazard ratio of pertussis comparing vaccinated and under-vaccinated individuals, keeping all other covariates constant
$\exp(\beta_{2.1-2.4})$: Hazard ratio of pertussis comparing children living in census tracts with NSES score quintiles 2,3,4,5 to children living in census tracts with NSES score in quintile 1, respectively.
$\exp(\beta_3)$: Hazard ratio of pertussis per unit increase in age (in years)

For comparison, we also estimated piecewise constant $VE_{s,IR}(t)$ by partitioning the data

into years since vaccination and fitting a separate Poisson regression model for each year as follows:

$$\log(Y) = \alpha + \beta_1 * Vacc.status + \beta_{2.1-2.4} * NSES + \beta_3 * Age + \log(Time - at - risk) + \epsilon \tag{6.2}$$

where Y is the pertussis case count in each year since start of follow-up and person time at risk is measured in months.

6.3 Results

The study included 134,950 children 5 years and older living in King County, WA, and registered in the WA-IIS who were followed between 1/1/2013 and 12/31/2017. The median duration of follow-up was 906 days (range: 2 – 1827 days). A total of 5,333 (3.9%) participants were lost to follow up. Of the WA-IIS participants, 1,755 (1.3%) had no recorded doses of DTaP (vaccine-naïve) and 21,808 (16.1%) were partially vaccinated. Of those partially vaccinated, 2,392 (11%) received 1 dose, 1,463 (6.7%) received 2 doses, 2,579 (11.8%) received 3 doses and 15,374 (70.5%) received 4 doses of DTaP by age 5 years. Under-vaccination decreased in successive birth cohorts (Fig. 6.1).

There were 55 pertussis cases in this cohort. Of these, 22 (40%) were fully-vaccinated and 33 (60%) were under-vaccinated. Of the 33 under-vaccinated cases, 19 (57%) had no DTaP vaccine doses recoded in the WA-IIS.

Fig. 6.2 shows the results of the log-minus-log plots of Kalpan-Meier estimates of the survival curves for the DTaP fully vaccinated and under-vaccinated participants. Separation between the curves indicates that being fully vaccinated with 5 doses of DTaP protects better against pertussis. The survival curves are parallel to each other which is consistent with leaky mode of action for DTaP vaccine [172]. However, as pertussis is a rare disease in this population and incidence rates are low, the survival curves would look similar whether the mode of action of DTaP vaccin is leaky or all-or-none [13]. It appears that the curves slowly approach each other at the end of 5 years, but this is difficult to infer based on cumulative incidence.

Fig. 6.3 shows non-parametrically smoothed plots of $VE_{s,\lambda}$(t) for the 5-dose DTaP vaccine series estimated using Schoenfeld residuals after fitting a Cox PH model to the data. Table 6.1 gives the efficacy estimates and approximate 95% confidence intervals

for roughly six month time intervals throughout the follow-up period. Testing for a linear association between scaled Schoenfeld residuals and time showed no significant time-dependent waning effects (individual test for vaccination status, p:0.44) (Supplementary Material 6.7.2 Fig. 1). This indicates that there is no statistically significant waning of vaccine effects for the DTaP vaccine. Vaccine efficacy remains high for up to 4 years after the fifth dose, with a $VE_{s,\lambda}$(t) of 0.83 (95% CI: 0.95, 0.39) in the fourth year. However, $VE_{s,\lambda}$(t) estimates for the fifth year are not statistically significant. A downward curve and large 95% confidence intervals towards the end of the study are due to very few events occurring among children who contributed more than 4 years of follow-up.

In the piecewise constant Poisson regression analysis, we again observe high vaccine efficacy for up to four years after vaccination with no evidence of waning, and $VE_{s,IR}(t)$ estimate for the fifth year is not significant (Supp Material 6.7.1 Table 1).

6.4 Discussion

We used survival methods to estimate waning vaccine effects after the fifth dose of DTaP vaccine among children aged 5-9 years and living in King County, WA. We found that contrary to other studies, vaccine efficacy after 5 doses of DTaP in this age group is high and does not wane with time for up to 4 years after vaccination.

Our estimate of vaccine effectiveness between 80% and 90% for 5 doses of DTaP in the four years of follow-up comparing vaccinated to under-vaccinated children is similar to other studies [83]. There were only 55 cases reported during the study period and incidence in the vaccinated group was low, which likely resulted in high $VE_{s,\lambda}$(t) estimates in year 1 and year 4 post vaccination when data was smoothed. Using smoothed Schoenfeld residuals to estimate waning vaccine effects might be more appropriate when number of events is sufficiently high [165].

Several modeling studies have shown that current pertussis dynamics can be explained by long-lasting DTaP immunity. One study estimated a 10% risk of vaccine protection waning within 10 years and a 55% chance of lifelong protection in vaccinated individuals in Massachusetts [22]. Studies that used age-structured pertussis incidence data from Thailand and Sweden also found evidence for long-lasting protection of DTaP vaccine

at the population scale [17] [163]. While our observational study could only estimate effects of waning for up to 4 years post-vaccination, it supports the overall conclusions of these modeling studies that DTaP immunity does not wane rapidly.

In stark contrast to these modeling studies and our results, some epidemiological studies that used case-control and retrospective cohort study designs have claimed that despite providing good protection against pertussis, immunity from DTaP vaccines wanes steadily between 27% to 42% annually [159] [160] [113] [173] [174] [138]. However, there could be issues with how these studies interpret their results with respect to speed of waning efficacy. Klein et al. estimated that risk of pertussis increased by 42% for every year since vaccination of fifth dose, interpreting the annual increase in odds of pertussis as rapid decrease of vaccine effectiveness [113]. Simply interpreting increasing yearly odds of pertussis as rapid loss of vaccine-derived immunity does not take into account the complex transmission dynamics of pertussis. [162] [175]. One study did not include a vaccine naïve or under-vaccinated comparison group in their estimation of waning VE and simply interpreted an increase in risk of pertussis with time since vaccination in fully vaccinated individuals as strong evidence of waning [160]. In our study, we tested the hypothesis that DTaP vaccine effects vary with time instead of interpreting increase in pertussis risk as rapidly waning immunity. Furthermore, there were only 55 pertussis cases in a cohort of over 130,000 children. If indeed protection after the fifth dose waned rapidly, we would expect to see many more cases of pertussis given how highly transmissible it is [162].

Our study has limitations. There were only 55 cases reported in this cohort which may have limited our ability to reliably estimate smoothed $VE_{s,\lambda}(\mathrm{t})$. Our comparison group consisted of partially vaccinated children and not just vaccine-naïve children. Most of the partially vaccinated children had received at least 4 doses of DTaP vaccine, and hence had some protection against pertussis. Including partially vaccinated children in the comparison group likely underestimated the $VE_{s,\lambda}(\mathrm{t})$ estimates. One could recategorize vaccination groups as unvaccinated, partially vaccinated, and fully vaccinated, but estimation of $VE_{s,\lambda}(\mathrm{t})$ might be difficult given the small number of cases. We were not able to follow-up all children in the cohort for the entire 5 years post vaccination because we did not have WA-IIS records for birth cohorts before 2008 or after 2017.

To summarize, we found that the 5-dose DTaP vaccine series provides good protection against pertussis disease and contrary to oft cited evidence, vaccine effectiveness does

not wane rapidly. Our results suggest that changing current vaccination schedule to delay dose 5 or administer booster doses before age 10 might not be necessary as some studies have proposed [138] [158].

6.5 Tables

Table 6.1: Estimated vaccine efficacy over time $VE_{s,\lambda}$(t) with 95% confidence interval (CI) for 5-dose series of DTaP vaccine among 5-9 year old children living in King County, Washington, followed from 1st, January, 2013, through 31st December, 2017.

Time since vaccination (months)	$VE_{s,\lambda}$	**95% CI**
5.5	0.90	0.96, 0.76
12.1	0.79	0.91, 0.51
17.8	0.76	0.91, 0.41
24.5	0.83	0.93, 0.63
30.2	0.89	0.96, 0.71
35.8	0.90	0.97, 0.71
42.6	0.89	0.97, 0.64
48.2	0.83	0.95, 0.39
53.8	0.69	0.95,-0.85
60	0.23	0.95, -11.04

Cox proportional hazards models are adjusted for age and neighborhood SES (NSES) index.

6.6 Figures

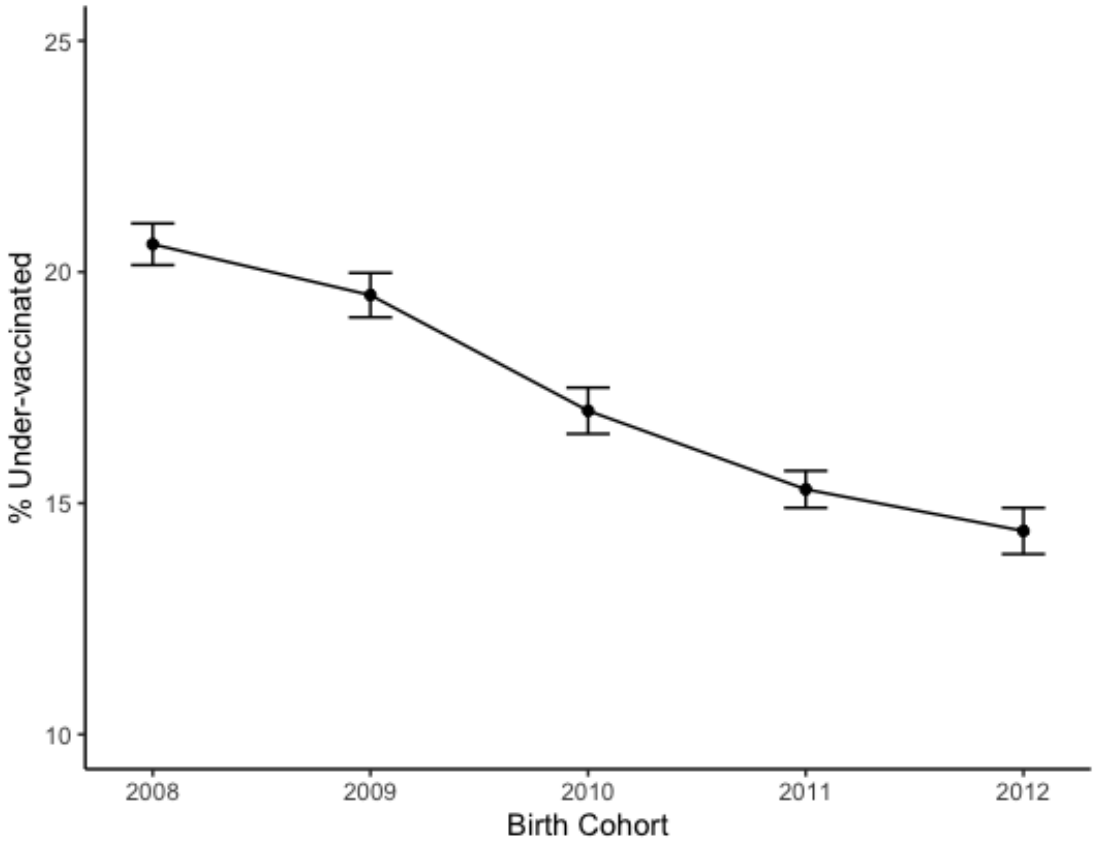


Figure 6.1: DTaP under-vaccination (proportion of WA-IIS participants who received fewer than 5 doses at age 5) by birth cohort and 95% CI in 5-9 year old residents of King County, WA between 1st January, 2013, and 31st December, 2017

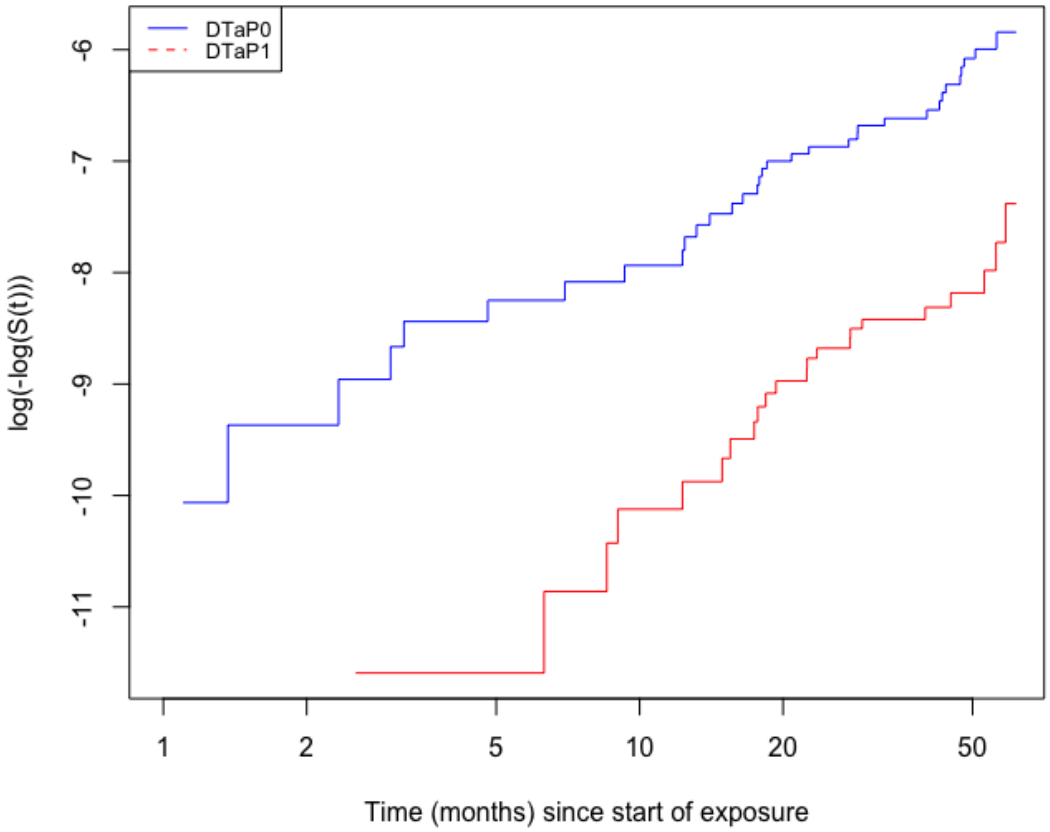


Figure 6.2: log-minus-log plots of Kalpan-Meier estimates of the survival curves for the DTaP fully vaccinated and under-vaccinated participants living in King County between 1st January, 2013, and 31st December, 2017.

Survival curve for vaccinated individuals is in red, under-vaccinated is in blue. The two curves are nearly parallel. Exposure here is being fully-vaccinated or under-vaccinated at age 5 years

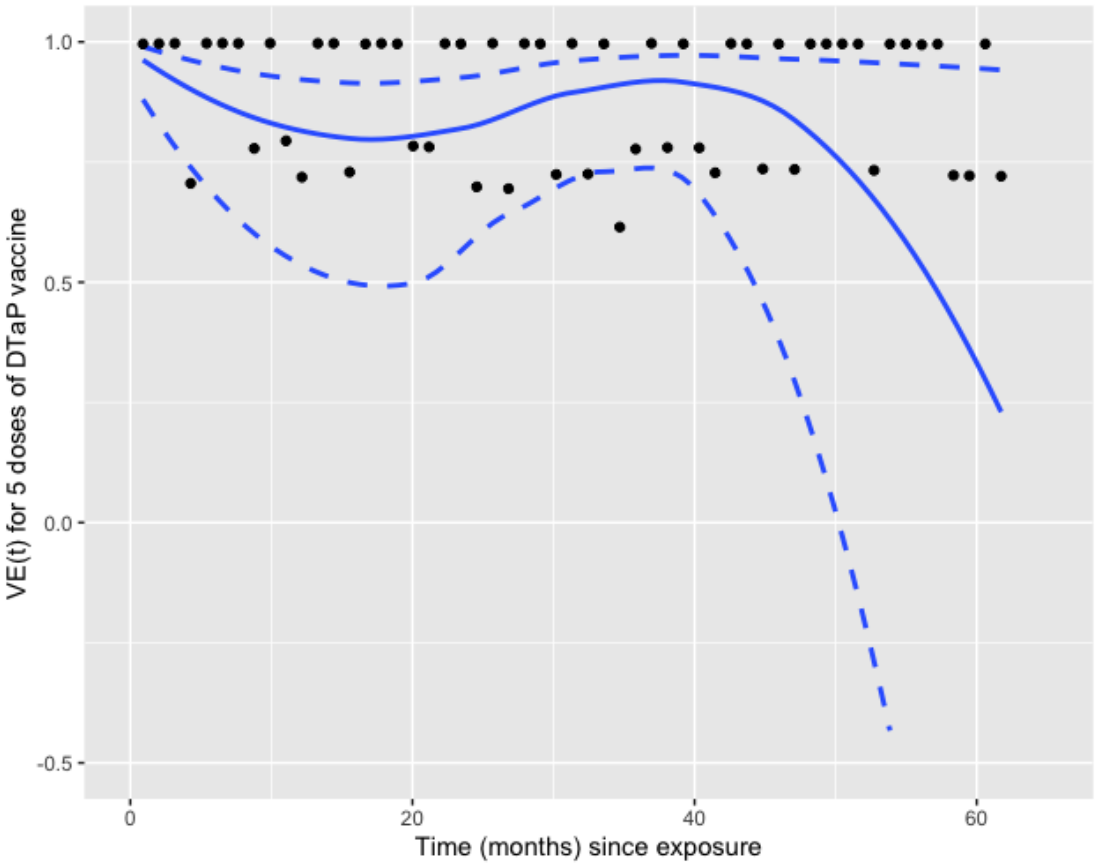


Figure 6.3: Non-parametric smoothed plots of $VE_{s,\lambda}$(t) (solid line) with 95% CI (dashed lines) versus time since exposure.

Black dots indicate scaled Schoenfeld residuals from the Cox proportional hazards model. Exposure starts at age 5 when children are supposed to receive dose 5 of DTaP vaccine.

6.7 Supplementary Materials

6.7.1 Supplementary Tables

Table 1: Piecewise constant 1- incidence rate ratio (1-IRR) estimates with 95% confidence intervals (CI) for 5 doses of DTaP vaccine among 5-9 year old children living in King County WA between 1st January, 2013, and 31st December, 2017

Time since vaccination (years)	**Cases in vaccinated group**	**Person-time at risk in vaccinated group**	**Cases in under-vaccinated group**	**Person-time at risk in under-vaccinated group**	$VE_{s,\lambda}$	**95% CI**
1	4	1214150	8	259809	0.91	0.97, 0.69
2	10	937871.1	12	206962	0.82	0.93, 0.59
3	3	678265.6	4	157275.4	0.84	0.96, 0.27
4	2	421702.2	7	102986.6	0.93	0.98, 0.67
5	3	166029.8	2	41190.5	0.64	0.94, -1.16

Poisson regression models are adjusted for age and neighborhood SES index.

Person time at risk measured in months

6.7.2 Supplementary Figures

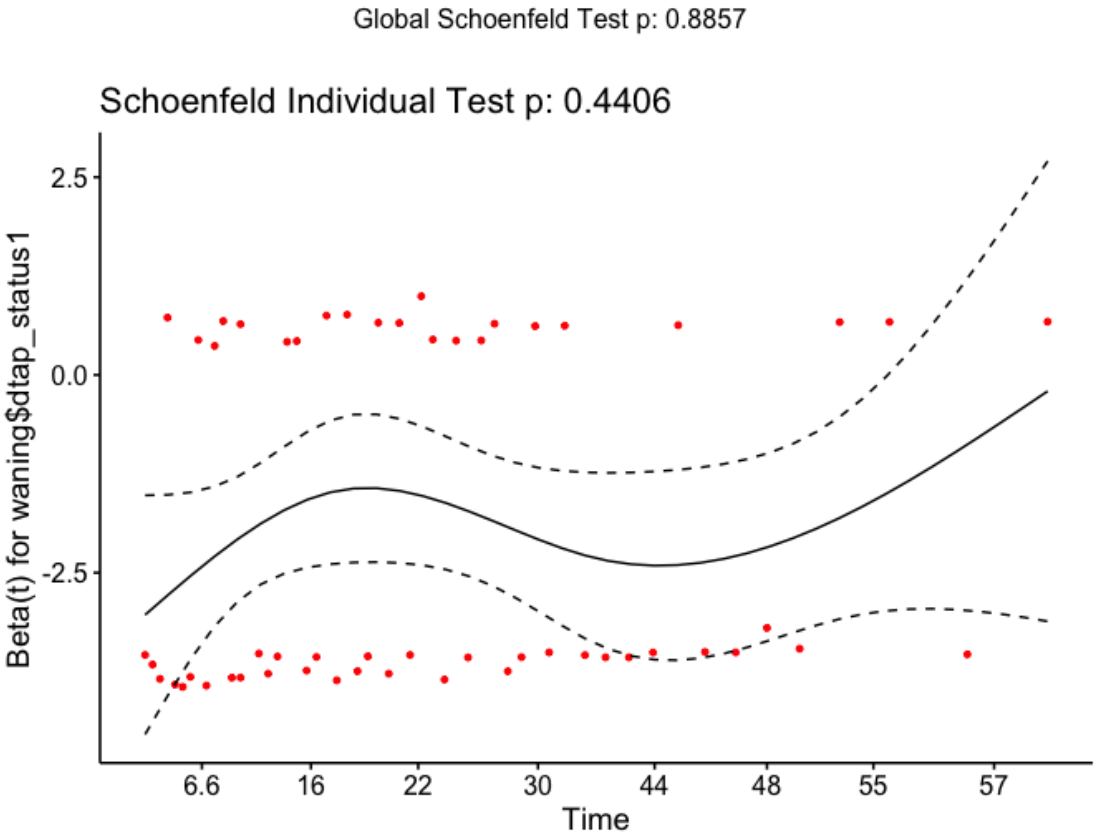


Figure 1: Schoenfeld residuals for DTaP vaccination status from Cox proportional hazards model fitted to estimate vaccine efficacy of 5-dose DTaP vaccine series on pertussis risk compared to under-vaccination.

Global P-value of 0.88 suggests proportionality of hazards is a valid assumption. There is lack of evidence for correlation between scaled Schoenfeld residuals and time (p=0.44), suggesting no significant waning of effects.

Chapter 7

Conclusion and Discussion

Despite widespread vaccination, pertussis continues to circulate endemically in most countries. Understanding vaccine-related factors that shape pertussis epidemiology is critical because vaccination is the most effective strategy for pertussis prevention. This dissertation used time-series surveillance data and individual-level incidence and immunization data to describe pertussis persistence in King County, WA, and examined the role of the acellular vaccines in shaping certain aspects of its epidemiology.

The pertussis incidence and immunization data used in this dissertation came from different sources, and data for different age groups and time periods were used for specific analyses. For cases, we used a time-series of reported pertussis cases from a surveillance database maintained by the Public Health Seattle and King County. That pertussis is severely under-reported is well-known, with an estimated reporting rate of $\sim 10\%$ in the US [8]. Moreover, reporting efficiency varies by age and vaccination status [72] [176]. Vaccinated adolescents and adults might not show typical symptoms of pertussis and may never get a pertussis diagnosis. Physicians may be less likely to suspect and test for pertussis if a patient is fully-vaccinated. Pertussis in very young babies also presents atypically, where cyanosis is a common symptom but there is no coughing [1]. Differential reporting may introduce bias in any analysis done using surveillance data, and it is challenging to quantify this bias at least in the analyses presented here. Moreover, pertussis-like disease can be caused by congeners of *B. pertussis*, such as *B. holmesii* and *B. parapertussis*, that surveillance data cannot identify, but could be essential to understand the overall epidemiology of pertussis disease [177]. Of note, however, is that mathematical models often estimate reporting rate from time-series of incidence data and adjust for it, which is an advantage over other methods [8].

While we had reports of pertussis case data for all ages between 1999 and 2017, we obtained individual-level immunization information for the population of King County only for children younger than 10 years of age. Thus, we could only link cases 0 to 9 years of age reported between 2008 and 2017 to the immunization data and could study the effect of vaccination only for this age group. Considering that adolescents formed $\sim 41\%$ of the cases reported between 2008 and 2017 and contributed substantially to the epidemiology of pertussis in King County, it is unfortunate that we did not have individual level immunization data for this age group to study the direct and indirect vaccine effects and waning of the adolscent TDaP booster. On the other hand, using data from the immunization registry for almost all children 0 through 9 years of age living or born in King County allowed us to conduct a population-based study with almost the entire population of interest, which was a rare opportunity. There could be misclassification bias because some immunizations could be missing from the registry, however this would result in conservative estimates of pertussis relative risk. One threat to external validity is if some children are not captured at all by the registry and are also at higher risk for pertussis, however, as birth certificates are loaded twice a month in the registry, this bias is likely to be minimal.

Area-level vaccination coverage was estimated in two different ways in chapters 3 and 5. In chapter 3 for the epidemic-endemic model and ecological models, school-district level vaccination coverage was estimated annually as proportion of 19-35 month old children who had received ≥ 4 doses of DTaP. This is consistent with how Public Health Seattle and King County and Washington State Departments of Health report area-level vaccination coverage for counties and the state. For obtaining population-level effects in chapter 5, census-tract, that is, cluster-level vaccination coverage was estimated as proportion of all children 0 through 9 years of age who were age-appropriately vaccinated with the recommended DTaP doses per US schedule, as measured on the last day of follow-up (12/31/2017). Vaccination coverage measured for population-level effects was not time-varying because there are no methods developed yet for estimating vaccine indirect effects for changing vaccination coverage. Measuring vaccination coverage at baseline was also not possible because participants started follow-up at different times. For a more complete measure of population immunity, we included everyone in the cohort when measuring vaccination coverage for this analysis. Both analyses could be sensitive to how vaccination coverage is estimated. As a sensitivity analysis, we re-estimated population-level vaccine effects in chapter 5 where vaccination coverage was estimated

as proportion of all children 0 through 9 years of age who were age-appropriately vaccinated with the recommended DTaP doses per US schedule, as measured on 12/31/2015 (this date was arbitrarily selected for sensitivity analysis). Our inference did not change, but appropriate estimation of vaccination coverage should be an important considertion when estimating population-level vaccine effects using administrative data. How to measure these effects when vaccination coverage varies with time is an area of future research.

We estimated direct vaccine effectiveness for the acellular pertussis vaccine using three different statistical models and datasets in this dissertation. We first used the pertussis time series data for all ages and school district level vaccination coverage to estimate vaccine direct effects with the ecological vaccine model. Vaccination coverage at the school district level was estimated using data from the WA-IIS. This model assumes that (1) Poisson approximates binomial distribution for a rare disease, (2) when the force of infection, λ_t is small, a Taylor expansion approximates $1 - e^{-\lambda_t} \approx \lambda_t$, and (3) number of infections are negligible compared to the susceptible population [92]. Thus, the ecological vaccine model does not account for depletion of susceptibles. Using this model, we estimated direct vaccine effectiveness for being vaccinated with four or more doses by age 35 months to be 83% (95% CI: 63%, 95%). We also used individual-level data on pertussis and vaccination status and Cox proportional hazards model to estimate direct effects of the DTaP vaccine for all children 3 months to 9 years of age in the WA-IIS. In this model, the 5-dose DTaP series was included as a time-dependent exposure and we assumed that each dose provides short-term immunity until the next dose is due, that the probability of receiving a DTaP dose at a given age is independent of previous doses received, and the effect of DTaP doses is not cumulative. With this model, we estimated direct vaccine effectiveness comparing vaccinated person-time at risk to unvaccinated person-time at risk to be 72% for the whole cohort (95% CI: 65%, 77%) and 75% (95% CI: 62%, 84%) in the lowest vaccination coverage quartile. Finally, using individual-level data on pertussis and vaccination status for children 5 through 9 years of age in the WA-IIS and Cox proportional models again, we estimated direct vaccine effects after at least 5 doses of DTaP to be 79% (95% CI: 91%, 51%) after 1 year, 83% (95% CI: 93%, 63%) after 2 years, 90% (95% CI: 71%, 97%) after 3 years and 83% (95% CI: 39%, 95%) after 4 years. Even though we used different models and data to estimate DTaP vaccine effectiveness in different age groups, our direct vaccine effectiveness estimates were consistent across analyses. Thus, this dissertation found that acellular pertussis vaccines provides good direct protection to infants and young children, and vaccine effectiveness is high. However, none of these statistical models

help us infer the mechanism of vaccine failure for DTaP vaccine.

We estimated vaccine effectiveness in chapters 5 and 6 as 1- (hazard rate of pertussis comparing vaccinated to under-vaccinated groups). The choice of measures of vaccine effectiveness depends on the mode of action of the vaccination, whether it is leaky or all-or-none [172]. But as pertussis is a rare disease, the hazard rate and the cumulative incidence rate are both low, and $VE_{S,CI}$ will approximate $VE_{S,\lambda}$. In this situation, our choice of measure was $VE_{S,\lambda}$ because using the hazard rate allowed us to account for different follow-up time in our cohort [13].

We found evidence for indirect vaccine effects, suggesting that pertussis vaccination with the primary series may protect unvaccinated individuals, thus generating herd effects. This result is consistent with epidemiologic trends in Sweden where mass vaccination with acellular pertussis vaccines resulted protection of unvaccinated individuals [17] as well as the results of a randomized controlled trial that directly measured vaccine indirect effects [18]. However, it cannot be deduced from the methods we used whether the indirect protection is due to reduction in susceptibility to infection or reduction in infectiousness after exposure in vaccinated individuals. Even so, this is an important result supporting the hypothesis that the current acellular vaccines reduce infection transmission at least to some degree.

Observational studies that have used administrative records and immunization registry data to estimate duration of immunity after vaccination with the DTaP vaccine have inferred immediate and rapid loss of immunity based on odds ratio slopes [160] [113]. This interpretation has been challenged in a study where, using a dynamic model, the authors showed that increase in annual odds of pertussis that the epidemiologic studies found are actually consistent with slowly waning. long-lived immunity [162]. Using an age-structured transmission model, they found that vaccine effectiveness was higher than 75% among 5 to 9 year old children and 65% remained immune 5 years after their last DTaP dose [162]. Our results are more consistent with the study with the dynamic transmission model despite differences in the methology used- we estimated vaccine effectiveness to be over 75% for up to 4 years after the fifth dose. Very high rates of waning are also inconsistent with our pertussis incidence data. We only had 55 cases of pertussis in our cohort of 5 to 9 year olds. If vaccine effectiveness did indeed drop rapidly within a few years as suggested by the observational studies, we would expect to see higher pertussis incidence in this age group and overall. Despite using a different

data source (individual-level pertussis and vaccination status vs. incidence time-series) and a different method (statistical model vs. dynamic transmission model), our results are in agreement with the conclusion made by De Celles, et al. (2019) that interpreting increasing temporal trends in the odds of pertussis as rapid decline of immunity may not be a reliable method to estimate waning immunity for pertussis vaccines [162]. We propose that Schoenfeld residuals be used to estimate the vaccine waning effects when using a statistical model.

Understanding the mechanism by which the acellular vaccine fails can give insights into vaccine-related reasons for pertussis resurgence and help guide policy regarding vaccine schedules and number of doses required to effectively prevent pertussis. While the statistical models used here cannot differentiate between the modes of vaccine failure, especially when the disease is rare, dynamic transmission models can be used to adress this question. Magpantay, et al. (2016) estimated vaccine failure due to primary vaccine failure, leakiness, and waning using time-series data from Italy. They derived a measure of the effectiveness of imperfect vaccines from a specific dynamic model and termed it "vaccine impact" [23] [178]. A vaccine can fail in more than one ways. So one might think of the vaccine impact measure as a composite measure of the vaccine effect. It is a measure of what a vaccinated person contributes to the basic reproduction number compared to an unvaccinated person. It was derived as:

$$\phi = (1 - \epsilon_A)[(1 - \theta) + \theta(1 - \epsilon_L)(1 - \epsilon_W)] \tag{7.1}$$

where ϕ is the vaccine impact, ϵ_A is the probability of primary vaccine failure, ϵ_L is the reduction in probability of infection after exposure in vaccinated compared to unvaccinated individuals, ϵ_W is the probability that the vaccine protection will wane in the individual's lifetime, and θ is the relative infectiousness of vaccinated to unvaccinated individuals. If we assume slowly waning immunity (ϵ_W=0) and that vaccinated and unvaccinated individuals are equally infectious ($\theta = 1$), and plug in the vaccine effectiveness estimate obtained from the ecological vaccine model in the equation, we find vaccine impact to be 0.83 for both primary failure and leaky modes of vaccine action. The vaccine impact is the same for both because of the form of the equation. We can include the vaccine effect estimate obtained from the ecological vaccine model in eq. (7.1) either as the probability of primary vaccine failure or as a measure of leakiness because the ecological model does not tell us anything about the mode of vaccine failure. In the Magpantay, et al. (2016) paper, the authors estimated that vaccinated and unvaccinated

individuals are equally infectious and that vaccinated cases are not reported at all. In our cohort, $\sim$ 60 % of the reported cases were vaccinated. So while we cannot estimate the differential rates of reporting for vaccinated vs. unvaccinated individuals from our data, reporting probability of vaccinated cases is not zero. The authors also estimated that vaccinated and unvaccinated cases are equally infectious. However, an observational study in Senegal found otherwise. Pertussis vaccination reduced infectiousness in breakthrough vaccinated cases and estimated vaccine effectiveness for infectiousness was found to be 85% (95% CI: 46%, 95%) [19]. One caveat for the Senegal study is that vaccinated individuals had both DTwP and DTaP vaccines, so it is not clear what role the DTaP vaccines played in reducing infectiousness. One might revisit these assumptions in future analyses.

Spatial structure of epidemics can reveal insights about the underlying risk factors of disease. Often, space is a proxy for risk factors such as socio-economic status, access to healthcare, and under-vaccination [55]. We found an overlap between clusters of census tracts with high pertussis risk and high vaccine refusal, which suggests some association between the two. One might include annual census-tract level non-medical vaccine exemption rates in the Bayesian heirarchical space time models as a covariate and assess its predictive effect on pertussis risk and its spatial dependence. An important finding from this chapter was the scale of spatial transmission of pertussis, which was found to transmit over short ranges. We found that pertussis can be transmitted to persons living in the immediate vicinity of cases. This type of analysis has not been done for pertussis before. One interesting future analysis would be to collect mobility data for cases and locations visited other than households such schools, offices, and daycares, and examine spatial clustering patterns by age and location. It is challenging to study who infected whom for pertussis using only clinical diagnosis data, [20] and these kind of analyses can identify plausible transmission events for cases whose souce of infection is unknown.

Taking all evidence into account, this dissertation makes the following conclusions:

1. King County has not experienced a resurgence in pertussis between 1999 and 2017. Even so, endemic transmission and occasional outbreaks in the county despite high vaccination coverage indicate that vaccine failure could play a role in pertussis persistence.

2. Not only do the current DTaP vaccines provide effective direct protection against pertussis, but there is evidence supporting the hypothesis that DTaP vaccines

reduce overall pertussis transmission- whether by reduction in susceptibility or infectiousness of vaccinated individuals could not be inferred.

3. DTaP vaccine effect does not wane rapidly after 5 doses and might play a limited role as the mechanism of vaccine failure that led to pertussis persistence or resurgence at least in King County.

4. The current vaccination schedule for the US is effective is preventing pertussis disease in infants and young children and no changes in number of doses or interval may be necessary.

7.1 Future Work

The statistical models and data used in this dissertation, while useful for estimating different types of vaccine effectiveness, do not differentiate between leaky and all-or-none vaccine failure. Pertussis transmission models with different types of vaccine failures built in have been used to explore which mode of vaccine failure best explains observed incidence data in many different settings [22] [23] [163]. Continued circulation of pertussis observed in King County and presence of vaccinated cases suggests that the vaccine fails to protect at least some individuals. While evidence suggests no rapid waning upto age 10 years in our analysis, we estimated waning in a restricted cohort of children 5 to 9 years of age who had 5 doses of DTaP. As next steps, to estimate different types of vaccine failures, we will fit dynamic mechanistic models of pertussis transmission and examine the effects of different type of vaccine failures on pertussis epidemiology in King County. We will fit the models to pertussis incidence data for all ages from 1999 to 2017. We will then use likelihood-based statistical inference methods within the Partially-Observed Markov Process (POMP) modeling framework to estimate the values of vaccine- and disease-related parameters and evaluate how well the model fits the observed data [179]. The dynamic models we fit will be similar to the ones used by Magpantay, et al. (2016). Specifically, we will fit extensions of the Susceptible-Exposed-Infected-Recovered (SEIR) models [101], and estimate the rate of leakiness of the vaccine and rate of its waning. We will make pragmatic assumptions about primary vaccine failure, and relative infectiousness and relative reporting probability between vaccinated and unvaccinated individuals. We will then compare the estimates of vaccine effectiveness and duration of immunity found from fitting these models to our estimates obtained from the statistical models, to see if they are consistent in explaining the observed data.

Bibliography

[1] National Center for Immunization and Respiratory Diseases — Division of Bacterial Diseases — Centers of Disease Control and Prevention. Pertussis (Whooping cough) Signs and Symptoms, 2017.

[2] Karene Hoi Ting Yeung, Philippe Duclos, E. Anthony S. Nelson, and Raymond Christiaan W. Hutubessy. An update of the global burden of pertussis in children younger than 5 years: a modelling study. *The Lancet Infectious Diseases*, 17(9):974–980, 2017.

[3] HP Lambert. The carrier state: Bordetella pertussis. *Journal of Antimicrobial Chemotherapy*, 18:13–16, 1986.

[4] J Cherry, P Brunell, G Golden, and D Karzon. Report of the Task Force on Pertussis and Pertussis Immunization-1988. *Pediatrics*, 81(supp):933–84, jan 1988.

[5] Nicola P Klein. Licensed pertussis vaccines in the United States. History and current state. *Human vaccines & immunotherapeutics*, 10(9):2684–90, 2014.

[6] WHO. Pertussis Vaccines: WHO position paper-August 2015. *The Weekly Epidemiological Record*, (40):145–168, 2018.

[7] Pejman Rohani and John M. Drake. The decline and resurgence of pertussis in the US. *Epidemics*, 3(3-4):183–188, 2011.

[8] Pejman Rohani and Samuel V. Scarpino. *Pertussis Epidemiology, Immunology, and Evolution*. Oxford University Press, 1 edition, 2019.

[9] Elena Chiappini, Alessia Stival, Luisa Galli, and Maurizio de Martino. Pertussis re-emergence in the post-vaccination era. *BMC Infectious Diseases*, 13(1):1, 2013.

[10] Matthieu Domenech de Celles, Felicia M. G. Magpantay, Aaron A. King, and Pejman Rohani. The pertussis enigma: reconciling epidemiology, immunology and evolution. *Proc R Soc B*, 283(1822):20152309–, 2016.

[11] D W Jackson and Pejman Rohani. Perplexities of pertussis: recent global epidemiological trends and their potential causes. *Epidemiology and infection*, 142(4):672–684, 2014.

[12] Benjamin M. Althouse and Samuel V. Scarpino. Contrasting ecological and evolutionary signatures of whooping cough epidemiological dynamics. In Pejman Rohani and Samuel V. Scarpino, editors, *Pertussis: Epidemiology,Immunology, and Evolution*, chapter 13, pages 211–224. Oxford University Press, 1st edition, 2019.

[13] M. Elizabeth Halloran, C J Struchiner, and Ira M. Longini. *Design and Analysis of Vaccine Studies*. Springer New York, New York, 2010.

[14] Elizabeth Wolf, Ali Rowhani-Rahbar, Azadeh Tasslimi, Jasmine Matheson, and Chas DeBolt. Parental country of birth and childhood vaccination uptake in Washington State. *Pediatrics*, 138(1), 2016.

[15] J. M. Warfel, L. I. Zimmerman, and T. J. Merkel. Acellular pertussis vaccines protect against disease but fail to prevent infection and transmission in a nonhuman primate model. *Proceedings of the National Academy of Sciences*, 111(2):787–792, 2014.

[16] Benjamin M. Althouse and Samuel V. Scarpino. Asymptomatic transmission and the resurgence of Bordetella pertussis. *BMC Medicine*, 13(1):146, 2015.

[17] Pejman Rohani, Xue Zhong, and Aaron King. Contact Network Structure Explains the Changing Epidemiology of Pertussis. *Science*, 1781(November):982–986, 2010.

[18] Birger Trollfors, John Taranger, Teresa Lagergard, Valter Sundh, Dolores Bryla, Rachel Scheenerson, and John Robbins. Immunization of children with pertussis toxoid decreases spread of pertussis within the family. *The Pediatric Infectious Disease Journal*, 17(3):196–199, 1998.

[19] Marie Pierre Préziosi and M. Elizabeth Halloran. Effects of pertussis vaccination on transmission: Vaccine efficacy for infectiousness. *Vaccine*, 21(17-18):1853–1861, 2003.

[20] Natasha S Crowcroft and Elizabeth Miller. Pertussis Epidemiology. In Pejman Rohani and Samuel V. Scarpino, editors, *Pertussis: Epidemiology,Immunology, and Evolution*, chapter 4, pages 66–86. Oxford University Press, New York, 1 edition, 2019.

[21] Helen J. Wearing and Pejman Rohani. Estimating the duration of pertussis immunity using epidemiological signatures. *PLoS Pathogens*, 5(10), 2009.

[22] Matthieu Domenech De Cellès, Felicia M.G. Magpantay, Aaron A. King, and Pejman Rohani. The impact of past vaccination coverage and immunity on pertussis resurgence. *Science Translational Medicine*, 10(434), 2018.

[23] F. M.G. Magpantay, M. Domenech De Cellès, P. Rohani, and A. A. King. Pertussis immunity and epidemiology: Mode and duration of vaccine-induced immunity. *Parasitology*, 143(7):835–849, 2016.

[24] Aaron M. Wendelboe, Annelies Van Rie, Stefania Salmaso, and Janet A. Englund. Duration of immunity against pertussis after natural infection or vaccination. *Pediatric Infectious Disease Journal*, 24(5 SUPPL.):58–61, 2005.

[25] Ashleigh Mcgirr and David N Fisman. Duration of pertussis immunity after DTaP immunization: A meta-analysis. *Indian Journal of Practical Pediatrics*, 17(1):15, 2015.

[26] Youyi Fong, M. Elizabeth Halloran, Jin Kyung Park, Florian Marks, John D. Clemens, and Dennis L. Chao. Efficacy of a bivalent killed whole-cell cholera vaccine over five years: A re-analysis of a cluster-randomized trial. *BMC Infectious Diseases*, 18(1):1–8, 2018.

[27] Wolfe S Hamborsky J, Kroger A. *Center of Disease Control and Prevention. Epidemiology and Prevention of Vaccine-Preventable Diseases.* Public Health Foundation, Washington D.C., 13th edition, 2015.

[28] Maria A Riolo, Aaron A King, and Pejman Rohani. Can vaccine legacy explain the British pertussis resurgence? *Vaccine*, 31(49):5903–5908, 2013.

[29] W. Scheil, S. Cameron, C. Roberts, and R. Hall. Pertussis in South Australia 1893 to 1996. *Communicable diseases intelligence*, 22(5):76–80, 1998.

[30] T Smith, J Rotondo, S Desai, and H Deehan. Pertussis Surveillance in Canada: Trends to 2012. *Canada Communicable Disease Report*, 40(3):21–30, 2014.

[31] Pertussis Working group. WHO SAGE pertussis working group: Background paper SAGE April 2014. Technical Report April, 2014.

[32] CDC — National Center for Immunization and Respiratory Diseases — Division of Bacterial Diseases. 2012 Final Pertussis Surveillance Report. Technical report, Centers for Disease Control and Prevention, 2013.

[33] James D. Cherry. Pertussis in young infants throughout the world. *Clinical Infectious Diseases*, 63(Suppl 4):S119–S122, 2016.

[34] Cici Bauer, Jon Wakefield, Harvard Rue, Steve Self, Zijian Feng, and Yu Wang. Bayesian Penalized Spline Models for the Analysis of Spatio- Temporal Count Data. *Stat Med.*, 35(11):1848–1865, 2016.

[35] Saad B Omer, Kyle S Enger, Lawrence H Moulton, Neal A Halsey, Shannon Stokley, and Daniel A Salmon. Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis. *American Journal of Epidemiology*, 168(12):1389–1396, 2008.

[36] Jessica E Atwell, Josh Van Otterloo, Jennifer Zipprich, Kathleen Winter, Kathleen Harriman, Daniel a Salmon, Neal a Halsey, and Saad B Omer. Nonmedical vaccine exemptions and pertussis in California, 2010. *Pediatrics*, 132(4):624–630, 2013.

[37] Carlin Aloe, Martin Kulldorff, and Barry R. Bloom. Geospatial analysis of nonmedical vaccine exemptions and pertussis outbreaks in the United States. *Proceedings of the National Academy of Sciences*, 114(27):7101–7105, 2017.

[38] Rubén Solano, Diana Gómez-Barroso, Fernando Simón, Sarah Lafuente, Pere Simón, Cristina Rius, Pilar Gorrindo, Diana Toledo, and Joan A. Caylà. Retrospective space-time cluster analysis of whooping cough re-emergence in Barcelona, Spain, 2000-2011. *Geospatial Health*, 8(2):455–461, 2014.

[39] Moises E Maravi, Lauren E Snyder, L Dean McEwen, Kathryn DeYoung, and Arthur J Davidson. Using Spatial Analysis to Inform Community Immunization Strategies. *Biomedical Informatics Insights*, 9:117822261770062, 2017.

[40] Justin Lessler, Henrik Salje, M. Kate Grabowski, and Derek A.T. Cummings. Measuring spatial dependence for infectious disease epidemiology. *PLoS ONE*, 11(5):1–13, 2016.

[41] Marta Blangiardo and Michela Cameletti. Spatio-temporal models. In *Spatial And Spatio-Temporal Bayesian Models With R-INLA*, chapter Spatio-tem, pages 235–258. John Wiley & Sons, Ltd, 2015.

[42] Matt Keeling and Pejman Rohani. Spatial Models. In *Modeling of Infectious Diseases in Humans and Animals*. 2008.

[43] M Kulldorff. Prospective time-periodic geographical disease surveillance using a scan statistic. ;164:61–72. *J R Stat Soc A Stat Soc.*, 164:61–72, 2001.

[44] US Census Bureau. US Census Bureau: American Community Survey, 2017.

[45] Center of Disease Control and Prevention. Pertussis / Whooping Cough (Bordetella pertussis) 2014 Case Definition, 2014.

[46] Office of Immunization Washington State Department of Health and Child Profile. Public Health Immunization Reports by County. Technical report, 2018.

[47] ESRI. ESRI 2012. ArcGIS Desktop: Release 10.1 Redlands, CA: Environmental Systems Research Institute., 2012.

[48] U.S. Department of Commerce Economics and Statistics Administration U.S. Census Bureau. GEOCODING USING ARCGIS & TIGER/LINE® SHAPEFILES. Technical report, U.S. Department of Commerce April 2013 Economics and Statistics Administration U.S. CENSUS BUREAU, 2013.

[49] Steven Manson, Jonathan Schroeder, David Van Riper and Steven Ruggles. IPUMS National Historical Geographic Information System: Version 12.0 [Database]. Minneapolis: University of Minnesota. 2017., 2017.

[50] 2017 TIGER/Line Shapefiles Technical Documentation / prepared by the U.S. Census Bureau, 2017.

[51] Washington State Department of Health. Office of Immunization and Child Profile. School Immunization Data Tables, 2018.

[52] Jennifer L. Liang, Tejpratap Tiwari, Pedro Moro, Nancy E. Messonnier, Arthur Reingold, Mark Sawyer, and Thomas A. Clark. Prevention of Pertussis, Tetanus, and Diphtheria with Vaccines in the United States: Recommendations of the Advisory Committee on Immunization Practices (ACIP) Morbidity and Mortality Weekly Report Recommendations and Reports Centers for Disease Control a. *Recommendations and Reports*, 67(2), 2018.

[53] Marie-Pierre Préziosi and M Elizabeth Halloran. Effects of pertussis vaccination on disease: vaccine efficacy in reducing clinical severity. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 37(January 1993):772–779, 2003.

[54] Jennifer L. Kelsey and Ellen B. Gold. Observational Epidemiology. In *International Encyclopedia of Public Health*, volume 5, pages 295–307. 2016.

[55] Jonathan C. Wakefield, N.G Best, and L Waller. Bayesian approaches to disease mapping. In P. Elliot, Jonathan C. Wakefield, N.G Best, and D.J Briggs, editors, *Spatial Epidemiology : Methods and Applications*, pages 104–127. London, 1 edition, 2000.

[56] Harvard Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392, 2009.

[57] AP Moran. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1):17–23, 1950.

[58] Kyra H. Grantz, Madhura S. Rane, Henrik Salje, Gregory E. Glass, Stephen E. Schachterle, and Derek A.T. Cummings. Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. *Proceedings of the National Academy of Sciences of the United States of America*, 113(48):13839–13844, 2016.

[59] John Giles, Henrik Salje, and Justin Lessler. IDSpatialStats: Estimate Global Clustering in Infectious Disease, 2019.

[60] Sarah Boslaugh. *Statistics in a Nutshell*. O'Reilley Media, 2nd edition, 2012.

[61] Ottar N. Bjornstad. ncf: Spatial Covariance Functions., 2019.

[62] Ottar N. Bjørnstad. Spatial and Spatiotemporal patterns. In *Epidemics: Models and Data using R*, chapter 13, pages 241–252. Springer Nature, Cham, Switzerland, 1 edition, 2018.

[63] Martin Kulldorff, Richard Heffernan, Jessica Hartman, Renato Assuno, and Farzad Mostashari. A space-time permutation scan statistic for disease outbreak detection. *PLoS Medicine*, 2(3):0216–0224, 2005.

[64] M Kulldorff. SatScan user guide, 2018.

[65] Kulldorff M. and Information Management Services Inc. SaTScanTM v8.0: Software for the spatial and space-time scan statistics., 2009.

[66] RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.

[67] R Core Team. R: A language and environment for statistical computing., 2019.

[68] Fiona P. Havers, Pedro L. Moro, Paul Hunter, Susan Hariri, and Henry Bernstein. Use of Tetanus Toxoid, Reduced Diphtheria Toxoid, and Acellular Pertussis Vaccines: Updated Recommendations of the Advisory Committee on Immunization Practices - United States, 2019. *MMWR. Morbidity and mortality weekly report*, 69(3):77–83, 2020.

[69] Centers for Disease Control and Prevention (CDC). Tdap vaccination coverage among women who had a recent live birth by state, Pregnancy Risk Assessment Monitoring System (PRAMS), 2017, 2017.

[70] Ana I. Bento, Maria A. Riolo, Yoon H. Choi, Aaron A. King, and Pejman Rohani. Core pertussis transmission groups in England and Wales: A tale of two eras. *Vaccine*, 36:1160–1166, 2018.

[71] Hélène Broutin, Jean François Guégan, Eric Elguero, François Simondon, and Bernard Cazelles. Large-scale comparative analysis of pertussis population dynamics: Periodicity, synchrony, and impact of vaccination. *American Journal of Epidemiology*, 161(12):1159–1167, 2005.

[72] Lucy A Mcnamara, Vaccine Preventable, Diseases Branch, Respiratory Diseases, Tami Skoff, Vaccine Preventable, Diseases Branch, Respiratory Diseases, Amanda Faulkner, Vaccine Preventable, Diseases Branch, Respiratory Diseases, Lisa Miller, Colorado Disease Control, Environmental Epidemiology Division, Kathy Kudish, Cynthia Kenyon, Marisa Bargsten, Santa Fe, Shelley Zansky, Amy D Sullivan, Communicable Disease Services, and Stacey Martin. Reduced Severity of Pertussis in Persons with Age-Appropriate Pertussis Vaccination — United States, 2010–2012. *Clinical Infectious Diseases*, 65(5):811–818, 2017.

[73] Kathleen Winter, James D. Cherry, and Kathleen Harriman. Effectiveness of prenatal tetanus, diphtheria, and acellular pertussis vaccination on pertussis severity in infants. *Clinical Infectious Diseases*, 64(1):9–14, 2017.

[74] Pejman Rohani, David J.D. Earn, and Bryan T. Grenfell. Impact of immunisation on pertussis transmission in England and Wales. *Lancet*, 355(9200):285–286, 2000.

[75] Henrik Salje, Justin Lessler, Timothy P. Endy, Frank C. Curriero, Robert V. Gibbons, Ananda Nisalak, Suchitra Nimmannitya, Siripen Kalayanarooj, Richard G. Jarman, Stephen J. Thomas, Donald S. Burke, and Derek A.T. Cummings. Revealing the microscale spatial signature of dengue transmission and immunity in an urban population. *Proceedings of the National Academy of Sciences of the United States of America*, 109(24):9535–9538, 2012.

[76] Piraya Bhoomiboonchoo, Robert V. Gibbons, Angkana Huang, In Kyu Yoon, Darunee Buddhari, Ananda Nisalak, Natkamol Chansatiporn, Mathuros Thipayamongkolgul, Siripen Kalanarooj, Timothy Endy, Alan L. Rothman, Anon Srikiatkhachorn, Sharone Green, Mammen P. Mammen, Derek A. Cummings, and Henrik Salje. The Spatial Dynamics of Dengue Virus in Kamphaeng Phet, Thailand. *PLoS Neglected Tropical Diseases*, 8(9):1–9, 2014.

[77] Mary K. Grabowski, Justin Lessler, Andrew D. Redd, Joseph Kagaayi, Oliver Laeyendecker, Anthony Ndyanabo, Martha I. Nelson, Derek A.T. Cummings, John Baptiste Bwanika, Amy C. Mueller, Steven J. Reynolds, Supriya Munshaw, Stuart C. Ray, Tom Lutalo, Jordyn Manucci, Aaron A.R. Tobian, Larry W. Chang, Chris Beyrer, Jacky M. Jennings, Fred Nalugoda, David Serwadda, Maria J. Wawer, Thomas C. Quinn, and Ronald H. Gray. The Role of Viral Introductions in Sustaining Community-Based HIV Epidemics in Rural Uganda: Evidence from Spatial Clustering, Phylogenetics, and Egocentric Transmission Models. *PLoS Medicine*, 11(3), 2014.

[78] Kristine M. Bisgard, F. Brian Pascual, Kristen R. Ehresmann, Claudia A. Miller, Christy Cianfrini, Charles E. Jennings, Catherine A. Rebmann, Julie Gabel, Stephanie L. Schauer, and Susan M. Lett. Infant pertussis: Who was the source? *Pediatric Infectious Disease Journal*, 23(11):985–989, 2004.

[79] N. S. Crowcroft, R. Booy, T. Harrison, L. Spicer, J. Britto, Q. Mok, P. Heath, I. Murdoch, M. Zambon, R. George, and E. Miller. Severe and unrecognised: Pertussis in UK infants. *Archives of Disease in Childhood*, 88(9):802–806, 2003.

[80] Seattle Department of Transportation. Seattle Department of Transportation 2018 Traffic Report. Technical report, 2018.

[81] Aaron M. Wendelboe, Michael G. Hudgens, Charles Poole, and Annelies Van Rie. Estimating the role of casual contact from the community in transmission of Bordetella pertussis to young infants. *Emerging Themes in Epidemiology*, 4:1–7, 2007.

[82] Office of Immunization and Child Profile. Washington State Department of Health. Public Health Immunization Reports by School District.

[83] T. Roice Fulton, Varun K. Phadke, Walter A. Orenstein, Alan R. Hinman, Wayne D. Johnson, and Saad B. Omer. Protective Effect of Contemporary Pertussis Vaccines: A Systematic Review and Meta-analysis. *Clinical Infectious Diseases*, 62(9):1100–1110, 2016.

[84] World Health Organization. WHO Immunization Coverage, 2020.

[85] Hélène Broutin, Eric Elguero, François Simondon, and Jean-François Guégan. Spatial dynamics of pertussis in a small region of Senegal. *Proceedings of the Royal Society B: Biological Sciences*, 271(September):2091–2098, 2004.

[86] Erin Duffy and Kimberly Shea. Effect of County-level Pertussis Vaccine Coverage on County-level Pertussis Incidence Rates in California During the 2010 Pertussis Epidemic. In *ID Week*, 2012.

[87] P Y Iroh Tam, J S Menk, J Hughes, and S L Kulasingam. An ecological analysis of pertussis disease in Minnesota, 2009-2013. *Epidemiology and infection*, (2016):1–9, 2015.

[88] X. HUANG, S. LAMBERT, C. LAU, R. J. SOARES MAGALHAES, J. MARQUESS, M. RAJMOKAN, G. MILINOVICH, and W. HU. Assessing the social and environmental determinants of pertussis epidemics in Queensland, Australia: a Bayesian spatio-temporal analysis. *Epidemiology and Infection*, 145(06):1221–1230, 2017.

[89] Jon Wakefield, Tracy Qi Dong, and Vladimir N. Minin. Spatio-Temporal Analysis of Surveillance Data. *Handbook of Infectious Disease Data Analysis*, pages 455–475, 2019.

[90] S Richardson and C Monfort. Ecological Correlation Studies. In P Elliott, JC Wakefield, Ng Best, and DG Briggs, editors, *Spatial Epidemiology: Methods and Application*. Oxford University Press, 2000.

[91] Jonathan Wakefield. Ecologic Studies Revisited. *Annual Review of Public Health*, 29(1):75–90, 2008.

[92] Leigh H. Fisher and Jon Wakefield. Ecological inference for infectious disease data, with application to vaccination strategies. *Statistics in Medicine*, 39(3):220–238, 2020.

[93] Sebastian Meyer, Leonhard Held, and Michael Höhle. hhh4: Endemic-epidemic modeling of areal count time series. *Journal of Statistical Software*, (1), 2016.

[94] CDC. Pertussis / Whooping Cough (Bordetella pertussis) 2014 Case Definition, 2014.

[95] Washington State Department of Health. Washington State Department of Health: Immunization Data –Technical Technical Notes. Technical report, 2015.

[96] US Census Bureau. 2010 Census, 2010.

[97] Leonhard Held, Michael Höhle, and Mathias Hofmann. A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling*, 5(3):187–199, 2005.

[98] Michael Höhle, Sebastian Meyer, M. Paul, and L. Held. Predictive assessment of a non-linear random effects model for multivariate time series of infectious disease counts. *Statistics in Medicine*, 30(10):1118–1136, 2011.

[99] S. A. Herzog, M. Paul, and L. Held. Heterogeneity in vaccination coverage explains the size and occurrence of measles epidemics in German surveillance data. *Epidemiology and Infection*, 139(4):505–515, 2011.

[100] Jiqiang Guo, Jonah Gabry, Ben Goodrich, and Daniel Lee. Package ' rstan ', 2020.

[101] Matt Keeling and Pejman Rohani. *Modeling of Infectious Diseases in Animals*. 2008.

[102] Leigh H. Fisher. Ecological inference for infectious disease data, with application to vaccination strategies, 2020.

[103] Sebastian Meyer, Johannes Elias, and Michael Höhle. A Space-Time Conditional Intensity Model for Invasive Meningococcal Disease Occurrence. *Biometrics*, 68(2):607–616, 2012.

[104] Sebastian Meyer and Leonhard Held. Power-law models for infectious disease spread. *Annals of Applied Statistics*, 8(3):1612–1639, 2014.

[105] Sebastian Meyer and Leonhard Held. Incorporating social contact data in spatio-temporal models for infectious disease spread. *Biostatistics*, 18(2):338–351, 2017.

[106] Rubén Solano, Inma Crespo, María Isabel Fernández, Carles Valero, María Isabel Álvarez, Pere Godoy, Joan A. Caylà, and Àngela Domínguez. Underdetection and underreporting of pertussis in children attended in primary health care centers: Do surveillance systems require improvement? *American Journal of Infection Control*, 44(11):e251–e256, 2016.

[107] Matthew J Ferrari, Ottar Bjornstad, and Andrew Dobson. Estimation and inference of R 0 of an infectious pathogen by a removal method. 198:14–26, 2005.

[108] E. J. Gangarosa, A. M. Galazka, C. R. Wolfe, L. M. Phillips, R. E. Gangarosa, E. Miller, and R. T. Chen. Impact of anti-vaccine movements on pertussis control: The untold story. *Lancet*, 351(9099):356–361, 1998.

[109] Victoria Romanus, Jonsell Ragnar, and Seth-Olof Bergquist. Pertussis in Sweden after the cessation of general immunization in 1979. *The Journal of Pediatric Infectious Diseases*, 6:364–371, 1987.

[110] Mineo Watanabe and Masaaki Nagai. Acellular pertussis vaccines in Japan: Past, present and future. *Expert Review of Vaccines*, 4(2):173–184, 2005.

[111] Kevin D. Forsyth, Tina Tan, Carl Heinz Wirsing von König, Ulrich Heininger, Amar J. Chitkara, and Stanley Plotkin. Recommendations to control pertussis prioritized relative to economies: A Global Pertussis Initiative update. *Vaccine*, 36(48):7270–7275, 2018.

[112] Judith Mueller, Thibaut Koutangni, Cheick Diallo, and Seydou Yaro. Diptheria and tetanus vaccines. Comparative efficacy/effectiveness of schedules in infant immunisation against pertussis, diphtheria and tetanus: Systematic review and meta-analysis. Technical report, 2014.

[113] Baxter R Klein NP, Bartlett J, Fireman B. Waning Protection after Fifth Dose of Acellular Pertussis Vaccine in Children. *Pediatrics*, 137(3):e20153326, 2016.

[114] Helen E. Quinn, Thomas L. Snelling, Kristine K. Macartney, and Peter B. McIntyre. Duration of protection after first dose of acellular pertussis vaccine in infants. *Pediatrics*, 133(3), 2014.

[115] Helen E Quinn and Peter B McIntyre. Pertussis epidemiology in Australia over the decade 1995-2005–trends by region and age group. *Communicable diseases intelligence*, 31(2):205–215, 2007.

[116] Elizabeth T. Luman, Lawrence E. Barker, Mary Mason McCauley, and Carolyn Drews-Botsch. Timeliness of childhood immunizations: A state-specific analysis. *American Journal of Public Health*, 95(8):1367–1374, 2005.

[117] Elizabeth T Luman, Lawrence E Barker, Kate M Shaw, Mary Mason McCauley, James W Buehler, and Larry K Pickering. Timeliness of childhood vaccinations in the United States: days undervaccinated and number of vaccines delayed. *Journal of the American Medical Association*, 293(10):1204–1211, 2005.

[118] Wan Ting Huang, Hui Chen Lin, and Chin Hui Yang. Undervaccination with diphtheria, tetanus, and pertussis vaccine: National trends and association with pertussis risk in young children. *Human Vaccines and Immunotherapeutics*, 13(4):757–761, 2017.

[119] Jason M. Glanz, Komal J. Narwaney, Sophia R. Newcomer, Matthew F. Daley, Simon J. Hambidge, Ali Rowhani-Rahbar, Grace M. Lee, Jennifer C. Nelson, Allison L. Naleway, James D. Nordin, Marlene M. Lugg, and Eric S. Weintraub. Association between undervaccination with diphtheria, tetanus toxoids, and acellular pertussis (DTaP) vaccine and risk of pertussis infection in children 3 to 36 months of age. *JAMA Pediatrics*, 167(11):1060–1064, 2013.

[120] Cameron C Grant, Mavis Roberts, Robert Scragg, Joanna Stewart, Diana Lennon, Denise Kivell, Rodney Ford, and Rosalie Menzies. Delayed immunisation and risk of pertussis in infants : unmatched case-control study Science commentary : Pertussis immunisation. *Bmj*, 326(April):852–853, 2003.

[121] Nina B. Masters, Abram L. Wagner, and Matthew L. Boulton. Vaccination timeliness and delay in low- and middle-income countries: a systematic review of the literature, 2007-2017. *Human Vaccines and Immunotherapeutics*, 15(12):2790–2805, 2019.

[122] Jodie McVernon and Hester De Melker. Role of vaccine schedules. In Pejman Rohani and Samuel V. Scarpino, editors, *Pertussis: Epidemiology,Immunology, and Evolution*, chapter 5. Oxford University Press, 1 edition, 2019.

[123] Jeremy N. Miles, Margaret M. Weden, Diana Lavery, José J. Escarce, Kathleen A. Cagney, and Regina A. Shih. Constructing a Time-Invariant Measure of the Socioeconomic Status of U.S. Census Tracts. *Journal of Urban Health*, 93(1):213–232, 2016.

[124] M. Soobader, C. Cubbin, G. C. Gee, A. Rosenbaum, and J. Laurenson. Levels of analysis for the study of environmental health disparities. *Environmental Research*, 102(2):172–180, 2006.

[125] Adrian Sayers, Yoav Ben-Shlomo, Ashley W. Blom, and Fiona Steele. Probabilistic record linkage. *International Journal of Epidemiology*, 45(3):954–964, 2016.

[126] Ted Enamorado, Ben Fifield, and Kosuke Imai. Fast Probabilistic Record Linkage with Missing Data, 2018.

[127] Center for Disease Control and Prevention. National Center for Immunization and Respiratory Diseases, 2016.

[128] Jason M. Glanz, Sophia R. Newcomer, Komal J. Narwaney, Simon J. Hambidge, Matthew F. Daley, Nicole M. Wagner, David L. McClure, Stan Xu, Ali Rowhani-Rahbar, Grace M. Lee, Jennifer C. Nelson, James G. Donahue, Allison L. Naleway, James D. Nordin, Marlene M. Lugg, and Eric S. Weintraub. A population-based cohort study of undervaccination in 8 managed care organizations across the United States. *JAMA Pediatrics*, 167(3):274–281, 2013.

[129] Gustavo H. Dayan, Kate M. Shaw, Andrew L. Baughman, Liliana C. Orellana, Raúl Forlenza, Alejandro Ellis, Jorge Chaui, Silvia Kaplan, and Peter Strebel. Assessment of delay in age-appropriate vaccination using survival analysis. *American Journal of Epidemiology*, 163(6):561–570, 2006.

[130] SAS Institute Inc. 2013. SAS® 9.4 Statements: Reference. Cary, NC: SAS Institute Inc.

[131] U.S. Department of Health and Human Services. Office of Disease Prevention and Health Promotion. Washington DC. Healthy People 2020.

[132] Michael L. Jackson, Nora B. Henrikson, and David C. Grossman. Evaluating Washington State's immunization information system as a research tool. *Academic Pediatrics*, 14(1):71–76, 2014.

[133] G. Amirthalingam, S. Gupta, and H. Campbell. Pertussis Immunisation and control in England and Wales, 1957 to 2012: A historical review. *Eurosurveillance*, 18(38):1–18, 2013.

[134] Nicoline A.T. Van der Maas, Frits R. Mooi, Sabine C. de Greeff, Guy A.M. Berbers, Marina A.E.Conyn van Spaendonck, and Hester E. de Melker. Pertussis in the Netherlands, is the current vaccination strategy sufficient to reduce disease burden in young infants? *Vaccine*, 31(41):4541–4547, 2013.

[135] Nicholas Wood, Helen E. Quinn, Peter McIntyre, and Elizabeth Elliott. Pertussis in infants: Preventing deaths and hospitalisations in the very young. *Journal of Paediatrics and Child Health*, 44(4):161–165, 2008.

[136] Helen Campbell, Gayatri Amirthalingam, Nick Andrews, Norman K. Fry, Robert C. George, Timothy G. Harrison, and Elizabeth Miller. Accelerating control of pertussis in England and Wales. *Emerging Infectious Diseases*, 18(1):38–47, 2012.

[137] Alexis Pillsbury, Helen E. Quinn, and Peter B. McIntyre. Australian vaccine preventable disease epidemiological review series: pertussis, 2006-2012. *Communicable diseases intelligence quarterly report*, 38(3):E179–E194, 2014.

[138] Lara K. Misegades, Kathleen Winter, Kathleen Harriman, John Talarico, Nancy E. Messonnier, Thomas A. Clark, and Stacey W. Martin. Association of Childhood Pertussis With Receipt of 5 Doses of Pertussis Vaccine by Time Since Last Vaccine Dose, California, 2010. *Jama*, 308(20):2126, 2012.

[139] Douglas J. Opel, Ashmita Banerjee, and James A. Taylor. Use of alternative childhood immunization schedules in King County, Washington, USA. *Vaccine*, 31(42):4699–4701, 2013.

[140] L Zhang, Prietsch Som, I Axelsson, Halperin Sa, Linjie Zhang, Sílvio O M Prietsch, Inge Axelsson, and Scott A Halperin. Acellular vaccines for preventing whooping cough in children (Review) Acellular vaccines for preventing whooping cough in children. *The Cochrane Collaboration*, (9), 2014.

[141] Marie Pierre Préziosi, Abdoulaye Yam, Steven G.F. Wassilak, Laurence Chabirand, Aminata Simaga, Malick Ndiaye, Marème Dia, Franọis Dabis, and François Simondon. Epidemiology of pertussis in a West African community before and after introduction of a widespread vaccination program. *American Journal of Epidemiology*, 155(10):891–896, 2002.

[142] M Elizabeth Halloran. The Minicommunity Design to Assess Indirect Effects of Vaccination. *Epidemiol Method*, 1(1):83–105, 2012.

[143] M. Elizabeth Halloran, C J Struchiner, and Ira M. Longini. Study Designs for Evaluating Different Efficacy and Effectiveness Aspects of Vaccines. 146(10):789–803, 1997.

[144] Richard Hayes and Lawrence Moulton. *Cluster Randomized Trials*. Taylor and Francis Group, 2nd editio edition, 2017.

[145] Michael G Hudgens and M Elizabeth Halloran. Toward Causal Inference With Interference. *American Statistician*, 103(482):832–842, 2008.

[146] M. Elizabeth Halloran and Michael G. Hudgens. Estimating population effects of vaccination using large, routinely collected data. *Statistics in Medicine*, (August 2016), 2017.

[147] Mohammad Ali, Young Ae You, Suman Kanungo, Byomkesh Manna, Jacqueline L. Deen, Anna Lena Lopez, Thomas F. Wierzba, Sujit K. Bhattacharya, Dipika Sur, and John D. Clemens. Assessing different measures of population-level vaccine protection using a case-control study. *Vaccine*, 33(48):6878–6883, 2015.

[148] Mohammad Ali, Michael Emch, Lorenz Von Seidlein, Mohammad Yunus, David A. Sack, Malla Rao, Jan Holmgren, and John D. Clemens. Herd immunity conferred by killed oral cholera vaccines in Bangladesh: A reanalysis. *Lancet*, 366(9479):44–49, 2005.

[149] Carolina Perez-Heydrich, Michael G. Hudgens, M. Elizabeth Halloran, John D. Clemens, Mohammad Ali, and Michael Emch. Assesing Effects of Cholera Vaccination in the Presence of Interference. 70(3):731–741, 2014.

[150] Michael Emch, Mohammad Ali, Jin Kyung Park, Mohammad Yunus, David A. Sack, and John D. Clemens. Relationship between neighbourhood-level killed oral cholera vaccine coverage and protective efficacy: Evidence for herd immunity. *International Journal of Epidemiology*, 35(4):1044–1050, 2006.

[151] Peter C. Austin and Elizabeth A. Stuart. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, 34(28):3661–3679, 2015.

[152] Ahmed M. Khatib, Mohammad Ali, Lorenz von Seidlein, Deok Ryun Kim, Ramadhan Hashim, Rita Reyburn, Benedikt Ley, Kamala Thriemer, Godwin Enwere, Raymond Hutubessy, Maria Teresa Aguado, Marie Paule Kieny, Anna Lena Lopez, Thomas F. Wierzba, Said Mohammed Ali, Abdul A. Saleh, Asish K. Mukhopadhyay, John Clemens, Mohamed Saleh Jiddawi, and Jacqueline Deen. Effectiveness of an oral cholera vaccine in Zanzibar: Findings from a mass vaccination campaign and observational cohort study. *The Lancet Infectious Diseases*, 12(11):837–844, 2012.

[153] World Health Organization. Immunization, Vaccines and Biologicals: National Passive Surveillance Pertussis, 2020.

[154] Tami H. Skoff, Stephen Hadler, and Susan Hariri. The epidemiology of nationally reported pertussis in the United States, 2000-2016. *Clinical Infectious Diseases*, 68(10):1634–1640, 2019.

[155] Danuta M. Skowronski, Gaston De Serres, Diane MacDonald, Wrency Wu, Carol Shaw, Jane Macnabb, Sylvie Champagne, David M. Patrick, and Scott A. Halperin. The Changing Age and Seasonal Profile of Pertussis in Canada. *The Journal of Infectious Diseases*, 185(10):1448–1453, 2002.

[156] F. R. MOOI, N. A. T. VAN DER MAAS, and H. E. De MELKER. Pertussis resurgence: waning immunity and pathogen adaptation – two sides of the same coin. *Epidemiology and Infection*, 142(04):685–694, apr 2014.

[157] Manoj Gambhir, Thomas A. Clark, Simon Cauchemez, Sara Y. Tartof, David L. Swerdlow, and Neil M. Ferguson. A Change in Vaccine Efficacy and Duration of Protection Explains Recent Rises in Pertussis Incidence in the United States. *PLoS Computational Biology*, 11(4):1–16, 2015.

[158] Maxwell A. Witt, Paul H. Katz, and David J. Witt. Unexpectedly limited durability of immunity following acellular pertussis vaccination in preadolescents in a north American outbreak. *Clinical Infectious Diseases*, 54(12):1730–1735, 2012.

[159] Roger Baxter, Joan Bartlett, Ali Rowhani-Rahbar, Bruce Fireman, and Nicola P Klein. Effectiveness of pertussis vaccines for adolescents and adults: case-control study. *BMJ (Clinical research ed.)*, 347(July):f4249, 2013.

[160] Ousseny Zerbo, Joan Bartlett, Kristin Goddard, Bruce Fireman, Edwin Lewis, and Nicola P. Klein. Acellular pertussis vaccine effectiveness over time. *Pediatrics*, 144(1), 2019.

[161] Ruth Koepke, Jens C. Eickhoff, Roman A. Ayele, Ashley B. Petit, Stephanie L. Schauer, Daniel J. Hopfensperger, James H. Conway, and Jeffrey P. Davis. Estimating the effectiveness of tetanus-diphtheria-acellular pertussis vaccine (Tdap) for preventing pertussis: Evidence of rapidly waning immunity and difference in effectiveness by Tdap brand. *Journal of Infectious Diseases*, 210(6):942–953, 2014.

[162] Matthieu Domenech De Cellès, Pejman Rohani, and Aaron A. King. Duration of Immunity and Effectiveness of Diphtheria-Tetanus-Acellular Pertussis Vaccines in Children. *JAMA Pediatrics*, 173(6):588–594, 2019.

[163] Julie C Blackwood, Derek A T Cummings, Hélène Broutin, Sopon Iamsirithaworn, and Pejman Rohani. Deciphering the impacts of vaccination and immunity on pertussis epidemiology in Thailand. *Proceedings of the National Academy of Sciences of the United States of America*, 110(23):9595–600, 2013.

[164] Kathryn l Durham, Ira M Longini, M Elizabeth Halloran, John D Clemens, Azhar Nizam, and Malla Rao. Estimation of Vaccine Efficacy in the Presence of Waning : Application to Cholera Vaccines. *American Journal of Epidemiology*, 147(10):948–959, 1998.

[165] Kathryn l Durham, M. Elizabeth Halloran, Ira M. Longini, and Amita Manatunga. Comparison of Two Smoothing Methods for Exploring Waning Vaccine Effects. *Royal Statistical Society*, 48(3):395–407, 1999.

[166] David W. Hosmer, Stanley Lemeshow, and Susanne May. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. John Wiley & Sons, 2nd edition, 2008.

[167] David Schoenfeld. Partial Residuals for The Proportional Hazards Regression Model A. *Biometrika*, 69(1):239–241, 1982.

[168] Patricia Grambsch and Terry Therneau. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika*, 81(3):515–526, 1994.

[169] Alboukadel Kassambara, Kosinski Marcin, Biecek Przemyslaw, and Fabian Scheipl. Package 'survminer', 2019.

[170] Terry Therneau. Package ' survival ', 2020.

[171] Youyi Fong, Krisztian Sebestyen, Jason Becker, Bendix Carstensen, Daryl Morris, Josh Pasek, and Dennis Chao. Package ' kyotil ', 2019.

[172] P. G. Smith, L. C. Rodrigues, and P. E.M. Fine. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. *International Journal of Epidemiology*, 13(1):87–93, 1984.

[173] Sara Y. Tartof, Melissa Lewis, Cynthia Kenyon, Karen White, Andrew Osborn, Juventila Liko, Elizabeth Zell, Stacey Martin, Nancy E. Messonnier, Thomas A. Clark, and Tami H. Skoff. Waning immunity to pertussis following 5 doses of DTaP. *Pediatrics*, 131(4), 2013.

[174] Anna M. Acosta, Chas DeBolt, Azadeh Tasslimi, Melissa Lewis, Laurie K. Stewart, Lara K. Misegades, Nancy E. Messonnier, Thomas A. Clark, Stacey W. Martin, and Manisha Patel. Tdap vaccine effectiveness in adolescents during the 2012 Washington State pertussis epidemic. *Pediatrics*, 135(6):981–989, 2015.

[175] Ayman Chit, Hossein Zivaripiran, Thomas Shin, Jason K.H. Lee, Antigona Tomovici, Denis Macina, David R. Johnson, Michael D. Decker, and Jianhong Wu. Acellular pertussis vaccines effectiveness over time: A systematic review, meta-analysis and modeling study. *PLoS ONE*, 13(6):1–11, 2018.

[176] Russell S. Barlow, Laura E. Reynolds, Paul R. Cieslak, and Amy D. Sullivan. Vaccinated children and adolescents with pertussis infections experience reduced illness severity and duration, Oregon, 2010-2012. *Clinical Infectious Diseases*, 58(11):1523–1529, 2014.

[177] Iain MacArthur and Andrew Preston. Congenerics: what can be learned about pertussis from pertussis-like disease caused by other Bordetella? In Pejman Rohani and Samuel V. Scarpino, editors, *Pertussis: Epidemiology,Immunology, and Evolution*, chapter 11, pages 182–192. Oxford University Press, 1st edition, 2019.

[178] Felicia M.G. Magpantay, Maria A. Riolo, Matthieu Domenech de Celles, Aaron A. King, and Pejman Rohani. Epidemiological consequences of imperfect vaccines for immunizing infections. *SIAM Journal on Applied Mathematics*, 74(6):1810–1830, 2014.

[179] Edward L. Ionides, Dao Nguyen, Yves Atchadé, Stilian Stoev, and Aaron A. King. Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences*, 112(3):719–724, 2015.