The Outbreak Potential for Measles and its Implications for Elimination

by

Shaun A. Truelove

A dissertation submitted to the Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
December 2017

© 2017 Shaun A. Truelove
All Rights Reserved
Abstract

When the goal of measles elimination, or the absence of continuous transmission in a defined geographical area, was first proposed in 1962 by Alexander Langmuir, immediately prior to licensure of the first measles vaccine, an estimated six million deaths and 135 million cases occurred annually. Since that time, numerous efforts, initiatives, and goals have been initiated with great success, cutting measles incidence and mortality to fractions of what it once was. However, even now, with an effective, safe, and widely available vaccine, measles remains one of the leading causes of death among young children, resulting in approximately 90,000 deaths in 2016. Vaccination is the major reason for this success, yet remains the challenge for complete elimination and eradication: 2016 estimates indicate that only 85% of the children receive the first dose of measles vaccine. However, to meet regional elimination goals, vaccination likely needs to reach the target 95% coverage or more. Essential to this continued progress and the ultimate success of these elimination goals are methods for measuring the risks of transmission and outbreaks. Unfortunately, as we get closer to measles elimination, our current methods have demonstrated several shortcomings. The aims of this dissertation were to explore some areas where our current understanding and methods regarding outbreak and transmission of measles virus and other vaccine-preventable diseases fall short, and to develop new tools for estimating the outbreak potential among populations at various stages of control and elimination, specifically addressing potential failings of current methods.
Thesis Readers

Justin Lessler, PhD, Advisor
Associate Professor, Department of Epidemiology

William Moss, MD
Professor, Department of Epidemiology

Daniel Salmon, PhD, Committee Chair
Professor, Department of International Health

Christopher Heaney, PhD
Associate Professor, Department of Environmental Health and Engineering
Acknowledgements

This dissertation is the product of an incredible journey, in life, academics, and work, with many challenges but also many joys and successes. I am immensely grateful to everyone who helped me get here, who supported me, and helped to shape the person who I am today.

First, I would be nowhere without my amazing and incredibly supportive spouse, April, who had to put up with a lot during this process, and who helped me keep that “balance” in my life that for many during a dissertation seems like some fabled fairytale. She followed me across the globe to Egypt, then back to Baltimore, and continues to share in the adventures while providing me with a grounded reality and happiness in the quiet times, and given me hope in the dark times. I also must thank our new adventure partner, our daughter Louisa, who joined us 9 months ago. No matter how difficult life gets, she can always put it in perspective and make the worries and stresses melt away. I also would not be where I am without the tremendous support and love from my parents and siblings.

I owe much of my success in this dissertation to Justin Lessler, my advisor and friend, who, despite constant and numerous demands on his schedule, always found time to work through problems, give advice on work or life, or provide opportunities for experience and growth. Justin has provided an endless source of inspiration, through his work and brilliance as a researcher, and his dedication and talents as an educator, advisor, and mentor, and I know I will continue learning from him for years to come. Justin embodies what a mentor should be, and to him I am eternally grateful.

I am extremely grateful to Bill Moss, who, through his expertise in measles, infectious diseases, and epidemiology helped guide this work to be both valid and relevant. Additionally, I am grateful to Derek Cummings who helped introduce me to infectious disease dynamics and helped advise me along with Justin. Thank you to Jessica Metcalf, Matt Ferrari, and Bryan Grenfell, who helped direct this work and
provided insight into the dynamics and details of measles epidemiology and policy. Thanks to Matt Graham for his continued assistance with “maths”. I also am extremely grateful to Luis Mier-y-Teran-Romero, who worked closely with me to develop and conduct my final project, and to Michael Johansson, Allison Taylor Walker, Paul Gastanaduy, Andre Berro, and the U.S. Centers for Disease Control and Prevention for their willingness to collaborate and continuous contribution.

My graduate education experience would also not have been what it was without my fellow classmates and the Infectious Disease Dynamics group. Through the late nights of classwork and exams in the first years, to the daily stresses of developing as researchers in the latter years, these colleagues provided support, entertainment, and inspiration.

This project would also not have been possible without the gracious funding from the Bill & Melinda Gates Foundation.
# Table of Contents

ABSTRACT ................................................................................................................................. II

ACKNOWLEDGEMENTS ................................................................................................................ IV

LIST OF TABLES ........................................................................................................................ XI

LIST OF FIGURES ........................................................................................................................ XIII

INTRODUCTION .......................................................................................................................... 1

1.1 Measles .................................................................................................................................. 1

1.2 Feasibility of measles elimination and eradication .................................................................. 2

1.3 Challenges to achieving measles control and elimination ......................................................... 2

1.4 Current goals and policy efforts ............................................................................................... 4

1.5 Current research efforts .......................................................................................................... 4

1.6 Gaps in current research efforts .............................................................................................. 5

1.7 Measles as a laboratory ........................................................................................................... 5

1.8 Dissertation Overview ............................................................................................................ 6

THE CHALLENGE OF SPATIAL CLUSTERING OF SUSCEPTIBILITY FOR MEASLES ELIMINATION .................................................................................................................. 16

2.1 Abstract ................................................................................................................................. 16

2.2 Introduction ............................................................................................................................ 17

2.3 Approach ............................................................................................................................... 19

2.4 Methods ................................................................................................................................. 21

2.4.1 Analytic Characterization ................................................................................................. 21

2.4.2 Simulation Studies ............................................................................................................ 22

2.4.3 Application ....................................................................................................................... 23

2.5 Results .................................................................................................................................. 24
A.10 Validation Challenges ........................................................................................................... 102
A.11 Additional Supplemental Tables and Figures ........................................................................ 104

SUPPLEMENT TO THE EXISTENCE OF LOCAL CLUSTERING OF NON-VACCINATION AND ITS IMPACT ON OUTBREAK RISK ........................................ 111

B.1 Effective Reproductive Number (R) and Critical Vaccination Threshold (Vc) Calculations .......... 111
B.2 Weighted Cluster Survey τ(r) Calculation ................................................................................ 112
B.3 τ(r) Distributions .................................................................................................................. 114
B.4 g(r) Distributions .................................................................................................................. 114
B.5 Methods for Application to Cluster Survey Data .................................................................... 114
    B.5.1 Moving Window τ(r) Estimation ....................................................................................... 114
    B.5.2 Jittered Bootstrap Estimates ............................................................................................ 115
    B.5.3 Identification of Outlying Clusters .................................................................................. 115
B.6 Comparison of Clustering ...................................................................................................... 115
    B.6.1 Outbreak probability ........................................................................................................ 115
    B.6.2 Regression analyses ......................................................................................................... 116
B.7 Sensitivity Analyses ............................................................................................................... 116
    B.7.1 Household Clustering Exclusion and Sample Weighting .................................................. 116
    B.7.2 Exclusion of Outliers ....................................................................................................... 118
    B.7.3 Impact of Urban-focused Vaccination Efforts ................................................................. 118
    B.7.4 Validation of Consistency of Cluster Survey Results ...................................................... 120
    B.7.5 Generalized Additive Models for Modeled Surface Estimation ....................................... 120
B.8 Extrapolation of Methods to Other Forms of Data ................................................................. 121
    B.8.1 Measuring spatial clustering using modeled surfaces ....................................................... 121
B.9 Additional Supplemental Tables ............................................................................................ 123
B.10 Additional Supplemental Figures .......................................................................................... 127

SUPPLEMENT TO BIOSECURITY IN A GLOBALIZED WORLD: THE CASE OF MEASLES ........................................................................................................ 131
C.1 Mathematical Model for the Importation of Measles-Infected Air travelers into the U.S. ..........131
  C.1.1 Monthly Measles Incidence .......................................................................................132
  C.1.2 Monthly Air Trips from Source Locations to U.S. States ..........................................132
  C.1.3 Infectious Person Days ...............................................................................................132

C.2 Model validation .............................................................................................................133

C.3 Estimated Numbers of Importations and Exportations ...............................................134

C.4 Estimating Relative Risk ................................................................................................136
  C.4.1 Relative risk versus the median risk across all time and locations ...............................136
  C.4.2 Relative risk versus median across time by location ..................................................139

C.5 Incidence and Travel Trends .........................................................................................140

CURRICULUM VITAE ........................................................................................................144
List of Tables

Table 2.1. Analytic $R^*$ and $Vc^*$ estimates from Equation 1 assuming $R0=15$ and using contact distributions derived from Read et al. 2014 ($g_A(r)$), rural Zambia ($g_B(r)$), and Noulas et al. 2012 ($g_C(r)$). 27–29 ............................................................................................................................................... 25

Table 2.2. Outbreak probability from simulations with SIR and individual-based spatially explicit model, assuming $\lambda=0.5$ and $g(r)$. ........................................................................................................................................ 28

Table 2.3. Incidence and estimated effective reproductive numbers for Tanzania from 1998 to 2014. .................................................. 30

Table 3.1. Regional estimates of cumulative relative probability of non-vaccination at $\leq$100m, $\leq$1km, $\leq$5km, and $\leq$10km. ........................................................................................................................................... 46

Table 3.2. Regional estimates of the clustering coefficient ($\phi$), effective reproductive number ($R$ and $R^*$), the adjusted critical vaccination threshold ($Vc$), and probability of outbreak given introduction. .................. 47

Table 3.3. Correlation of measles incidence (per 100,000) with the effective reproductive number ($R$) and probability of outbreak. Assumes $R0=15$. ........................................................................................................................................ 48

Table 4.1. Importations and incidence from nine epidemic periods during 2006-2015. ................................................................. 66

Table 4.2. Overall estimated relative risks of measles importation by state, the top 20 states during the study period (January 2006 – December 2015), sorted by state mean population. Relative risk uses the median risk across all states as the reference. See Appendix C, Table C.2 for full table. .................. 68

Table 4.3. Overall estimated relative risk of exporting measles by European countries, the top 20 countries during the study period (January 2006 – December 2015), sorted by country mean population. Relative risk uses the median risk across all countries as the reference. See Appendix C, Table C.3 for full table. ...................................................................................................................................................... 71

Table A.1. Non-vaccination clustering adjustment factor estimates for the three contact distance distributions and four levels of clustering of non-vaccination .......................................................................................................................... 104

Table A.2. Non-vaccination clustering-adjusted effective reproductive number estimates for 85, 90, and 95% effective vaccination coverage at the three contact distance distributions and four levels of non-vaccination clustering. Assumes $R0=15$. ...................................................................................................................................................... 104

Table A.3. Non-vaccination clustering-adjusted critical vaccination thresholds at none, low, medium, and high clustering with vaccine efficacy (VE) of 100% and 95%, with $\lambda=0.5$. ........................................................................................................................................ 105

Table A.4. Non-vaccination clustering-adjusted critical vaccination thresholds at none, low, medium, and high clustering with vaccine efficacy (VE) of 100% and 95%, with $\lambda=0.5$. ........................................................................................................................................ 105

Table A.5. Incidence and estimated effective reproductive numbers for Tanzania from 1998 to 2014. .. 106
Table B.1  Exclusion criteria and numbers of surveys included/excluded with each criterium. ............ 123

Table B.2. Estimated clustering adjusted critical vaccination thresholds ($V_c$) for the most recent survey from each of the 33 included countries. $R_0$=15. ........................................................................................................ 124

Table B.3  Estimated effective reproductive numbers ($R_*$) for the most recent survey from each of the 33 included countries. $R_0$=15. ........................................................................................................ 125

Table B.4. Regression model likelihood estimates and likelihood ratio comparing clustering adjusted and unadjusted models........................................................................................................... 126

Table C.1. Model Parameters................................................................................................................................................................................................................................. 133

Table C.2. Validation model likelihoods and likelihood ratio compared to the null model. .............. 133

Table C.2. Overall estimated relative risks of measles importation by state during January 2006 – December 2015, sorted by state mean population. Relative risk uses the median risk across all states as the reference.................................................................................................................................................................................................................................................................................................................. 138

Table C.3. Overall estimated relative risk of exporting measles by European countries, during January 2006 – December 2015, sorted by country mean population. Relative risk uses the median risk across all countries as the reference.................................................................................................................................................................................................................................................................................................................. 139
List of Figures

Figure 1.1. U.S. Measles Cases by Year. Source: U.S. Centers for Disease Control. ........................................ 3

Figure 2.1. Synthetic populations and corresponding $\tau(r)$ functions with no, low, medium, and high spatial clustering of susceptible individuals. Vaccinated (grey) and unvaccinated (red) individuals are represented as points within the plotted space. Individuals are evenly spatially distributed across the space. ............................... 23

Figure 2.3. Critical vaccination threshold, as defined by the overlapping densities of the $\tau r$ and $g r$ distributions. Plotted curves are $\tau r$ where $r$ is that which corresponds to the value of $1 - Gr$, and color corresponds to the resulting $Vc$ when $R0=15$. Three levels of clustering ($\theta=.25, .50, 1.0$) and $529 g(r)$ distributions, with both $\alpha$ and $\beta$ range=0.1-10. ................................................................. 27

Figure 2.4. (A) Probability of an outbreak with increasing clustering ($\phi$) and effective reproductive number ($R$). This assumes an introduction has occurred and an outbreak is defined as 5% of susceptible individuals becoming infected. (B) Probability ratio (PrR) of an outbreak with increasing clustering ($\phi$) compared with homogeneous ($\phi=1$). ........................................................................... 29

Figure 2.5. (A) Vaccination coverage among children aged 1-5yr by district and survey cluster in Tanzania during the 1999 and 2010 Demographics and Health Surveys. The DHS are cross-sectional data on individuals sampled using cluster randomized sampling. GPS location is only available for clusters, not individuals. (B) Empirical $\tau(r)$ functions and fitted exponential for Tanzania and Zambia, using DHS cluster data. .................................................................................................................... 31

Figure 3.1. Estimated $\tau r$ fits for five countries using DHS data. Forms of the fitted distributions follow an exponential decay function. ........................................................................................................ 45

Figure 3.2. Estimated vaccination coverage needed above the 95% target to maintain herd immunity, DHS 2000-2015. .................................................................................................................. 46

Figure 3.3. Probability that the needed critical vaccination threshold ($Vc*$) is greater than the target of 95%, DHS 2000-2015. ........................................................................................................ 47

Figure 4.1. (A) Measles incidence in Europe and (B) reported measles case importations from twenty-eight European countries into the U.S. during January 2006 – December 2015. ...................................................... 65

Figure 4.2. Measles importation risk ratios by state or state group within the United States during January 2006 –July 2016. (a) Overall risk ratio across time and (b) per capita risk ratio across time. Risk ratios are calculated as the cumulative risk across time for each state as compared to the median risk across all states. ................................................................. 67

Figure 4.3. Relative risk of measles importation by month among U.S. states and state groups, January 2006 – July 2016, sorted by population. Relative risk is calculated against the median risk of importation across all states and all months. ................................................................. 70
Figure 4.4. Relative risk of measles exportation by month from European countries to the United States, January 2006 – July 2016, sorted by population. Relative risk is calculated against the median risk of exportation of Austria. Austria was used the reference country due to having median overall risk among countries with multiple months of non-zero exportation risk.

Figure 4.5. Correlation between reported numbers of importation and predicted, by source country during 2008, 2011, and 2015.

Figure A.1. Contact distance probability distribution g(r) derived from contact diary and interview acquired contact data from China (g_C(r)), GPS logger data from Zambia (g_B(r)), and cell phone data from the U.S. (g_U(r)). From these data, gamma distributions were fit to produce parameterized approximations of each distribution.

Figure A.2. Example of pre-clustering τr in a synthetic population and the resulting τr after the vaccination clustering procedure. Red line is the pre-defined parametric τ(r) and the black line is the fitted form.

Figure A.3. Effect of increasing spread of the τ(r) when θ is held constant. (A) Parametric τ(r) distributions with θ=1.5. (B) Associated resulting R * as λ increases. Assumes a constant stable g(r).

Figure A.4. gr decay rate with exponential distribution. (A) Exponential PDF forms of gr with rate parameters 0.1, 0.2, 0.5, 1.0. (B) Exponential rate versus R for varying levels of spatial clustering.

Figure A.5. Impact of τr rate parameter, λτr, on the clustering coefficient, φ. As λτr increases and the τr function becomes steeper, resembling clustering centered within households, φ approaches 1 and the R * estimate will approach the unclustered R estimate.

Figure A.6. Empirical CDFs of spatial SIR simulation final sizes compared to the mean ECDF of homogeneous SIR simulations (A), and homogeneous final size ECDFs vs mean ECDF of spatial simulations. Spatial SIRs result in a mean final size around half that of homogeneous SIRs, but a similar, though slightly lower probability of outbreak.

Figure A.7. Individual initial R for a synthetic population with medium clustering and 95% vaccination coverage. (A) Clustering of unvaccinated individuals in the population space (vaccinated individuals are not shown; individuals are distributed evenly in the space). Unvaccinated individuals are colored by initial R. (B) Distribution of initial R estimates for individuals in the population.

Figure A.8. (A) Probability of an outbreak with increasing clustering (φ) and effective reproductive number (R). This assumes an introduction has occurred and an outbreak is defined as 5% of susceptible individuals becoming infected. (B) Probability ratio (PrR) of an outbreak with increasing clustering (φ) compared with homogeneous (φ=1).

Figure A.9. Correlation between φ and AUC of 1 − G(r) versus τ(r) curves. All three levels of clustering (θ=0.25, 0.5, 1.0) have perfect correlation with AUC (Pearson’s r=1.0). For each combination of τ(r) and g(r), φ = AUC + 1.

Figure A.10. The association between R0 and Vc * at the four defined levels of clustering and three gr distributions. As clustering increases, the required vaccination coverage to maintain R=1 increases.
low $R_0$, this increase due to clustering is much greater than at high $R_0$. (A) Linear $Vc$ * scale. (B) Logarithmic $Vc$ * scale.

Figure A.11. Association between $R_0$, clustering maximum ($\theta$ parameter of $\tau(r)$), and the critical vaccination threshold ($Vc$).

Figure A.12. Sensitivity analyses of the varied impact of $\tau x$ parameters ($\theta$ and $\lambda$) and $g x$ parameters (shape and scale) on the resulting $Reff$ estimate. $\tau x$ follows an exponential distribution characterized as $\tau x = \theta e - \lambda x + \psi$, while $g x$ follows a standard gamma distribution parameterized by shape and scale parameters. Regardless of the $g x$ parameters, as $\theta$ increases, $Reff$ increases (A and B).

Figure A.13. Relationship between $R_0$, $\phi$, and $Vc$.

Figure A.14. Analytic $R^*$ estimates for measles assuming $R_0=15$ and using contact distributions derived from Read et al. 2014 ($g_A(r)$) and Noulas et al. 2012 ($g_C(r)$). As clustering increases, $R$ increases; however, this difference decreases as vaccination coverage increases.

Figure B.1. Empirical and fitted $\tau(r)$ functions including and excluding household-based clustering and sample weighting.

Figure B.2. Empirical $\tau(r)$ results with and excluding four outlying clusters. No obvious difference was found.

Figure B.3. Results of simulated urban rural disparities in vaccination coverage.

Figure B.4. Cluster locations of simulated clusters and true sampling clusters from Zambia.

Figure BS.5. Results from GAM smoothing.

Figure B.6 Modeled surface estimation of clustering, Togo 2013.

Figure B.7. Case study countries.

Figure B.8. (a) Unadjusted and (b) spatially clustering adjusted effective reproductive estimates, using Demographics and Health Survey data and WHO/UNICEF vaccination coverage estimates.

Figure B.9 Relative probability of outbreak, clustering adjusted probability versus unadjusted probability estimates.

Figure C.1. Reported measles incidence, Europe, June 1999 to April 2016.

Figure C.2. Estimated measles cases imported to the United States from Europe, by country of origin, June 1999 to April 2016.

Figure C.3. Estimated measles cases imported to the United States from Europe, by U.S. state of destination, June 1999 to April 2016.

Figure C.4. Relative risk of measles importation monthly by state. Here the relative risk is calculated as the estimated risk during each month divided by the median estimated risk of importation across the study
period for that state. This figure highlights the periods of elevated and decreased risk of measles importation.

Figure C.5. Monthly travel fraction from all of Europe to the United States, 2005-2015.

Figure C.6. Total monthly measles incidence in Europe, 2005-2015. (A) absolute monthly incidence is highly variable from one year to the next, governed by the occurrence of outbreaks. However, (B) monthly incidence fraction is somewhat consistent, as Europe typically experiences seasonality in measles incidence, though substantial variation does occur (i.e. 2005, 2007, 2013).
CHAPTER 1

Introduction

1.1 Measles

Alexander Langmuir first described the importance of eradicating measles in 1962.\(^1\) A decade and a half later, in 1978, the Centers for Disease Control and Prevention (CDC) set the first goal of measles elimination for the United States for 1982.\(^2,3\) Although it took two more initiatives and 22 years, the U.S. achieved measles elimination in 2000, meeting the elimination criterion of the absence of continuous disease transmission for >12 months.\(^3,4\) Worldwide elimination was first initiated in 1994 when the Pan-American Sanitary Conference set the goal to eliminate measles from the Western hemisphere.\(^5\) All six WHO regions have now set goals for measles elimination by 2020, and in 2016 the Region of the Americas became the first to accomplish theirs.\(^6,7\) Unfortunately, with only three years remaining, accomplishing worldwide measles elimination and eradication is likely much further away than 2020.

The burden of measles, as with other infectious diseases, has dramatically decreased during the last century. During the first half of the 20\(^{th}\) century, before the measles vaccine, measles deaths were already dropping. In the United States, measles deaths declined from 5,000 to 500 annually.\(^8,2\) This pre-vaccine progress coincided with dramatic reductions in illness and mortality due to other infectious diseases and resulted largely from improvements in diet, sanitation and hygiene, health care and administration, and medical breakthroughs like antibiotics and gamma globulin.\(^9\) Despite the reductions in measles deaths, measles continued to infect nearly every person on the planet until John Enders and colleagues licensed the first vaccine in 1963, making prevention of measles virus infection possible. Worldwide, measles
vaccine took much longer to become widespread – in 1980, measles caused an estimated 2.6 million deaths.\textsuperscript{10} Thanks in part to efforts spurned by the Millennium Development Goals to reduce childhood mortality and the Measles and Rubella Initiative to increase vaccination coverage, in 2016 this had dropped to 90,000 deaths from measles in children.\textsuperscript{11,12} Despite the successes, however, some estimate 20 million cases continue to occur annually.

\textbf{1.2 Feasibility of Measles Elimination and Eradication}

The case for measles eradication is one with great merit and potential feasibility. Measles has long been a cause of widespread disease and death, and still remains one of the top causes of child mortality five decades after the vaccine was developed.\textsuperscript{13} The measles vaccine provides a powerful tool to prevent these cases and eliminate transmission, with an estimated efficacy of 85% with one dose (when given at nine months of age) and 97% with two, and this immunity is not believed to wane substantially, unlike pertussis and others.\textsuperscript{14} As Sencer et al. argued in 1967, just as success was achieved with smallpox, vaccinating enough individuals would cause the disease to die out, thus creating herd protection.\textsuperscript{15} Without environmental or zoonotic reservoirs, once transmission among humans is interrupted, measles virus cannot persist and will eventually be eradicated. Measles elimination is defined as the absence of continuous transmission in a defined geographical area for at least twelve months; eradication is defined as the permanent reduction to zero of the worldwide incidence of infection by measles virus.

\textbf{1.3 Challenges to Achieving Measles Control and Elimination}

Despite checking all the boxes for elimination and eradication feasibility and the incredible progress already made, measles has one dominant challenging characteristic: its extreme transmissibility. This transmissibility is quantified by the basic reproductive number, $R_0$, which is the average expected number of secondary infections due to a single case, in a completely naïve population. For measles, this is often estimated to be 12-18, though a recent review indicates it is likely context dependent and could range
dramatically; in contrast for influenza, $R_0=1.5$, and for smallpox $R_0=3.5-6$. As a result, measles vaccination coverage must be extremely high to reach the critical vaccination threshold and prevent disease through herd immunity, and vaccination coverage must be maintained with each birth cohort. Because of this transmissibility, measles requires substantially more effort than previously envisioned, and we are still many years away from global measles eradication.

Recent outbreaks in regions where measles control should be high exemplify these challenges. Even with elimination status, high vaccination rates, and extensive spending and public health efforts by health authorities, the Americas continue to face measles outbreaks. In the U.S., outbreaks have increased in frequency and size since 2000, and often are focused in populations where vaccine uptake is low (Figure 1.1). Canada and Brazil have also experienced increases in cases, including recent outbreaks of over 700 in both countries.21,22

Europe has experienced even greater obstacles to elimination efforts. In 2002, the WHO European Region established a goal of measles elimination by 2010, yet over the last decade, Europe has faced repeated measles outbreaks, several with thousands of cases.25 From 2006-09, Switzerland experienced an outbreak with 4,415 confirmed cases.26 In 2010, Bulgaria experienced an outbreak of 24,364 cases and 24 deaths.26

Figure 1.1. U.S. Measles Cases by Year. Source: U.S. Centers for Disease Control.
The following year, Europe experiences an even larger outbreak centered in France, with around 15,000 cases in France, and more than 30,000 cases across Europe. In 2016-17 measles hit Europe once again in Romania, with 7,500 cases and 31 deaths through June 2017.

1.4 CURRENT GOALS AND POLICY EFFORTS

Despite these challenges and setbacks, the successes of the U.S. and the Americas demonstrate that establishing elimination is possible, and with the right goals, tools, and efforts, it can be accomplished in the near future. All six WHO Regions have established or reaffirmed commitments to achieving measles elimination by 2020, establishing a de facto eradication goal for measles. The World Health Organization’s Global Measles and Rubella Strategic Plan provides a rubric and driving force with which action can be directed to accomplish these goals and sustain progress that is made.

1.5 CURRENT RESEARCH EFFORTS

Measles has proven to be a disease that is simultaneously very easy and an extremely difficult to prevent. The vaccine is cheap, highly efficacious, and readily available. However, vaccinating enough people to interrupt transmission through herd immunity requires tremendous efforts, and maintaining those levels can be even harder. While recent and current work to improve our understanding of measles risk and how to reduce it is extensive, this single challenge of high transmissibility and high vaccination coverage needs largely governs current research efforts. These efforts include ongoing work to map vaccination coverage in different parts of the world and to estimate burden of disease. There are also research efforts to improve the ways that we calculate the rates of vaccination coverage and others to determine at what ages and intervals to conduct vaccination campaigns. Other research is focused on how to understand who are not getting vaccinated and why.
Even with these efforts, requiring substantial financial investment and effort from those involved, we continue to falter with measles control. Each of these efforts contributes a piece to the puzzle, and measles is a very challenging puzzle with an unknown number of pieces.

1.6 GAPS IN CURRENT RESEARCH EFFORTS

Even with all the current and past research, outbreaks continue to occur and the possibility of reaching worldwide measles eradication by 2020, only 3 years from the time of this writing, looks bleak. Thus, the need remains to add to the collage of knowledge and tools, particularly where critical gaps remain. With this dissertation, I set out to provide improved methods for calculating and understanding the risk of measles transmission in populations at different stages of measles elimination, from those with improving control like many countries in Africa, to countries struggling to maintain control and establish elimination like in Europe, and finally countries like the U.S. where measles elimination has been established, yet outbreaks continue to occur.

1.7 MEASLES AS A LABORATORY

One final factor motivating this work is that measles serves as an excellent laboratory with which to develop theories, frameworks, and methods that can then be applied to other diseases. The reasons for this are some of the same reasons that measles is an excellent candidate for disease eradication: everyone infected is symptomatic, both infection and vaccination provide lifelong immunity, the measles virus is monotypic and its dynamics are simple and consistent, and the measles vaccine is highly effective. Starting with such a “simple” disease lets us initially avoid some of the complexities that complicate work with other diseases, including multiple competing and complicating serotypes like influenza and dengue viruses, waning immunity like mumps, latent or asymptomatic infections like tuberculosis and chikungunya, or non-human reservoirs. To the simple, demonstrably valid methods developed with measles, we can then add the complexity needed to apply these to more complex diseases.
1.8 DISSERTATION OVERVIEW

The first aim of this dissertation (Chapter 2) was to develop a framework for incorporating heterogeneity in susceptibility, which we refer to as “clustering”, into metrics used to understand transmission and outbreak potential (i.e. the effective reproductive number, $R$) and the level of vaccination needed to prevent such transmission (i.e. the critical vaccination threshold, $V_c$). The impetus of this was previous work that I was undertaking to estimate the proportion of the population considered “inaccessible” to repeated vaccination efforts (unpublished). That work was an extension of previous work by Lessler et al. and used a Bayesian probabilistic framework to estimate the proportion of children aged $\leq 5$ years that remained unvaccinated despite multiple vaccination activities for which they were eligible. In short, an individual is considered inaccessible if they remained unvaccinated at the time of the survey, even though multiple vaccination opportunities occurred for which they were eligible (i.e. routine vaccination or supplementary immunization activity (SIA)). Understanding the size of this inaccessible population can be critical for accurately estimating the proportion of individuals truly vaccinated in a population, and the disregard of this quantity results in overestimation of the proportion vaccinated.

While working to quantify inaccessible populations, we began to wonder who these people were, and what impact their existence has on the total population in terms of outbreak risk, beyond the implications for vaccination estimates. This became particularly critical because we were finding that these individuals were either spatially or social clustered, with the lack of vaccination often being the result of health care access (spatial clustering) or belief (social clustering). We hypothesized that this clustering of non-vaccination, thus susceptibility, violates the assumptions governing the methods we use to make vaccination targets: in a population with susceptibility clustering, contact between unvaccinated individuals is not random, but rather unvaccinated individuals are more likely to be in contact with other unvaccinated individuals than in a population without clustering. Through this aim, we develop a way to quantify the impact of increased contact between susceptible individuals, and incorporate this into
adjusted metrics ($R^*$ and $V_c^*$) that are accessible and can be applied and understood readily by the global public health community.

As no new method is useful without the ability to apply it, for Aim 2 (Chapter 3), we endeavored to find a way to consistently apply these methods without having to collect data on vaccination status and residence location from every individual in the population. The initial intention of the clustering adjusted $R^*$ and $V_c^*$ was a set of statistics that could be applied to large populations, such as entire countries, to assess the transmission risk and vaccination need in those populations, just as their unadjusted counterparts are commonly used. However, in contrast to their unadjusted counterparts that only require the $R_0$ of the disease and an estimate of the proportion vaccinated, $R^*$ and $V_c^*$ require the addition of a population-wide assessment of susceptibility clustering. Without the ability to perform this assessment, our methods become little more than theory.

Fortunately, the Demographics and Health Surveys (DHS) provide an excellent available source of comprehensive survey data from more than 90 countries, and most surveys in the last 15 years include geospatial information. We developed methods to use these DHS data to estimate susceptibility clustering, which we then applied to the methods from Aim 1 to produce adjusted $R^*$ and $V_c^*$ estimates. We examined how these estimates corresponded to recent history of measles outbreaks and transmission in these countries. While susceptibility clustering did not have a major impact on the risk of outbreaks across the included countries, this could be expected by the generally low vaccination coverage in these countries. As we found that the impact of clustering increases as vaccination increases, had we surveys from countries in Europe or the Americas, we likely would have observed more dramatic increases in measles incidence among countries with high clustering. Unfortunately, the availability of standardized fine-scale spatially-explicit data like DHS is limited for countries in these regions.
Clustering of susceptibility within a population, whether spatial, social, or otherwise, amplifies the risk of an infectious disease outbreak, and through applying the methods developed in Aims 1 and 2, we can accurately assess the extent and impact of such clustering. Through quantifying clustering with the $\tau(r)$ function, we can estimate an adjusted effective reproductive number and critical vaccination threshold, statistics that enable us to adjust vaccination goals and implement targeted vaccination campaigns.

Through clustering adjustment, we can understand the risk of outbreaks within a population assuming introduction is occurring. For countries in regions that with endemic measles transmission, and thus the potential for frequent measles virus introductions from neighboring countries, susceptibility clustering can be critical to outbreak risk. However, just as the best firewood requires a spark to start a fire, within regions where measles has been eliminated, like the Americas, importations of measles from elsewhere are first required to spark outbreaks. To understand the risk of outbreaks in the United States, for example, we must determine the risk of these importations. Once we quantify the risk of importations, then we can apply methods like the clustering framework to assess the risk of an outbreak once a measles case has been introduced.

For this final Aim (Chapter 4), the objective was to create a way to estimate the numbers or risk of importations of measles into the U.S. through air travel. Through incidence and travel data from Europe, we estimated the relative risks of importations by destination states, time periods, and source countries. These estimates corresponded well with true reported importation data provided by the Centers for Disease Control and Prevention.

We now exist in a new elimination landscape for measles virus where the assumptions of homogeneity make our old methods invalid, and simply pushing to continue increasing vaccination coverage overall no longer provides the benefits it once did. To fill this void, we have developed three new tools to aid in understanding and assessing the risk of measles virus transmission and outbreaks during this new era of
elimination. Through accounting for heterogeneity in populations and risks of importations from endemic to eliminated countries, we can begin to think about more effective strategies for reaching and maintaining elimination goals around the world. From this work, it is now evident that standard national vaccination targets are no longer sufficient – with the increased subnational and global connectedness of today, local vaccination targets are essential in order to maintain high enough vaccination coverage to prevent measles virus transmission and outbreaks and eventually establish measles virus eradication worldwide.
References


PART II

SPATIAL CLUSTERING OF SUSCEPTIBILITY AND ITS IMPACT ON OUTBREAK POTENTIAL
CHAPTER 2

The challenge of spatial clustering of susceptibility for measles elimination

Shaun Truelove, Matthew Graham, William Moss, Jessica Metcalf, Bryan Grenfell, Matthew Ferrari, and Justin Lessler

2.1 ABSTRACT

Vaccination coverage goals, including the WHO goal of 95% for measles, are based in part on estimates of the critical immunity threshold ($V_c$): the percentage of the population that must be successfully vaccinated to prevent sustained epidemics. Traditionally, estimates of $V_c$ assume populations mix evenly, which we know to be untrue. If susceptible individuals preferentially contact one another, communities may remain vulnerable to epidemics even when vaccination coverage goals are met at larger spatial scales. Here we present a simple method to estimate the effective reproductive number, $R$, in a population with only vaccine derived immunity, that captures the impact of spatial clustering of non-vaccination combined with spatially assortative mixing on $R$ and $V_c$. For measles, assuming $R_0=15$ and 95% vaccination coverage, high clustering of non-vaccination would increase $R$ from 0.75 to 1.31 and the probability of an outbreak after a single introduction from <1% to 23% (95% CI: 0-1% and 22-24%). This reflects an increase in $V_c$ from 93% to 96%, We illustrate our approach using Demographic and Health Survey data from Tanzania, and show how clustering of non-vaccination potentially contributed to the continued endemic transmission of measles during the last two decades. Our approach demonstrates why high vaccination coverage at the national level sometimes fails to achieve measles elimination, and can help countries appropriately adjust vaccination targets when there is clustering of unvaccinated individuals.
2.2 INTRODUCTION

Despite declared goals of measles elimination by 2020 by all six WHO regions and substantial improvements in global measles vaccination coverage (from 73% to 85% for a single dose between 2000 and 2015), large outbreaks continue to occur. These outbreaks are not only in countries with low measles vaccination coverage but also in countries with perceived measles control throughout Europe (e.g., France, Romania), Africa (e.g., Malawi), and Asia (e.g., the Philippines). While there are multiple factors contributing to these outbreaks, their continued occurrence reflects the considerable logistical, political and social challenges to measles elimination, and highlights a need for novel approaches as we enter a new stage in our battle against this disease.

WHO’s vaccination coverage goals of 95% for the first and second doses of measles-containing vaccine (MCV) in the Global Measles and Rubella Strategic Plan derive, in part, from theoretical estimates of the immunity level needed to prevent ongoing transmission of a pathogen. These theoretical estimates are based on the basic reproductive number \(R_0\), or the number of people a single infected person is expected to infect in a fully susceptible population, which for measles has been estimated to be between 12 and 18. An important rule of thumb for setting vaccination targets is that if countries can successfully immunize, via vaccination, a portion of their population equal to \(1-(1/R_0)\), the population will have sufficient herd immunity to avoid significant epidemics if the pathogen is introduced into the population. Hence, for measles, 91-94% of the population would need to be immune to achieve this “critical immunity threshold” \(V_c\).

Although useful for setting rough vaccination coverage targets, this rule of thumb is based on simple models of disease spread that ignore population structure. That is, they assume an infected individual is equally likely to have a potentially infectious contact with all other individuals in that population. If this assumption is violated, particularly if infected individuals are more likely to contact susceptible
individuals than immune ones, then the predictions from these models will start to break down. If vaccination coverage is heterogeneous by geographic location or social group, the vaccination process itself can lead to a violation of this critical assumption.\textsuperscript{10-12}

The existence of heterogeneities in vaccination coverage and association with disease incidence have been demonstrated across the globe. Within African countries, heterogeneity in measles vaccination coverage is associated with distance from major population centers.\textsuperscript{13} In the United States, spatial clustering of nonmedical exemptions to school immunization requirements in census tracts in Michigan overlapped with clusters of pertussis incidence.\textsuperscript{14} Similar results have been found in California.\textsuperscript{15} In 2014, an outbreak of 380 confirmed measles cases occurred in a large spatially clustered Amish community in Ohio, the largest such outbreak in recent U.S. history.\textsuperscript{16} While 92\% of Ohio kindergarteners are vaccinated against measles, only 11\% of cases from this outbreak had previously been vaccinated, indicating high spatial clustering of non-vaccination in this Amish community.

The impact of these heterogeneities has been previously theorized. Work by Funk \textit{et al.} has demonstrated the importance of maintaining immunity particularly in the 5- to 9-year-old age group due to age-specific heterogeneity in contact.\textsuperscript{17} Other studies have demonstrated that the assumption of homogeneous contact in network structures can underestimate $R_0$.\textsuperscript{9,18,19} Using network theory, one study found that to eliminate disease transmission, a higher level of vaccination coverage is needed than previously believed.\textsuperscript{19} These studies, however, relied on contact networks and theoretical scenarios, and do not extend easily to application with data available for real populations. This is the gap we intend to fill.

Simple, broadly applicable methods that provide guidance on the critical vaccination threshold where heterogeneity exists may be useful in planning measles control activities as regions approach elimination. Here we present a method for estimating the effective reproductive number and critical vaccination threshold in the presence of spatial heterogeneity in susceptibility. Using Demographic and Health Survey
(DHS) data from Tanzania, we demonstrate its application and examine how spatial clustering of vaccination may have contributed to continued measles virus transmission while vaccination coverage, measured at the national level, was presumed to be high enough to achieve the critical vaccination threshold and maintain herd immunity.

2.3 APPROACH

Estimates of a disease’s outbreak potential and the vaccination coverage needed to eliminate it are based on the basic and effective reproductive numbers, $R_0$ and $R$. Like $R_0$, $R$ is a measure of how many people a single infectious individual is expected to infect, but unlike $R_0$, $R$ reflects a population with some level of immunity. If every individual is equally likely to interact with every other individual in a population, the probability of a contact being susceptible is equivalent to the proportion susceptible in the population. If immunity is derived exclusively from vaccination, this will be $1 - v$, where $v$ is the proportion who successfully develop protective immunity following vaccination. Hence:

$$ R = R_0 (1 - v) = R_0 \times \Pr(\text{contact is susceptible}) \quad [1] $$

Here we focus on maintaining protection from measles through successful vaccination. To maintain simplicity for demonstration purposes, we assume there is no naturally acquired immunity. Hence, throughout, susceptibility is synonymous with not being successfully vaccinated.

We can incorporate heterogeneity into Equation 1 by appropriately adjusting the probability that a contact is susceptible based on a measure of the heterogeneity. Here we focus the synergistic effect of spatial heterogeneity in vaccination and the fact that individuals are more likely to have infectious contacts with those living near their home than those living farther away.\textsuperscript{20-23} Hence, the more spatially clustered susceptibility is, the higher the probability a potentially infectious contact will occur between two susceptible individuals. Such spatial clustering of susceptibility is a common and well-documented
phenomenon, and can occur as a result of spatial heterogeneities in access to vaccination or vaccination refusal. We quantify the tendency of susceptible individuals to cluster together using the statistic $\tau(r)$, defined as the relative probability that an individual living at spatial distance $r$ from a susceptible individual is also susceptible versus the probability that anyone in the population is susceptible, corrected for population density. Hence:

$$\Pr(\text{contact is susceptible } | \text{contact lives at distance } r) = (1 - v) \frac{1 - v(r)}{1 - v(\infty)} = (1 - v)\tau(r) \quad [2]$$

where $1 - v(\infty)$ is equivalent to $1 - v$ for the full population. If we know $g(r)$, the probability that a contact occurs at distance $r$ given that a contact did occur, we can calculate the number of infectious contacts that a susceptible individual is expected to make (i.e., the adjusted effective reproductive number):

$$R^* = \beta c (1 - v) \int_0^{\infty} g(r)\tau(r) \, dr = R_0 (1 - v) \int_0^{\infty} g(r)\tau(r) \, dr = R_0 (1 - v)\phi \quad [3]$$

where $c$ is an individual’s expected number of potentially infectious contacts over the entire course of their infectious period, $\beta$ is the per contact transmission probability (note that $R_0 = \beta c$), and $\phi$ is a clustering adjustment factor for $R$ (i.e., $R^* = R\phi$).

Assuming the susceptibility clustering and contact distance remain constant, thus at a constant $\phi$, we can define an adjusted critical immunity threshold, $V^*_C$, as the level of immunity (or successful immunizing vaccination) that must be maintained in the population to achieve herd immunity.
2.4 METHODS

For these analyses, we assumed that the relative probability that an individual at distance \( r \) from an susceptible individual is also susceptible follows an exponential decay function, of the form \( \tau(r) = \theta \exp(-\lambda r) + 1 \), where \( \theta \) is the maximum relative probability of a contact being susceptible and \( \lambda \) characterizes the rate of decay in this ratio. Again, we are assuming successful vaccination to be the only source of immunity in the population. The distance from the home at which contacts occur follows a gamma probability distribution, \( g(r) \sim \text{gamma}(r, \alpha, \beta) \), where \( r \) is the distance of the contact. Fitting \( g(r) \) to empirical and simulated contact data was done through mathematical optimization using the Nelder-Mead method. We note that the method could be applied using other parametric and empirical forms of \( \tau(r) \) and \( g(r) \).

2.4.1 Analytic Characterization

To illustrate our approach, we considered a measles-like disease with four levels of spatial clustering of susceptible individuals: none (\( \tau(r) \) parameters: \( \theta=0, \lambda=0.5 \), low (\( \theta=0.25, \lambda=0.5 \), moderate (\( \theta=0.5, \lambda=0.5 \)), and high (\( \theta=1.0, \lambda=0.5 \)). We applied three contact distributions based on real data from distinct populations and locations, for each of which a gamma distribution was parameterized: \( g_A(r) \), based on data from southern China and characterized by highly local contacts (\( \alpha=0.238, \beta=0.162 \); \( g_B(r) \), based on data from rural villages in Zambia, where contacts are either highly local or widely dispersed (\( \alpha=0.086, \beta=0.042 \)); and \( g_C(r) \), based on cell phone data from the U.S., where contacts are generally more dispersed, representing a commuting culture like much of the U.S. (\( \alpha=0.701, \beta=0.076 \)) (see supplemental Figure A.1 and Text A.4 for full details). Additionally, we considered three levels of
successful vaccination coverage, 85, 90, and 95%, and, for simplicity, assumed $R_0=15$ for measles when applying these distributions to Equations 3 and 4 above.

To understand the impact of different $g(r)$ and $\tau(r)$ distributions, we also conducted sensitivity analyses to a wide range of the parameters for each (Text A.7). Additionally, we examined the effect of overlap between the distributions by calculating curves of $g(r)$ quantiles versus $\tau(r)$ values at $r$ of the quantiles, using a broad parameter range for both gamma $g(r)$ ($\alpha$ and $\beta = 0.1$-$10.0$) and $\tau(r)$ ($\theta = 0.25, 0.5, 1.0$).

### 2.4.2 Simulation Studies

To validate our analytic calculations, we performed simulation studies using spatially-explicit synthetic populations and compared these results to those from simulations of well-mixed homogenous populations with the calculated $R^*$. We considered the same scenarios as before (vaccination: 85, 90, and 95%; clustering: none, low, medium, and high) and used $g_A(r)$ for all simulations. Ten populations of 100,000 individuals were stochastically generated for each level of vaccination coverage and clustering, with additional population sizes and densities examined in the sensitivity analyses (Figure 2.1, Appendix A.7). We assumed measles virus characteristics with $R_0=15$ and a serial interval of two weeks. With each population, we performed 10,000 simulations of spatially-explicit individual-based measles virus transmission following an introduction event. $R^*$ was calculated analytically for each scenario using the pre-defined target $\tau(r)$ and the exact $g(r)$, which was calculated from each simulation result.
Figure 2.1. Synthetic populations and corresponding τ(r) functions with no, low, medium, and high spatial clustering of susceptible individuals. Immune (grey) and susceptible (red) individuals are represented as points within the plotted space. Individuals are evenly spatially distributed across the space.

Using the resulting $R^*$ estimates, we performed stochastic, homogeneous mixing SIR simulations for each combination of clustering level and vaccination coverage, with $R$ was set at $R^*$. From both spatial and homogeneous mixing simulations, we estimated the final size and probability of an outbreak occurring; to do this we set a threshold of 5% of the susceptible population becoming infected to indicate an outbreak occurred.

2.4.3 Application

While explicitly measuring the spatial clustering of non-vaccination and distribution of contacts is possible, cost and logistical constraints mean doing so in a timely manner is likely infeasible. Demographics and Health Surveys (DHS) provide a widely available alternative that can be used to measure clustering in vaccine uptake at a coarse scale. The DHS are cluster randomized surveys, conducted every 3-5 years in over 90 countries, and more recent surveys include GPS coordinates for
sampling cluster locations. These data specifically include reported and documented individual-based vaccination, and provide a snapshot of vaccination coverage through a combination of routine and supplemental immunization activities for children under 6 years. We used Tanzania DHS data to demonstrate application of our approach.

To account for the characteristics of DHS surveys, and to produce a final parametric $\tau(r)$ function with confidence intervals, we estimated $\tau(r)$ using a modified cluster-based method employed during 10,000 Monte Carlo iterations with bootstrapping. Because cluster locations are displaced for de-identification purposes, during each iteration, clusters locations were jittered based on urban/rural status (urban: $\leq 2$km, rural: $\leq 5$km), and clusters included were bootstrapped. From this jittered and bootstrapped dataset, we calculated the $\tau(r)$ function using a modified cluster-based method, which calculates $\tau(r)$ through cluster pairs, rather than individual pairs, and applies the normalized sampling weights provided by DHS (see Text S5). This calculation also employed a moving window approach for $r$, where $\tau(r)$ was calculated for overlapping $r$ windows (e.g. $r = 0$-1km, 0.5-1.5km, 1-2km, ...). The exponential decay form was then fit to the calculated empirical $\tau(r)$ from each iteration using non-linear least squares with inverse variance weighting. Means and confidence intervals of fitted parameters from each iteration are used as the final fitted $\tau(r)$ for the DHS data (see Text A.8).

2.5 RESULTS

2.5.1 Analytic Analysis

To assess the impact of clustering and contact patterns on risk of disease transmission, we calculated three statistics: the clustering adjustment factor ($\phi$), the adjusted effective reproductive number ($R^* = \phi R$), and the adjusted critical vaccination threshold ($V_c^*$). We show clustering of non-vaccination can have a significant impact on both $R$ and $V_c$ (Tables 2.1 and A.1). Under contact distribution $g_A(r)$ and high clustering, $R^*$ is nearly double that assumed by standard homogenous estimates and $\phi=1.71$, and $R^*>1$. 
with 95% vaccination coverage (Table 2.1). These effects translate into significant increases in $V_c^*$; under $g_A(r)$ and high clustering, the critical vaccination threshold increases from 93 to 96% (Table 1.1).

Table 2.1. Analytic $R^*$ and $V_c^*$ estimates from Equation 1 assuming $R_0=15$ and using contact distributions derived from Read et al. 2014 ($g_A(r)$), rural Zambia ($g_B(r)$), and Noulas et al. 2012 ($g_C(r)$).

<table>
<thead>
<tr>
<th>Vaccination Coverage</th>
<th>Clustering</th>
<th>$g_A(r)$</th>
<th>$g_B(r)$</th>
<th>$g_C(r)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%</td>
<td>None</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0.89</td>
<td>0.90</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>1.03</td>
<td>1.06</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>1.31</td>
<td>1.36</td>
<td>0.94</td>
</tr>
</tbody>
</table>

†Assumes $R_0=15$ and $\lambda=0.5$.

The absolute increase in $V_c^*$ due to clustering is inversely related to $R_0$ (Figure 2.2). For diseases with high $R_0$ like measles ($R_0=15$), the required $V_c^*$ with high clustering is 3% higher than without clustering ($V_c^*: 96%$ versus 93%; Table 2.1). Yet diseases with lower $R_0$, like pertussis ($R_0=10$) and polio ($R_0=6$), required increase in $V_c^*$ of 4% ($V_c^*: 90%$ versus 94%) and 7% ($V_c^*: 83%$ versus 90%; using $g_A(r)$).
Figure 2.2. The association between $R_0$, $V^*_c$, and spatial clustering of non-vaccination ($\phi$). As clustering increases, the required vaccination coverage to maintain $R=1$ increases. At low $R_0$, this increase due to clustering is much greater than at high $R_0$. (A) The association between $R_0$ and $V_c$ at the four defined levels of clustering, using the three $g(r)$ distributions. The relative increase in $V_c$ with increased clustering is relatively equivalent for $g_A(r)$ and $g_B(r)$, but lower for $g_C(r)$. (B) The relationship between $V^*_c$, $\phi$, and $R_0$, where $\phi$ represents the relative impact of spatial clustering on $R^*$. This relationship is highlighted for measles ($R_0=15$), pertussis ($R_0=10$), and polio ($R_0=6$). For diseases like polio, with lower $R_0$, there is a substantially greater increase in absolute $V^*_c$ with increasing $\phi$, as compared to diseases with higher $R_0$.

The shapes of both the contact distance and clustering distributions affect the resulting estimates of $R^*$, $V^*_c$, and $\phi$. Of the three contact distributions applied, $g_B(r)$ produced the highest $R^*$ values in the
presence of clustering due to having the highest proportion of contacts within very short distances, whereas the more dispersed $g_c(r)$ distribution produces the lowest $R^*$ estimates (Table 2.1). Curves of $1-G(r)$ (where $G(r)$ is the cumulative distribution function of $g(r)$) versus $\tau(r)$ demonstrate the effect of overlap between $\tau(r)$ and $g(r)$ (Figure 2.3). The areas under these curves are equal to $\phi^{-1}$ (Figure A.3).

We see a consistent maximization of the effect of clustering with maximal overlap of $g(r)$ probability and maximum $\tau(r)$ value (i.e. moving toward top right), occurring when 100% of contacts between individuals are at the distance where $\tau(r)$ is maximized, in this case at $r=0$. We also observe consistent minimization with minimal overlap, producing $V_c^*$ approaching the unadjusted $V_c=93\%$. Theoretically, if the shapes of the $\tau(r)$ and $g(r)$ functions were mismatched, we could see $V_c^*$ below $V_c$; however, for this study these were constrained to monotonically decreasing distributions, thus the lowest possible $V_c$ is produced when $g(r)$ is equal across all $r$ (i.e. equal contact distance probability, or homogeneous contact).

Figure 2.3. Critical vaccination threshold, as defined by the overlapping densities of the $\tau(r)$ and $g(r)$ distributions. Plotted curves are $\tau(r)$ where $r$ is that which corresponds to the value of 1 −
2.5.2 Simulation Studies

We found near perfect correlation between the expected outbreak probabilities from homogeneous mixing SIR simulations and spatial mixing simulations (Pearson’s $\rho=0.98$) and high comparability between these estimates (Pearson’s $\chi^2 p=0.24$) (Table 1.2). However, the final sizes of these epidemics differ significantly (51% [95% CI: 42-60%] vs 82% [95% CI: 80-84%] at $R^*=2.0$), due to well-known differences between spatial and homogenous mixing populations. The probability of outbreak ratio (PrR) comparing with clustering to without, which demonstrates the relative impact of susceptibility clustering, increases exponentially as $R$ decreases (Figure 2.4). For example, with measles comparing high clustering to none, at 95% immunity PrR>20 (Pr(outbreak)=23% [95% CI: 22-24%] vs. 0% [95% CI: 0-0.1%]), yet at 85% immunity, PrR=1.3 (95% CI: 1.2-1.4) (Tables 1.3 and A.3).

Table 2.2. Outbreak probability from simulations with SIR and individual-based spatially explicit model, assuming $\lambda=0.5$ and $g(r)$.

<table>
<thead>
<tr>
<th>Vaccination Coverage</th>
<th>Clustering</th>
<th>$R^*\dagger$</th>
<th>$G(r)$, and color corresponds to the resulting $V_c$ when $R_0=15$. Three levels of clustering ($\theta=0.25, 0.50, 1.0$) and 529 $g(r)$ distributions, with both $\alpha$ and $\beta$ range=0.1-10.</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%</td>
<td>None</td>
<td>0.75</td>
<td>Spatial Simulations, mean (95% CI)</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0.89</td>
<td>0 (0-0.01)</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>1.03</td>
<td>0.01 (0.00-0.02)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>1.24</td>
<td>0.07 (0.06-0.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.23 (0.22-0.24)</td>
</tr>
</tbody>
</table>

\* Outbreak is defined by $\geq 5\%$ of the susceptible individuals becoming infected. Assumes introduction happens at equal rates randomly among susceptibles.

\dagger $R^*$ calculated analytically using the empirical $g(r)$ from spatial simulations and defined $\tau(r)$ distributions.

\¥ Relative risk of outbreak occurring in spatial simulations, where “None” is the reference for each vaccination coverage level. Reference is no clustering. RR for SIR simulation results.
Chapter 2: The challenge of spatial clustering of susceptibility for measles elimination

2.5.3 Application

We illustrate our methods using DHS data from Tanzania, which has a recent history of largely uninterrupted measles virus transmission concurrent with increasing vaccination coverage. Visual examination of cluster and district-level vaccination coverage demonstrates apparent clustering of non-vaccination during both 1999 and 2010 Tanzania DHS surveys, with a possible increase in clustering in 2010 compared to 1999 (Figure 2.5a). We fit an exponential $\tau(r)$ to DHS data to characterize clustering of non-vaccination in Tanzania (Figure 2.5b), resulting in $\phi$ of 1.26 (95% CI: 1.1-1.5) and 1.56 (95% CI: 1.2-1.9) for 1999 and 2010 (using $g_d(r)$; Figure 2.5b). Assuming static clustering, we applied these $\phi$s to 5-year means of UNICEF/WHO vaccination coverage for Tanzania and estimated the reduction in reproductive number achieved from vaccination alone (compared to if clustering had not existed). In 1999, $R^*=$4.50 (95% CI: 3.90-5.30) versus $R=$3.57, and in 2010, $R^*=$1.83 (95% CI: 1.46-2.25) versus $R=1.17$ (Table 3). Starting in 2012, $R^*<1$ and vaccination $>V^*_c$, as implied by the current vaccination rates, and this was followed by measles incidence dropping and remaining below 1 per 100,000 (Table...
S2). Assuming clustering remains consistent, and $R_0=15$, Tanzanian will likely need to maintain successful vaccination coverage of at least 95-96% to maintain interrupted measles transmission, especially without measles virus circulation providing some naturally-acquired immunity.

Table 2.3. Incidence and estimated effective reproductive numbers for Tanzania from 1998 to 2014.

<table>
<thead>
<tr>
<th>Year</th>
<th>Routine Vaccination</th>
<th>5-yr Incidence (per 100,000)</th>
<th>$R_e$, 5-year mean</th>
<th>$V_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unclustered</td>
<td>Clustered†</td>
</tr>
<tr>
<td>1999</td>
<td>76.2%</td>
<td>29.78</td>
<td>3.57</td>
<td>4.5 (3.9-5.3)</td>
</tr>
<tr>
<td>2010</td>
<td>92.2%</td>
<td>3.75</td>
<td>1.17</td>
<td>1.8 (1.5-2.3)</td>
</tr>
</tbody>
</table>

† Estimate mean and 95% CI.
Figure 2.5. (A) Vaccination coverage among children aged 1-5yr by district and survey cluster in Tanzania during the 1999 and 2010 Demographics and Health Surveys. The DHS are cross-sectional data on individuals sampled using cluster randomized sampling. GPS location is only available for clusters, not individuals. (B) Empirical $\tau(r)$ functions and fitted exponential for Tanzania and Zambia, using DHS cluster data.
2.6 DISCUSSION

Elimination efforts for diseases like measles are based almost entirely on reaching and maintaining a vaccination target to produce a level of population immunity at which disease transmission cannot be sustained in a population, or herd immunity. Achieving and maintaining vaccination coverage targets, which at 95% are believed to be high enough to produce herd immunity, has been an immense challenge. Despite decades of efforts, global MCV1 coverage hovers around 85%, and only 41% (79/194) of countries met the 95% target in 2015.33,34 These efforts might become even more difficult as our work indicates that the measles vaccination target of 95% may be insufficient in some settings to achieve and maintain herd immunity in populations with high susceptibility clustering.

Clustering of susceptibility to diseases like measles can lead to dramatically increased risk of transmission and outbreaks, particularly when we rely on herd immunity to maintain protection with imperfect vaccination coverage. Accounting for this clustering becomes increasingly more critical as vaccination approaches herd immunity levels. Specifically, the relative impact of clustering increases exponentially with increasing vaccination coverage, a phenomenon previously demonstrated by Salathe and Bonhoeffer through network model simulations.19 This presents particular challenges for measles elimination as it can result in continued outbreaks and transmission where interruption is expected. However, with population-wide spatially-resolved data, such as DHS or serological surveys, we can readily estimate additional vaccination coverage needs or the impact of tackling pockets of susceptibility.

In an effort to provide a useful general rule of thumb for vaccination, we made several simplifying assumptions, specifically focusing on immunity acquired through successful vaccination. As such, we did not demonstrate a model that explicitly accounted for naturally-acquired immunity, vaccine efficacy, or multiple doses of vaccination. Vaccine efficacy, which is believed to be around 85% for a single dose (when given at nine months of age) and 97% for two doses of MCV, particularly complicates use of empirical data such as DHS where we only know whether any vaccine was received or not. Additionally,
low efficacy increases $R$ and $V_c$ substantially, leading to scenarios where herd immunity is not technically achievable with single dose regimens, particularly when clustering exists (Table A.3). However, vaccine efficacy is theoretically unbiased, assuming no differential cold chain or vaccine batch issues, and does not affect clustering, thus to account for it, we can simply assume $v^* = VE \times v$ for Equation 3.

Naturally-acquired immunity presents additional challenges to accurately estimating $R$, and if it occurs in a spatially localized manner, such as through an outbreak in a community, the nature of clustering of susceptibility may change. However, as noted above, the greatest impact of clustering, where this framework is most valuable, is when populations are close to disease elimination. In such populations, transmission of measles virus is limited and immunity is almost completely derived from successful vaccination. Use of serological surveys could circumvent this challenge if needed, as they have additional benefit of capturing clustering of naturally-acquired immunity.

One immediate challenge with our approach is accurate measurement of the spatial distribution of potentially infectious contacts. Spatial contact surveys have been used, but are expensive and context specific. As an alternative, we demonstrate the applicability of widely available data such as DHS data. Through a simplistic comparison of $R$ and $R^*$ estimates and incidence, we see evidence in Tanzania for the impact of clustering on the risk for continued outbreaks despite increasing vaccination coverage. However, this illustrative example uses an imperfect method for estimating $\tau(r)$ from Tanzania DHS data and a contact distribution from an entirely different context. Despite these limitations, we see both evidence of spatial clustering of non-vaccination and evidence that measles virus transmission is higher than would be expected using standard metrics. Other sources of available data may provide similar usefulness for measurement of clustering of potentially infectious contacts, including mobile phone, social network, and school-based data.
Describing human contact presents further challenges. While we used spatial proximity as basic metric with which to describe human contact, this is a simplistic approximation and does not fully capture the complexity of human social networks and mixing. Contact between individuals depends on where people work, where children go to school, and where people socialize. Human contact also differs significantly by age, both in distance of and number of contacts. This is further complicated by age assortativity, as it has been shown that individuals are much more likely to be in contact with others of a similar age, or with other specific age groups (e.g. children with parents or teachers). Although we do not capture age assortativity, age assortativity is essentially age clustering and produces similar impact to spatial clustering; we see this effect in the work previously mentioned that has demonstrated the impact of age-assortative social contact. This “age-specific clustering” could be incorporated into this framework in the future.

Although specific and exact population contact data, whether simple or complex, are generally not available, we found that we can likely apply distributions based on population characteristics with limited error. $V_c^*$ was especially robust to minor misspecification of $g(r)$ (Figures 2.3 and 2.4), particularly among populations with highly local contact, as demonstrated with $g_A(r)$ and $g_B(r)$. While derived from extremely different populations, these produced comparable $R^*$, $\phi$, and $V_c^*$ estimates at all levels of clustering, which differed greatly from those of $g_C(r)$, where contact is much closer to homogeneous. Through identifying a set of characteristic contact distributions that capture the range of possibilities, we can likely enable ready application of this approach without need for specific contact distributions.

We developed these methods using measles, but their application readily extends to other infectious diseases including pertussis, mumps, rubella, and polio, all of which currently present challenges in the U.S. and abroad. With lower $R_0$ values, our methods may more readily translate into feasible strategies to combat these diseases. For example, recent outbreaks of mumps among college students and hockey teams were likely contributed to by high social and spatial clustering of susceptibility among these...
populations. Assuming high mumps susceptibility clustering and $R_0=6$, an adjusted immunity level of 90% would be needed to achieve herd immunity, and it is likely already above 80%. Efforts to implement vaccination within college campuses are likely to be highly successful at producing herd immunity among the student populations.

As we struggle to achieve and maintain measles elimination goals across the globe, new challenges present themselves, among which is the impact of heterogeneity of susceptibility and contact, where even minimal clustering can produce substantial increases in outbreak risk. Through simple adjustments to the current rules of thumb guiding vaccination policy, we can estimate the impact of heterogeneity and adjust strategies accordingly, whether through increasing overall vaccination goals, or through targeted vaccination efforts to reach clusters of susceptibility.

2.7 Policy Implications

<table>
<thead>
<tr>
<th>Policy Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.</strong> The potential for disease transmission and outbreaks is greater than expected when vaccination coverage and susceptibility are clustered.</td>
</tr>
<tr>
<td><strong>2.</strong> Vaccination targets are likely too low to achieve herd immunity in many settings because of susceptibility clustering.</td>
</tr>
<tr>
<td><strong>3.</strong> Susceptibility clustering has the strongest impact as vaccination coverage approaches herd immunity levels.</td>
</tr>
<tr>
<td><strong>4.</strong> Two strategies are possible:</td>
</tr>
<tr>
<td>a. Increase vaccination targets for the entire population, or</td>
</tr>
<tr>
<td>b. Target clusters or causes of clustering</td>
</tr>
</tbody>
</table>
References


CHAPTER 3

The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

Shaun Truelove and Justin Lessler

3.1 ABSTRACT

Recent measles outbreaks in parts of the world with previously high control point to both increasing non-vaccination and local clustering of that non-vaccination. While we know that clustering of non-vaccination can increase the risk of outbreaks, the extent of this clustering and its impact are not known. Using the Demographics and Health Surveys, we developed methods to quantify spatial clustering of non-vaccination, and applied these to 59 surveys from 33 countries. We found that many countries show evidence of substantial local clustering of non-vaccination, which has the potential to increase the risk of outbreaks, and for some countries, vaccination targets greater than 95% may be required to achieve herd immunity.

3.2 INTRODUCTION

Measles vaccination has increased around the globe since its introduction in 1963, and with it, measles incidence has decreased dramatically.¹ Entire regions, like the Americas, are now declaring measles eliminated and measles eradication no longer seems impossible.² In many parts of the world, particularly high-income countries, the risk of childhood measles has been all but forgotten, so much that some among new generations of parents believe their children no longer need the vaccine. Unfortunately, recent outbreaks in many parts of the world tell us otherwise.
Previous work by Salathé and Bonhoeffer demonstrated that within high-income countries near elimination, opinion clustering, specifically vaccination refusal, resulted in outbreak frequency expected for a population with much lower vaccination coverage.\(^3\) We see direct evidence of this with recent outbreaks within large religious communities, such as occurred in Ohio, U.S.A. and the Netherlands.\(^4,5\) In both Ohio and the Netherlands, vaccination coverage at the state/national level (92% and >95%) would have suggested minimal outbreak risk.\(^4,5\)

Recent increases in measles incidence in Europe and the Americas have been blamed on faltering vaccination uptake; many of these countries, however continue to maintain vaccination coverage overall that would demonstrate minimal risk.\(^6-8\) Recent work looking at nonmedical exemptions to school immunization requirements has demonstrated localized clustering in parts of the United States.\(^9,10\) These increases in non-vaccination, are likely not occurring randomly, but rather clustering (i.e. opinion clustering). Though this clustering does not likely resemble that of the well-delineated religious populations of Ohio and the Netherlands, instead occurring at smaller scales within neighborhoods and communities, the result is potentially just as problematic, causing substantially greater risk of outbreaks than solely due to the decreasing vaccination coverage alone.

We previously demonstrated an approach for efficiently quantifying the impact of spatial clustering of unvaccinated individuals and applying this impact to produce adjusted estimates of the effective reproductive number \((R)\), the critical vaccination threshold \((V_c)\), and the clustering adjustment factor \((\phi)\) [Chapter 2]. Using these methods, we explore the impact of small-scale spatial clustering of non-vaccination on the risk of measles virus transmission and outbreaks. Through applying our methods to readily available and standardized data available through the Demographics and Health Surveys (DHS), we can quantify the impact of this phenomenon and provide another metric for measuring the performance of vaccination efforts of countries around the world.
3.3 METHODS

3.3.1 Measuring spatial clustering using cluster surveys

We used data from Demographics and Health Surveys (DHS) to quantify spatial susceptibility clustering for measles across multiple countries. DHS are cluster randomized surveys, conducted every 3-5 years in more than 90 countries, and many surveys include GPS coordinates for cluster locations. For this analysis, we used all available surveys with reliable data on location and receipt of measles virus-containing vaccination (MCV) between 2000-2015. Fifty-nine surveys from 33 countries met our inclusion criteria (Table B.1). We used the reported vaccination status at the time of the survey of the children included as a metric for vaccination coverage.

To assess clustering, we applied a modified version of the spatial-τ function adapted for spatially clustered data. We employed a series of techniques to provide a smoothed estimate for \( \tau(r) \), including moving window estimation, location jittering, and bootstrapping (see Appendix B). We removed the contribution of clustering within households and explored methods to identify and remove outlying survey clusters (see Appendix B). These methods together provide consistent empirical estimates of the \( \tau(r) \) function. Exponential decay and smoothing spline functions were fit to the empirical \( \tau(r) \) to define a smoothed, functional form that can be applied to the \( R^* \) model from Chapter 2.

3.3.2 Comparison of non-vaccination clustering levels

We compared levels of clustering of non-vaccination among individuals aged 12-60 months between countries and regions using the fitted \( \tau(r) \) functions. To provide a simple quantitative comparison of \( \tau(r) \) functions across countries and regions, we calculated mean probability ratios of there being another unvaccinated individual within 0.1, 1.0 km, 5.0 km, and 10.0 km from an unvaccinated individual versus any individual in the population. We also compared the clustering coefficient (ϕ) calculated using the \( g(r) \) derived from Read et al. for each survey, where \( \phi > 1 \) demonstrates spatial clustering of non-vaccination.
Chapter 3: The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

3.3.3 Evaluation of the association between clustering adjusted estimates and measles incidence

To assess whether non-vaccination clustering contributed to measles incidence, we compared the Spearman’s rank correlation coefficient between the unadjusted and clustering adjusted estimates of $R$ ($R$ vs. $R^*$) and between estimates of the probability that an outbreak occurs given an introduction. These probabilities were estimated analytically and through simulations from the calculated $R$ and $R^*$. We also used regression models to estimate the likelihood ratio of $R^*$ versus $R$ given incidence data and the likelihood ratio of adjusted versus unadjusted outbreak probabilities (Appendix B).

3.3.4 Calculation of R

In addition to the DHS data and $g(r)$ distributions, we used several sources of publicly available data to assess the impact of clustering on incidence. The unadjusted effective reproductive numbers were estimated using publicly available estimates of vaccination coverage from WHO/UNICEF and assumed $R_0=12-18$. Reported measles case counts were taken from public WHO data and used with population data from the United Nations to calculate the population incidence of measles per 100,000.

3.3.5 Sensitivity Analyses

Sensitivity analyses were performed using a multitude of variations to the methods and data. These included comparisons of $\tau(r)$s calculated including and excluding household clustering, including and excluding outlier clusters, weighting with DHS provided sampling weights, weighting by population data from WorldPop, and variations of included age ranges (see Supplement). We explored additional methods for fitting of the empirical $\tau(r)$ functions, including additional distributions (e.g. gamma, lognormal) and smoothing techniques (e.g. GAM). Additionally, we used the two additional contact distributions ($g(r)$) from Searle et al. and Noulas et al. Finally, we validated our fitted $\tau(r)$ distributions by comparing distributions estimated from the full sets of clusters to distributions using sampled subsets (Appendix B).
All statistical analyses and mapping were performed with R version 3.4.1.

### 3.4 RESULTS

Quantified $\tau(r)$ functions demonstrate significant clustering of measles non-vaccination for most DHS surveys (Figures 3.1, B.2). Across the most recent DHS surveys for 42 countries, clustering was found to be significant for each, and on average unvaccinated individuals were 1.57 (95% CI: 1.04-2.4) times more likely to have another unvaccinated individual within 100m as compared to the entire population. This varied significantly between countries (CoV$_{ln}$=0.48; range=1.2-3.7). Regional variation in clustering was not significant, though the European Region demonstrated substantially higher clustering than the other regions (Table 3.1).

![Figure 3.1](image-url)

**Figure 3.1.** Estimated $\tau(r)$ fits for five countries using DHS data. Forms of the fitted distributions follow an exponential decay function.
Table 3.1. Regional estimates of cumulative relative probability of non-vaccination at ≤100m, ≤1km, ≤5km, and ≤10km.

<table>
<thead>
<tr>
<th>WHO Region</th>
<th>Countries (no.)</th>
<th>≤100m (mean [95% CI])</th>
<th>≤1km (mean [95% CI])</th>
<th>≤5km (mean [95% CI])</th>
<th>≤10km (mean [95% CI])</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFR</td>
<td>21</td>
<td>1.53 (1.13-2.06)</td>
<td>1.19 (1.01-1.4)</td>
<td>1.04 (1-1.08)</td>
<td>1.02 (1-1.04)</td>
</tr>
<tr>
<td>EMR</td>
<td>4</td>
<td>1.43 (1.08-1.9)</td>
<td>1.14 (1.01-1.29)</td>
<td>1.03 (1-1.06)</td>
<td>1.02 (1-1.03)</td>
</tr>
<tr>
<td>EUR</td>
<td>2</td>
<td>2.54 (0.89-7.21)</td>
<td>1.57 (0.8-3.05)</td>
<td>1.13 (0.92-1.37)</td>
<td>1.06 (0.96-1.18)</td>
</tr>
<tr>
<td>SEAR</td>
<td>4</td>
<td>1.61 (1.2-2.16)</td>
<td>1.22 (1.02-1.47)</td>
<td>1.05 (1-1.1)</td>
<td>1.02 (1-1.05)</td>
</tr>
<tr>
<td>WPR</td>
<td>2</td>
<td>1.55 (1.44-1.66)</td>
<td>1.16 (1.09-1.24)</td>
<td>1.03 (1.02-1.05)</td>
<td>1.02 (1.01-1.02)</td>
</tr>
<tr>
<td>Overall</td>
<td>33</td>
<td>1.57 (1.04-2.37)</td>
<td>1.2 (0.95-1.52)</td>
<td>1.04 (0.98-1.11)</td>
<td>1.02 (0.99-1.06)</td>
</tr>
</tbody>
</table>

Adjustment for non-vaccination clustering significantly increased estimates of the effective reproductive number ($\phi$=1.49 [95% CI: 1.03-2.14]; mean $R^*$=3.46 vs. $R=2.33$, assuming $R_0=15$) and significantly increased corresponding critical vaccination thresholds (93% vs. 95% [94-97%]; Table 3.1). We estimated that only 3/33 (9%) countries require a $V_c^*$ that is statistically higher than the prescribed 95% target (Figure 3.2, Table B.2). However, for 13/33 (39%) countries, $V_c^*$ estimates were greater than 95% more than half the time and all 33 countries have some probability that $V_c^* > 95\%$ (Pr($V_c > 95\%$); median=.45 [range: 0.12-1.0]; Figure 3.3). Mapping demonstrates no obvious trend with geospatial location.

Figure 3.2. Estimated vaccination coverage needed above 95% to maintain herd immunity, DHS 2000-2015.
Chapter 3: The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

Figure 3.3. Probability that the needed critical vaccination threshold ($V_c^*$) is greater than the target of 95%, DHS 2000-2015.

Table 3.2. Regional estimates of the clustering coefficient ($\phi$), effective reproductive number ($R$ and $R^*$), the adjusted critical vaccination threshold ($V_c$), and probability of outbreak given introduction.

<table>
<thead>
<tr>
<th>WHO Region</th>
<th>Surveys (no.)</th>
<th>Countries (no.)</th>
<th>$\phi$ (mean [95% CI])</th>
<th>$R$ (mean [95% CI])</th>
<th>$R^*$ (mean [95% CI])</th>
<th>$V_c$ (mean [95% CI])</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFR</td>
<td>38</td>
<td>21</td>
<td>1.4 (1.1-2.0)</td>
<td>3.1 (1.1-9.2)</td>
<td>4.5 (1.5-13.2)</td>
<td>0.95 (0.94-0.97)</td>
</tr>
<tr>
<td>EMR</td>
<td>6</td>
<td>4</td>
<td>1.4 (1.1-1.9)</td>
<td>1.0 (0.2-5.5)</td>
<td>1.4 (0.2-8.6)</td>
<td>0.95 (0.94-0.97)</td>
</tr>
<tr>
<td>EUR</td>
<td>2</td>
<td>2</td>
<td>2.3 (0.9-5.9)</td>
<td>0.7 (0.3-1.4)</td>
<td>1.5 (1.3-1.9)</td>
<td>0.97 (0.94-1.0)</td>
</tr>
<tr>
<td>SEAR</td>
<td>8</td>
<td>4</td>
<td>1.5 (1.2-2.0)</td>
<td>3.0 (1.2-7.6)</td>
<td>4.5 (1.6-12.5)</td>
<td>0.96 (0.94-0.97)</td>
</tr>
<tr>
<td>WPR</td>
<td>5</td>
<td>2</td>
<td>1.4 (1.3-1.5)</td>
<td>2.5 (0.9-6.8)</td>
<td>3.6 (1.4-9.3)</td>
<td>0.95 (0.95-0.96)</td>
</tr>
<tr>
<td>Overall</td>
<td>59</td>
<td>33</td>
<td>1.5 (1.0-2.1)</td>
<td>2.6 (0.7-10.2)</td>
<td>3.8 (0.99-14.4)</td>
<td>0.95 (0.94-0.97)</td>
</tr>
</tbody>
</table>

3.4.1 Association between Non-Vaccination Clustering and Measles Incidence

Correlations of reported measles incidence per 100,000 population with $R$ and outbreak probability was weakly positive, with slight improvements when adjusting for clustering (Table 3.3). Correlation was slightly improved when using 5-year mean incidence compared with single year incidences, and correlation was higher for incidence with probability of outbreak than with $R$. When restricting to surveys
Chapter 3: The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

with 90% ($n=24$) and 95% ($n=12$) vaccination coverage, the correlation was much stronger, and adjusting for clustering produced a profound improvement in correlation with incidence.

Table 3.3. Correlation of measles incidence (per 100,000) with the effective reproductive number ($R$) and probability of outbreak. Assumes $R_0=15$.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Surveys</th>
<th>Correlation Coefficient</th>
<th>Incidence vs. $R$</th>
<th>Incidence vs. Pr(Outbreak)</th>
<th>Unadjusted</th>
<th>Adjusted</th>
<th>Unadjusted</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>All surveys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incidence, 1yr</td>
<td>59</td>
<td></td>
<td>0.27</td>
<td>0.29</td>
<td>0.29</td>
<td>0.29</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>Incidence, 5yr mean</td>
<td>59</td>
<td></td>
<td>0.34</td>
<td>0.37</td>
<td>0.36</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;90% Vaccination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incidence, 5yr mean</td>
<td>24</td>
<td></td>
<td>0.51</td>
<td>0.59</td>
<td>0.50</td>
<td>0.50</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>&gt;95% Vaccination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incidence, 5yr mean</td>
<td>12</td>
<td></td>
<td>0.08</td>
<td>0.65</td>
<td>0.13</td>
<td>0.13</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>Regions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFR</td>
<td>38</td>
<td></td>
<td>0.17</td>
<td>0.17</td>
<td>0.32</td>
<td>0.32</td>
<td>0.36</td>
<td>0.36</td>
</tr>
<tr>
<td>SEAR</td>
<td>8</td>
<td></td>
<td>0.34</td>
<td>0.52</td>
<td>0.22</td>
<td>0.22</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>WPR</td>
<td>5</td>
<td></td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>EMR</td>
<td>6</td>
<td></td>
<td>0.93</td>
<td>0.93</td>
<td>0.88</td>
<td>0.88</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>EUR</td>
<td>2</td>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Non-AFR</td>
<td>21</td>
<td></td>
<td>0.64</td>
<td>0.70</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Correlation within WHO regions demonstrated better correlation when adjusting for clustering for each region except the Africa Region (AFR), which demonstrate the worst correlation among the regions ($\rho=0.17$, Table 3.3). The Eastern Mediterranean Region (EMR) and Europe Region (EUR) had the highest correlations, though low numbers of countries included ($n=6$ and $n=2$). Among all non-AFR countries, correlation was moderate and improved when accounting for clustering.

Linear regression demonstrated significant association between both $R$ and $R^*$ and measles incidence variables, reported and estimated, with an increase in $\log(R)$ and $\log(R^*)$ of 1 associated with increases of 1.15 and 1.13 in 5-year mean $\log$(incidence). Comparison of regression model fits demonstrated improved fit when using the adjusted $R^*$ as compared with the unadjusted $R$, with the $R^*$ model 2.1 times as likely when using the reported 5-year mean population incidence (Table B.4). Regression of estimated incidence demonstrated slightly worse fit with the adjusted $R^*$ (Likelihood ratio=0.74). Regression of outbreak
probability demonstrated better fit with the adjusted $R^*$ for all incidence variables. Reported 5-year mean incidence improved the most, with the $R^*$ model being ten times more likely than $R$.

Overall, we observed that increased vaccination coverage is associated with increased clustering of non-vaccination, with $\phi$ increasing by 0.7% (95% CI: 0.3-1.0%) for each percent increase in vaccination coverage. However, within countries with multiple DHS available, no association was found between vaccination and $\phi$ ($p=0.8$).

3.4.2 Sensitivity Analyses

Various sensitivity analyses were performed to help determine the final form of the $\tau(r)$ function generating procedure. From these we determined that there were no outliers of substantial enough impact, particularly with smoothing (Appendix B). Thus, to simplify the process, outliers were not excluded. We also chose not to adjust for sampling weight using the provided weights from DHS. These weights are calculated at a regional or district basis and are not valid for use when applied to individuals.

Various sensitivity analyses were also performed to determine the sensitivity of the $\tau(r)$ functions to the stochastic processes of sampling clusters and households. Assuming sampling is based on population and that enough clusters were included, measurable clustering and resulting $\tau(r)$ functions were consistent.

We simulated prioritization of vaccination in urban or rural areas to determine if this was the cause of the initial dip around 2-10km in empirical $\tau(r)$ functions. These results confirmed that these dips were caused by disproportionately higher vaccination within urban compared to rural populations (Appendix B).

We also explored using smoothed surfaces generated using generalized additive models and as well as the modeled surface data provided by DHS for some recent surveys. While this produced $\tau(r)$ functions that
were smoother, they also washed out some of the initial peak of the function as a result of the smoothing process. Because of this, these data types did not provide as defined assessment of very local clustering.

### 3.5 Case Studies

Because most countries with DHS data have not been near elimination levels of measles control for long, if at all, our ability to detect strong trends between non-vaccination clustering and measles incidence was limited, and this was further complicated by real-world disease and population dynamics and the poor quality of reported incidence data. However, we did identify several countries for which clustering likely directly contributed to elevated measles virus transmission.

#### 3.5.1 Case study: Malawi

In 2010, Malawi had a massive measles outbreak with 134,000 reported cases, while having MCV1 coverage of 83-93% during the previous 5 years. While multiple factors contributed, including low vaccination rates during the decade prior resulting in infection in older children and teenagers, it is likely susceptible clustering contributed. We estimated that in 2010 Malawi had high non-vaccination clustering, with $\phi=1.5$ (95% CI: 1.42-1.66), resulting in an $R^*=1.6$ (95% CI: 1.5-2.1) compared with the unadjusted $R=1.0$ [95% CI: .84-1.3] for 2010. This translates to about nine times greater risk of outbreak per introduction of measles in 2010 ($pr(outbreak | introduction)=.43$ [.21-.55] versus .05 [0-.28]). This clustering, coupled with a buildup of susceptibles from minimal transmission during the previous decade (0 to 9 reported cases per 100,000 population), produced a population where minimal risk was perceived, yet high true potential, culminating in an outbreak of this magnitude.

#### 3.5.2 Case study: Countries at measles elimination levels

Measles incidence in Albania dropped from 21/100,000 to <1/100,000 in 2000 following substantial increases in vaccination coverage. However, despite having high overall vaccination coverage, Albania also has extremely high non-vaccination clustering among the Roma population ($\phi=3.4$ [95% CI: 2.28-...
5.35]), and in 2006, two outbreaks totaling 68 reported cases occurred, resulting in an incidence of 2.2/100,000. Following the outbreaks, vaccination coverage has increased to 98%, resulting in $R^*<1$ and no additional outbreaks since.

In Armenia and Morocco, clustering appears to have contributed to continued outbreaks and transmission despite high vaccination coverage (89-99%). While in both countries vaccination was close to the unadjusted critical vaccination threshold (93%) and $R \approx 1$ or $<1$, moderately high clustering ($\phi = 1.43$ and 1.42) likely contributed to incidence remaining $\geq 1/100,000$. Increased vaccination eventually caused $R^*$ to drop and maintain below 1 in both countries, and in Armenia, this coincided within one year with the drop to $<1/100,000$ measles incidence. For Morocco, this shift occurred 5 years after $R^*$ dropped.

Jordan demonstrates another likely example of the impact of non-vaccination clustering. From 2004-2012, measles incidence was $<1$ case per 100,000 (range: 0-0.7), maintained through high vaccination coverage of 95-99%. However, in 2013 amidst the Syrian crisis, an outbreak of 205 measles cases occurred, bringing the incidence to 2.9 per 100,000, the highest since 1998. Our clustering adjustment with DHS data did not detect this risk ($\phi = 1.2; R^* = 0.4$), as the most recent survey was in 2012 and likely does not include refugee populations, a significant limitation of DHS data. While these refugees made up only a small part of the population, work by UNICEF and UNHCR indicates high levels of non-vaccination among refugee children, and high clustering of this non-vaccination, thus likely a much higher $R^*$ than we estimated and a higher risk for outbreaks.

### 3.6 DISCUSSION

We quantified the expected impact of spatial clustering of measles non-vaccination, measured through use of Demographics and Health Surveys (DHS). We found that among 33 countries for which surveys were conducted from 2000-2015, all had significant spatial clustering of non-vaccination, with similar levels of clustering observed across WHO Regions (Table 3.2). As a result of this increased clustering, the required
vaccination coverage to achieve herd immunity was significantly higher than would be expected without clustering. However, only 3 countries (Albania, Tanzania, and Rwanda) had high enough clustering to dictate a statistically significant need to for greater vaccination than the prescribed 95% target (Table B.2).

The impact of this clustering appeared to be most relevant among countries and regions where measles vaccination is high, supporting theoretical findings from ours and other’s previous work. We see this with Albania, Morocco, and Armenia, all of which continued to experience measles virus transmission until the adjusted $R^*$ dropped below one. The clustering-conferred risk is likely to be exacerbated by the lack of naturally-acquired immunity: in countries with interrupted measles virus transmission, pockets of susceptibility persist that would otherwise be continuously depleted. This highlights the importance of countries being aware of non-vaccination spatial clustering as they near vaccination and control goals.

While we estimated that clustering should substantially increase measles virus transmission in multiple countries, we did not observe a pronounced overall association between non-vaccination clustering and measles incidence. A likely contributor to this disparity is that the occurrence of observed measles incidence is a complex process. We are assuming non-vaccination is equivalent to measles susceptibility, yet for many countries with endemic measles, the levels of susceptibility and susceptibility clustering may be much lower than expected due to naturally-acquired immunity. We also do not account for rates of introduction of measles virus into populations, which likely differ widely. For island nations like Madagascar and Comoros, the rates of introduction are likely different than those of landlocked countries like Malawi. Additionally, for cases to be observed they must seek care and be reported; it is generally held that measles incidence is much higher than reported.24,25

Also contributing to this disparity is the exponential relationship between clustering impact and vaccination coverage. As demonstrated in Chapter 2, the relative impact of susceptibility clustering is
Chapter 3: The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

very different between 80% vaccination coverage and 95% (Chapter 2, Figure 2.3). As the majority (65%) of countries for which DHS data are available have vaccination coverage ≤80%, the potential impact of clustering on this group of countries is likely limited. If we could conduct the same assessment for Europe, we suspect we would see strong positive correlation between clustering of measles susceptibility and incidence within countries.

Using survey data such as the DHS, while often the best data available, presents multiple challenges. Due to the selection of clusters to best capture the population, distance between clusters can be large, particularly in rural areas. This presents particular challenges for our model as we rely in information from pairs of individuals or clusters at each specified distance. We see with the empirical \( \tau(r) \) function that for many surveys there is a characteristic dip around the 2-10km range, which we can ascribe to a combination of higher densities of urban clusters sampled and higher vaccination within urban centers. Identification of this characteristic, while it presents challenges to our purpose of estimating a smooth, quantifiable \( \tau(r) \) function, may provide another important metric of vaccination practices. When this dip is prominent, it likely indicates focused vaccination activities within higher density areas like cities and towns, and potentially highlights disparities in vaccination access.

3.6.1 CONCLUSION

This analysis demonstrates quantification of spatial clustering of non-vaccination using available cluster survey data from DHS and, through application of our clustering adjustment model, quantitative assessment of the impact of this clustering. Through these methods, we can provide simple and familiar statistics like \( R^* \) and \( V_c^* \) with which policy makers and health authorities can readily use to assess the risks and successes associated with vaccination efforts in countries around the globe.

Broad scale heterogeneities in vaccination coverage (i.e. districts, regions, countries) have long been identified as clear contributors to the risk of measles outbreaks and continued transmission. Finer scale
local spatial heterogeneities, or clustering, are more challenging to define and thus get overlooked. As we demonstrate here, many countries show evidence of substantial local clustering of non-vaccination, which has the potential to increase the risk of outbreaks. While the results of this work are somewhat discouraging programmatically, particularly for those countries where we estimated a needed immunity above 95%, these results do not account for risk of introduction, neither into the country overall, or into specific locations within countries, nor do they account for naturally acquired immunity. The dynamics of measles and population behavior are complex and spatial clustering of susceptibility is only one component influencing disease transmission.
Chapter 3: The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

References


14. WHO Measles and Rubella Surveillance Data. *WHO* Available at:

15. United Nations Department of Economic and Social Affairs. Available at:


20. WHO | Immunization coverage. *WHO* Available at:


(Accessed: 16th May 2013)


23. *Jordan - Nortern Governorates - Percentage of Syrian households with no child vaccinated against measles by basic service unit (BSU).* (UNICEF, UNHCR, 2013).


PART III

RISK OF MEASLES IMPORTATION

INTO MEASLES-FREE POPULATIONS
CHAPTER 4

Biosecurity in a Globalized World: The Case of Measles

Shaun Truelove, Luis Mier-y-Teran-Romero, Allison Taylor Walker, Paul Gastanaduy, Andre Berro, Justin Lessler, Michael Johansson

4.1 ABSTRACT

The risk of domestic infectious disease outbreaks is often determined by incidence elsewhere in the world, the potential for those infections to cause cases in the U.S., and the probability those introduced cases spark outbreaks. The world is more connected than ever: 78 million international visitors arrived in the U.S. in 2015, and 44 million U.S. citizens traveled abroad. We developed a model for assessing the risk of measles importations to the U.S. from other countries through air travel, and then applied this model to assess the risk from Europe. From January 2006 through December 2015, 123 direct importations were reported from Europe, 37% of which originated from the 2011 outbreak centered in France. Despite equivalent incidence between the 2010 Bulgarian and 2011 French outbreaks, 2011 produced three times the estimated risk of measles in the United States, the result of travel volume differences. In the U.S., the highest risk of importation was among states with the largest populations and highest travel. Through real-time global infectious disease surveillance, we can readily assess domestic biosecurity threats and rapidly prepare and respond.
4.2 INTRODUCTION

The United States has 78 million tourists and 44 million residents enter its borders from abroad by air annually – equivalent to 334,000 travelers daily.\textsuperscript{1,2} With each of these individuals comes the potential for an infectious pathogen like measles virus to be brought into the U.S. Even with a miniscule probability per individual traveler, this high traveler volume could quickly result in hundreds or thousands of importations annually.

The looming biosecurity threat from air travel-based transmission of infectious diseases often reemerges during emerging outbreaks. From November 2002 through July 2003, SARS coronavirus spread to 37 countries, largely through air travel, and resulted in 8,098 cases and 774 deaths.\textsuperscript{3} Six years later the 2009 H1N1 influenza outbreak, originating in Mexico, caused widespread fear of a repeat of the 1918 influenza pandemic. While fortunately never reaching 1918’s devastation, the 2009-10 pandemic is estimated to have caused 284,000 deaths worldwide.\textsuperscript{4,5} The recent deadly outbreaks of MERS-CoV and Ebola in 2012 and 2014 spurned calls for flight restrictions and airport screening.\textsuperscript{6,7} While these and other emerging outbreaks highlight substantial biosecurity threats, the continuous threat of importation of diseases like measles often goes unnoticed.

The case of measles with U.S. biosecurity is a critical one. The U.S. has maintained measles elimination status since 2000, defined as the absence of continuous transmission for ≥12 months.\textsuperscript{8} However, in the time since, the U.S. has reported more than 100 measles outbreaks, over 2,000 reported cases, and 3 measles-related deaths, all the result of importations.\textsuperscript{9–11} Along with increasing globalization, less than ideal vaccination coverage in the U.S. and increased connectivity of populations within the U.S. have contributed to an increasing potential for these importations to become outbreaks. Recent outbreaks have already begun to exemplify this changing disease landscape. In 2014-15, the Disneyland amusement park was central to a large measles outbreak, and among the 110 California cases, 45% were unvaccinated and
43% had unknown or undocumented vaccination status.\textsuperscript{12} Another outbreak, occurring among the Amish community in Ohio in 2014, was imported from the Philippines by unvaccinated Amish missionaries returning home and infected at least 383 individuals.\textsuperscript{13}

Greater risk of importation also comes from greater incidence and more frequent and larger outbreaks elsewhere in the world, particularly where large numbers of traveler to the U.S. originate. Numerous large measles outbreaks have occurred in Europe during the last two decades, sparking multiple outbreaks in the US and Canada. In 2011, a large outbreak in France likely led to 15-17 reported importations and several resulting outbreaks in Quebec, Canada, and resulted in 16 reported direct importations in the U.S. that year.\textsuperscript{14–16} With the extensive travel between Europe and North America and outbreaks in Europe every year, the risk for measles to be brought into the U.S. or elsewhere in the Americas is high.

Using airline itinerary data for international arrivals into the U.S. and disease incidence data for countries abroad, we developed a model to estimate the risk of disease importation into the U.S., with measles serving as a case study. With this model, we hope to better understand the risks for measles virus importation through travel, whether by American residents returning from abroad, short-term visitors to U.S., or immigrants and refugees. We aim to provide better understanding of which U.S. states and cities are at greatest risk for importation, what times of the year have the highest importation risk, which countries are most likely involved, and what types of travelers contribute the highest risk. Through these findings, we can: (1) estimate future risk of importations, (2) determine whether events like the Disneyland outbreak can be predicted and prevented, (3) improve vaccination recommendations for travelers entering the U.S. and U.S. residents travelling abroad, and (4) assess probable origins of previous U.S. outbreaks.
Chapter 4: Biosecurity in a Globalized World: The Case of Measles

4.3 METHODS

4.3.1 Model

Our approach is based on previous models of importation for arboviruses.\textsuperscript{17,18} We assume that importation primarily occurs through air travel, and that it is directly proportional to the number of travelers and the number of cases of infection in the source population. We estimated the monthly number of infected travelers arriving at a destination ($d$) as a binomial process dependent on the number of infected individuals ($I_s$) in source country ($s$), the average duration of incubation and infection for measles ($D$), and the monthly probability that one of those individuals travels from $s$ to $d$ and is detected in $d$ ($p_{d,s,m}$):

$$n_{d,s,m} = Bin(p_{d,s,m}, I_s, mD)$$

Using the monthly number of travelers from $s$ to $d$ ($T_{d,s,m}$), the total population size of $s$ ($N_s$), we approximated $p_{d,s,m}$:

$$p_{d,s,m} = \frac{T_{d,s,m}}{N_s \delta_m}$$

To estimate the risk of an importation, we conducted 10,000 stochastic simulations, each with a sampled set of parameters. At the beginning of each simulation, parameter values were drawn from predetermined distributions, which were then applied to the functions above. In each simulation, these parameter sets were held constant across source countries, destinations, and months.

4.3.2 Data

Measles incidence data. Monthly European measles incidence data were acquired from the European Centre for Disease Prevention and Control (CDC) for 1999-2017.\textsuperscript{19} These data were checked against country incidence reports to verify accuracy.
Travel data. Travel data from the International Air Transport Association (IATA) were provided by the CDC. These data include locations of initial departure and final destination. As international airports like Chicago’s O’Hare or New York’s John F. Kennedy International Airports also serve nearby states, we have grouped several of these states together to limit bias.

Measles importations to the U.S. To validate the model, we used data on the reported number of cases of measles virus infection imported into the U.S., which were provided by the U.S. Centers for Disease Control and Prevention. These data included the month and state of detection, and country of origin or travel.

Disease Parameters. We assumed the infectious period for measles typically lasts from 4 days prior to through 4 days after presentation of the characteristic maculopapular rash, which typically presents 14 days after infection. Accordingly, we defined the mean time from infection to infectiousness as a normal distribution with a mean of 12 and a standard deviation of 2.

All data were restricted to January 2006 through December 2015 for this analysis, as this was the period for which data were complete for all sources. We also restricted the analysis to 28 European countries for which we were able to obtain complete data for this time period.

4.3.3 Model validation
Model performance was evaluated using Pearson’s correlation between the estimated reported risk and true reported importation cases, by year and by month. We further validated the model through comparison with probability models that excluded data (i.e. travel data, incidence data, both), through fitting to true reported importation data (see Supplement).
4.3.4 Relative probability of importation

Because we cannot estimate with confidence the absolute numbers of importations, we used the relative risk of importation or exportation as a metric with which to compare between importation destinations (i.e. U.S. states) and exportation sources (i.e. European countries). Through use of a ratio, we do not have to estimate the exact number of importations. We estimated the relative risk (RR) of importation as the ratio of the estimated number imported cases at a monthly and yearly basis, using a common number of importations across all time periods and sources or destinations as the reference, for each simulation.

Relative risk estimates were first calculated for each of the 10,000 simulations. For the risk of importation into the U.S., we used the median across all destinations (i.e. state) and all months or years as the reference value. Thus, for each unique set of parameter values, a single reference value is established. The mean RR and confidence intervals across the 10,000 simulations were estimated for each state/period.

For the risk of exportation to the U.S. from Europe, Austria was chosen as the reference group, thus, for each of the 10,000 simulations, RR of exportation was calculated as estimated risk (i.e. number of cases) divided by the risk with Austria. Because of the excessive number of zeros in the risk of exportation, as it is tied to occurrence of incident cases, we opted to select a reference country, and Austria was chosen as it is the median country among those with non-zero risk for more than one month.

To adjust for zeros in the importation estimates, which make risk ratios challenging, we used a pseudo-Bayesian approach where the probability of non-zero importation estimates was added to the importation estimates. For this we calculated the monthly probability for non-zero estimates for each destination or source across the 10,000 simulation results. These probabilities were then added back to the simulation results. This method was used rather than the traditional ‘add 1’ as it introduced less bias.

All analyses were performed using R 3.4.1 and RStudio 1.0.153.
4.4 RESULTS

From January 2006 through December 2015, the CDC reported 123 confirmed cases of direct importation of measles from 14 of 28 European countries. During those ten years, these countries reported 118,010 confirmed cases of measles, and individuals traveled 256 million times from Europe to U.S. destinations. European incidence was characterized by 9 epidemic periods, with two distinct large outbreaks of more than 30,000 reported cases during 2010 and 2011, primarily affecting Bulgaria and France, respectively (Figure 4.1). While the 2011 outbreak coincided with an appreciable spike in the number of reported importations of measles to the U.S. (n=46), 16 of which were from France, no notable increase in reported importations occurred during the 2010 outbreak (n=13; Table 4.1). No direct importations from Bulgaria were reported.

![Figure 4.1](image.png)

Figure 4.1. (A) Measles incidence in Europe and (B) reported measles case importations from twenty-eight European countries into the U.S. during January 2006 – December 2015.
Table 4.1. Importations and incidence from nine epidemic periods during 2006-2015.

<table>
<thead>
<tr>
<th>Epidemic Dates</th>
<th>Primary Outbreak Country</th>
<th>Incidence, Reported Importations</th>
<th>Reported Importations per 1000 cases</th>
<th>Estimated Risk of Importation*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan/2006 - Nov/2006</td>
<td>Germany</td>
<td>4,768</td>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>Jan/2008 - Sep/2008</td>
<td>Italy</td>
<td>8,438</td>
<td>10</td>
<td>1.2</td>
</tr>
<tr>
<td>Dec/2008 - Oct/2009</td>
<td>Multiple</td>
<td>5,264</td>
<td>11</td>
<td>2.1</td>
</tr>
<tr>
<td>Oct/2009 - Sep/2010</td>
<td>Bulgaria</td>
<td>31,202</td>
<td>13</td>
<td>0.4</td>
</tr>
<tr>
<td>Sep/2010 - Oct/2011</td>
<td>France</td>
<td>34,083</td>
<td>46</td>
<td>1.3</td>
</tr>
<tr>
<td>Oct/2011 - Sep/2012</td>
<td>Romania</td>
<td>12,043</td>
<td>4</td>
<td>0.3</td>
</tr>
<tr>
<td>Dec/2012 - Dec/2013</td>
<td>Multiple</td>
<td>2,216</td>
<td>18</td>
<td>1.6</td>
</tr>
<tr>
<td>Dec/2013 - Sep/2014</td>
<td>Italy</td>
<td>11,136</td>
<td>8</td>
<td>2.2</td>
</tr>
<tr>
<td>Sep/2014 - Sep/2015</td>
<td>Germany</td>
<td>3,667</td>
<td>7</td>
<td>1.6</td>
</tr>
</tbody>
</table>

*Relative risk is calculated relative to median risk across all outbreak periods

Using monthly European incidence data and international arrival data from Europe to the United States, we estimated the relative risk of measles importation into each destination (n=43; states/state groups, including D.C. and Puerto Rico) from the 28 European countries and the relative risk of exportation from each European country to the U.S. The overall estimated risk of importation during January 2006 – December 2015 correlated highly with state populations (Pearson’s $\rho=0.90$), and was highest for New York/New Jersey/Connecticut (NY/NJ/CT), California, and Florida (RR = 40 [95% CI=29-54], 20 [13-30], and 18 [11-26]; Figure 4.4a, Table 4.3). When adjusted for state population, Nevada and NY/NJ/CT had the highest risk of importation per capita, with 5-6 times the risk of importation per individual (RR = 5.6 [95% CI=1.9-13] and 5.2 [=3.3-9.0]; Figure 4.2b, Table 4.3). Risk of importation is dispersed geographically, though without major trends after accounting for population size (Figure 4.3).
Figure 4.2. Measles importation risk ratios by state or state group within the United States during January 2006–July 2016. (a) Overall risk ratio across time and (b) per capita risk ratio across time. Risk ratios are calculated as the cumulative risk across time for each state as compared to the median risk across all states.
Table 4.2. Overall estimated relative risks of measles importation by state, the top 20 states during the study period (January 2006 – December 2015), sorted by state mean population. Relative risk uses the median risk across all states as the reference. See Appendix C, Table C.2 for full table.

<table>
<thead>
<tr>
<th>Destination</th>
<th>Mean Population</th>
<th>Overall</th>
<th>Per Capita</th>
<th>Per Traveler</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR  95% CI</td>
<td>RR  95% CI</td>
<td>RR  95% CI</td>
<td>RR  95% CI</td>
</tr>
<tr>
<td>NY/NJ/CT</td>
<td>31,859,927</td>
<td>40 (29-54)</td>
<td>5.2 (3.3-9.0)</td>
<td>1.0 (0.74-1.4)</td>
</tr>
<tr>
<td>California</td>
<td>37,647,061</td>
<td>20 (13-30)</td>
<td>2.2 (1.3-4.0)</td>
<td>1.0 (0.66-1.5)</td>
</tr>
<tr>
<td>Florida</td>
<td>19,212,100</td>
<td>18 (11-26)</td>
<td>3.8 (2.1-7.0)</td>
<td>1.0 (0.62-1.5)</td>
</tr>
<tr>
<td>IL/WI/IN</td>
<td>25,006,142</td>
<td>9.3 (4.4-16)</td>
<td>1.5 (0.76-3.0)</td>
<td>1.1 (0.55-1.8)</td>
</tr>
<tr>
<td>DC/MD/VA</td>
<td>14,530,686</td>
<td>8.6 (4.4-15)</td>
<td>2.4 (1.1-4.8)</td>
<td>1.1 (0.57-1.9)</td>
</tr>
<tr>
<td>MA/ME/RI</td>
<td>8,994,548</td>
<td>8.4 (4.4-14)</td>
<td>3.8 (1.7-7.6)</td>
<td>1.1 (0.54-1.9)</td>
</tr>
<tr>
<td>Texas</td>
<td>25,634,112</td>
<td>7.4 (3.5-13)</td>
<td>1.2 (0.55-2.3)</td>
<td>1.1 (0.53-2.0)</td>
</tr>
<tr>
<td>PA/DE</td>
<td>13,610,750</td>
<td>4.6 (1.8-8.8)</td>
<td>1.4 (0.52-3.1)</td>
<td>1.2 (0.45-2.4)</td>
</tr>
<tr>
<td>Georgia</td>
<td>9,786,657</td>
<td>3.8 (1.7-7.9)</td>
<td>1.6 (0.56-3.7)</td>
<td>1.2 (0.47-2.5)</td>
</tr>
<tr>
<td>Nevada</td>
<td>2,734,504</td>
<td>3.7 (1.7-7.9)</td>
<td>5.6 (1.9-13)</td>
<td>1.1 (0.43-2.4)</td>
</tr>
<tr>
<td>North Carolina</td>
<td>9,609,849</td>
<td>2.9 (0.83-6.1)</td>
<td>1.3 (0.36-3.1)</td>
<td>1.2 (0.36-2.8)</td>
</tr>
<tr>
<td>Washington</td>
<td>6,817,587</td>
<td>2.9 (0.82-6.1)</td>
<td>1.7 (0.49-4.4)</td>
<td>1.2 (0.36-2.7)</td>
</tr>
<tr>
<td>Michigan</td>
<td>9,926,141</td>
<td>2.8 (0.82-6.1)</td>
<td>1.2 (0.33-2.8)</td>
<td>1.2 (0.36-2.8)</td>
</tr>
<tr>
<td>Colorado</td>
<td>5,122,700</td>
<td>2.5 (0.79-6.1)</td>
<td>2.0 (0.62-5.2)</td>
<td>1.2 (0.36-2.8)</td>
</tr>
<tr>
<td>Arizona</td>
<td>6,485,336</td>
<td>2.0 (0.73-5.1)</td>
<td>1.2 (0.4-3.4)</td>
<td>1.3 (0.47-3.3)</td>
</tr>
<tr>
<td>Minnesota</td>
<td>5,346,635</td>
<td>1.9 (0.72-5.1)</td>
<td>1.5 (0.49-4.1)</td>
<td>1.3 (0.46-3.3)</td>
</tr>
<tr>
<td>Ohio</td>
<td>11,549,682</td>
<td>1.5 (0.65-4.2)</td>
<td>0.6 (0.18-1.5)</td>
<td>1.3 (0.52-3.7)</td>
</tr>
<tr>
<td>Missouri</td>
<td>5,992,765</td>
<td>1.5 (0.64-4.2)</td>
<td>1.0 (0.36-3.0)</td>
<td>1.4 (0.54-3.8)</td>
</tr>
<tr>
<td>Tennessee</td>
<td>6,392,024</td>
<td>1.3 (0.6-4.1)</td>
<td>0.9 (0.3-2.5)</td>
<td>1.4 (0.58-4.0)</td>
</tr>
<tr>
<td>Kentucky</td>
<td>4,350,982</td>
<td>1.3 (0.57-3.2)</td>
<td>1.2 (0.44-3.6)</td>
<td>1.4 (0.59-3.8)</td>
</tr>
</tbody>
</table>
Chapter 4: Biosecurity in a Globalized World: The Case of Measles
Figure 4.3. Relative risk of measles importation by month among U.S. states and state groups, January 2006 – July 2016, sorted by population. Relative risk is calculated against the median risk of importation across all states and all months.

Monthly measles risks were consistent with overall risk, and highlight periods of elevated and reduced risk (Figure 4.3a). The months with the greatest aggregate importation risk were March – July 2011, during the large French measles outbreak, and when measles was spiking across Europe. During these months, Europe reported 16,577 measles cases (Figure 4.1), and the mean monthly risk of importation was six times normal (95% CI=3.8-7.6), and state-specific relative risks, as compared to the median monthly risk across all states, were among the highest estimated, with several states experiencing risks of several hundred times the overall median risk (median=11.3, range=0.15-1087).

The seasonal risk pattern of monthly risk was consistent with the seasonal patterns of measles incidence in Europe, with the highest risk consistently occurring during March through June and the lowest risk occurring during August through November (Figure 4.3). Both estimated monthly risk and reported importations correlate with incidence (Pearson’s ρ=0.8 and 0.5).

During January 2006 – December 2015, the top five exporting countries were Great Britain, France, Italy, Germany, and Ireland (Table 4.3). Exportation risk was highest for countries with high travel volume to the U.S. and either persistent measles incidence (e.g. Great Britain) or large outbreaks (e.g. France). Exportation risk varied by month, due to the variation in monthly traveler volume and measles incidence (Figure 4.4).
Table 4.3. Overall estimated relative risk of exporting measles by European countries, the top 20 countries during the study period (January 2006 – December 2015), sorted by country mean population. Relative risk uses the median risk across all countries as the reference. See Appendix C, Table C.3 for full table.

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Mean Population</th>
<th>Measles Incidence</th>
<th>Overall RR</th>
<th>95% CI</th>
<th>Per Capita RR</th>
<th>95% CI</th>
<th>Per Traveler RR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>63,225,965</td>
<td>9,925</td>
<td>87 (31-170)</td>
<td>8.9 (3.7-20)</td>
<td>3.7 (1.6-8.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>65,327,311</td>
<td>24,207</td>
<td>68 (24-140)</td>
<td>6.8 (2.7-15)</td>
<td>7.7 (3.3-18)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>59,550,475</td>
<td>14,554</td>
<td>47 (16-98)</td>
<td>5.1 (2-12)</td>
<td>5.5 (2.3-13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>81,555,942</td>
<td>10,131</td>
<td>30 (9.7-66)</td>
<td>2.4 (0.97-5.8)</td>
<td>2.7 (1.1-6.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>4,553,341</td>
<td>1,213</td>
<td>21 (6.4-49)</td>
<td>30 (10-78)</td>
<td>5.7 (2-16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>46,184,219</td>
<td>6,288</td>
<td>16 (4.7-40)</td>
<td>2.3 (0.89-6.1)</td>
<td>3.2 (1.1-8.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>16,672,271</td>
<td>3,044</td>
<td>15 (4.4-37)</td>
<td>5.9 (1.9-17)</td>
<td>4.2 (1.3-12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulgaria</td>
<td>7,357,095</td>
<td>24,427</td>
<td>13 (3.5-33)</td>
<td>11 (3.4-32)</td>
<td>52 (15-150)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romania</td>
<td>20,258,908</td>
<td>14,827</td>
<td>9.1 (2.4-26)</td>
<td>2.9 (0.95-9.2)</td>
<td>18 (4.7-58)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>10,978,721</td>
<td>1,748</td>
<td>7.2 (1.6-21)</td>
<td>4.2 (1.2-13)</td>
<td>4.3 (1.3-14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>8,437,769</td>
<td>1,308</td>
<td>4.6 (1.3-15)</td>
<td>3.5 (0.92-13)</td>
<td>4.3 (1-16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>5,572,569</td>
<td>196</td>
<td>1.6 (0.39-7.4)</td>
<td>1.8 (0.58-8.6)</td>
<td>0.96 (0.29-3.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td>10,995,472</td>
<td>415</td>
<td>1.2 (0.31-6.1)</td>
<td>0.69 (0.21-2.5)</td>
<td>0.84 (0.27-3.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>9,462,933</td>
<td>213</td>
<td>1.1 (0.28-4.2)</td>
<td>0.72 (0.22-2.7)</td>
<td>0.67 (0.21-2.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>38,063,160</td>
<td>787</td>
<td>1.0 (0.27-4.3)</td>
<td>0.18 (0.05-0.86)</td>
<td>0.73 (0.23-3.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Czech Republic</td>
<td>10,454,079</td>
<td>295</td>
<td>0.7 (0.19-3.6)</td>
<td>0.43 (0.13-1.9)</td>
<td>0.9 (0.33-5.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>4,951,461</td>
<td>98</td>
<td>0.6 (0.17-3.5)</td>
<td>0.79 (0.27-4)</td>
<td>0.57 (0.21-3.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iceland</td>
<td>320,477</td>
<td>1</td>
<td>0.22 (0.07-2.7)</td>
<td>4.5 (1.6-56)</td>
<td>0.49 (0.19-4.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lithuania</td>
<td>3,058,435</td>
<td>110</td>
<td>0.2 (0.07-2.7)</td>
<td>0.43 (0.16-3.5)</td>
<td>1.8 (0.71-32)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>5,386,179</td>
<td>51</td>
<td>0.11 (0.04-1.4)</td>
<td>0.14 (0.05-1.6)</td>
<td>0.24 (0.1-2.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4: Biosecurity in a Globalized World: The Case of Measles

Figure 4.4. Relative risk of measles exportation by month from European countries to the United States, January 2006 – July 2016, sorted by population. Relative risk is calculated against the median risk of exportation of Austria. Austria was used the reference country due to having median overall risk among countries with multiple months of non-zero exportation risk.

4.4.1 Model Validation

The annual estimated risk of measles importation correlated highly with reported importation cases (Figure 4.5; Pearson’s ρ=0.91). The validation models further demonstrated the validity of our approach, demonstrating that importation risk is most accurately estimated through combination of incidence and travel data, and the full model, given true reported importations, was significantly more likely than models the without travel or incidence data (likelihood ratio=10^97), with only incidence (LR=10^54), or with only travel (LR=10^22).
4.5 DISCUSSION

Despite measles elimination in the United States since 2000 and in the whole Americas region since 2016, the U.S. continues to experience numerous annual outbreaks, both large and small.8,23 Without circulating measles virus in the U.S. or neighboring countries, these outbreaks must be sparked by measles cases imported from elsewhere, likely via air travel. With 93% vaccination coverage in the U.S. for children aged 2-11, this conservatively translates to 20 million people in the U.S. susceptible to measles virus infection.24 If we assume the immunity level is closer to that found among U.S. Air Force recruits at 81%, closer to 60 million Americans may be at risk for measles, if they come in contact with an infected individual.25 Given the number of individuals potentially at risk for measles and the continual transfer of individuals between the U.S. and other parts of the world, it is critical to understand from where this risk originates and to where it in the U.S. it comes.

Here we assessed the potential for measles virus importation from Europe. While other continents are often perceived as greater threats, the frequency of travel between the U.S. and Europe produces unexpectedly high risk. For example, though measles incidence in sub-Saharan Africa is consistently 2-10
times higher than in Europe, travel volume between the U.S. and Europe is 40 times the volume between the U.S. and sub-Saharan Africa.\textsuperscript{26} Importation risk was typically highest among states with large populations, specifically New York, Florida, and California, which also serve as common ports of entry into the U.S.. While population correlated with importation risk, several states had disproportionately high rates once adjusting for population. Notably, these included states with high tourism, including Nevada (home of Las Vegas), New York, and Florida (Figure 3). The consistency of per traveler \(RR=1\) for all states indicates that none of the states have disproportionate travel with high incidence countries (Table 4.2).

Our model performed well for estimating the relative risk of measles importation and exportation from Europe to the U.S., with accurate risk estimation requiring both source country incidence and travel data. This is exemplified by the contrast between the 2010 and 2011 outbreaks, which were concentrated in Bulgaria and France, respectively. Despite similar reported incidences, the 2011 outbreak period was accompanied by three times the reported importations (46 versus 13) and three times the estimated risk (\(RR=4.1\) versus \(RR=1.5\)) of 2010 (Table 4.1). Travel volume explains these differences – France has 37 times the travel volume to the U.S. as Bulgaria. This resulted in 16 reported direct importations from France but none from Bulgaria.

U.S. residents constitute forty percent of international travelers between the U.S. and other countries and 65\% of measles importation cases, repeatedly causing outbreaks, including some of the largest.\textsuperscript{13,27–30} A recent study found that more than 16\% of adult U.S. travelers who attend pre-travel appointments are likely in need of measles vaccination at the time of international travel.\textsuperscript{31} The proportion needing vaccination is likely significantly higher among travelers not seeking pre-travel care. Reducing the risk contributed by U.S. travelers should be an essential first step toward mitigating this biosecurity threat, whether though limiting missed opportunities to vaccinate travelers seeking pre-travel consultation, or
through more prominent vaccination warnings for countries experiencing outbreaks or elevated measles virus transmission.

This work stresses the value of information and the importance of this information to U.S. biosecurity. Our world is highly connected, and infectious diseases are global problems that require global cooperation and efforts. Accurate estimation of the risk of any of these diseases landing in a country is impossible without accurate and complete data from the countries from which it might be coming. It is to the direct benefit of domestic populations to continue funding for measles, polio, and other surveillance efforts globally. The better these systems are for collection and sharing data, the more timely the data and the more likely these data can be used to prevent outbreaks and protect people. Here we focus on measles importations into the US, but the methods presented can be useful for other countries and other diseases. Virulent pathogens like MERS, Nipah virus, and highly resistant bacteria continue to emerge in different locations around the world, and with the right conditions they can take hold in the populations within which they emerge. When this occurs, we need accurate case information and effective tools, like our model, to understand how to control or prevent global dispersal and potential epidemics.

There were several limitations in this study. We cannot easily assess many of the factors that impact the exact number of disease cases that are imported. These factors include vaccination coverage, probability of detection, traveler characteristics, and others that contribute to the likelihood of an infected individual traveling to or being detected at their destination. As inclusion of these variables was not possible, we used relative risks of importation and exportation, allowing us to us to ignore these additional variables, assuming constancy between source countries and destination states.

Among these factors for which we lack specific data is reporting rates in both the source country and the U.S. We hypothesize that for an imported case to be reported, the person must either seek medical care, and then be reported by that provider, or result in secondary cases. For the first criteria, cases among foreign visitors might be less likely to seek care, particularly if they depart from the U.S. during the early
phase of the infectious period. Furthermore, we know reporting rates among U.S. health care providers are low, estimated as low as 3% of suspected cases.\textsuperscript{32} Cases causing secondary cases are more likely to be detected, but this requires contact with susceptible individuals.

We also do not account for traveler demographics, which can affect the risk of contracting measles or traveling when infectious. Age, for example, is associated with susceptibility status and contact patterns vary by age, with different age-assortativity and contact frequencies.\textsuperscript{33,34} While random variability in traveler demographics should have minimal impact, if consistent differences exist between sources or destinations or time periods, assumptions of similarity could be problematic.

Unfortunately, with broad models like this, it will likely always be impossible to predict specific outbreaks in populations with low probability of importation, such as was experienced in 2014 with the outbreak in Ohio Amish community.\textsuperscript{13} These high-risk populations are likely better served by vaccination education and outreach activities and targeting travelers pre-departure than attempts to understand their risk of importations and resulting outbreaks.

### 4.5.1 CONCLUSIONS

As long as measles virus continues to circulate in the world, the threat of importation persists. Vaccination remains the best defense against measles; however, because of its high transmissibility and perpetual struggles to vaccinate everyone, measles importations continue to result in outbreaks annually in the U.S. Locations in the U.S. with the greatest risk of importation are those with the greatest volume of travel with countries where measles, or any other diseases, is circulating. Often these are places with the largest populations (i.e. New York) or highest tourism (i.e. Las Vegas, Nevada). Additionally, not all source outbreaks are equivalent – with equivalent incidence. France contributed five times the risk of exporting measles cases as Bulgaria. With effective tools like this model and timely surveillance data, we can identify high risk destination locations and high-risk countries of origin, potentially enabling us to mitigate future biosecurity threats from infectious disease importations through preemptive actions.
Acknowledgements

Robert Whittaker and the European Centers for Disease Prevention and Control for providing the monthly measles incidence data for countries part of the European Union.
References


11. Data Access - Vital Statistics Online. Available at:

12. Measles Outbreak — California, December 2014—February 2015. Available at:


16. Measles — United States, 2011. Available at:
https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6115a1.htm. (Accessed: 26th June 2017)


Chapter 4: Biosecurity in a Globalized World: The Case of Measles


PART IV

CONCLUSIONS
CHAPTER 5

Conclusions

The success of smallpox eradication is often cited as proof that eradication of a human disease through vaccination is possible. For the work presented in this dissertation, smallpox eradication success also provides important evidence for our main findings, that to eliminate and eradicate measles, the focus needs to shift from national targets and efforts to local vaccination needs and risks of transmission. The ultimate success of the smallpox eradication efforts did not come from setting high overall vaccination goals and hoping those would produce enough herd immunity to halt the disease. Rather, eradication was eventually achieved through intensive, reactive localized ring vaccination, whereby cases of smallpox were identified and their contacts and the contacts of those contacts were vaccinated, thus preventing outbreaks and further spread from those cases. These efforts, combined with overall increasing vaccination coverage, enabled public health workers to cut off transmission of smallpox and eliminate remaining foci of disease, slowly eliminating both the existence of reservoirs of virus and potential for it to transmit.

The work in this dissertation demonstrates that a shift in strategies to one similar to that used in smallpox eradication is likely necessary to achieve measles elimination. For several decades now we have celebrated the continued successes of declining measles incidence and mortality resulting from a strategy of increasing widespread immunity through broad vaccination efforts. As highlighted in Chapter 1, measles vaccination coverage now stands at 85% globally, reported deaths due to measles have dropped below 100,000 globally for the first time ever in 2016, and one of the six WHO regions, the Americas, has established measles elimination. However, as we approach greater control of measles virus around
the globe, we also find that the current focus on meeting national vaccination targets fails to prevent outbreaks and incremental gains achieved by increasing vaccination have diminishing returns. To make the situation more challenging, these successes, particularly in developed countries, have erased the memory of the devastation that measles once caused from our collective minds, causing some to question the danger of measles and the need to vaccinate. As we demonstrate in this dissertation, a major shift is needed from the national/broad level of measles control efforts to the local level, both local immunity levels and local risks of transmission.

The importance of this shift is demonstrated theoretically in Chapter 2, with a new model we developed to account for heterogeneity in the population. This new model effectively quantifies the impacts of heterogeneous contact and distribution of susceptibility to measles virus infection, through directly increasing the effective reproductive number and critical immunity threshold. The effect of this “susceptibility clustering” becomes particularly critical when vaccination coverage approaches the estimates for the critical vaccination or immunity threshold ($V_c$; Figure 2.4). Chapter 3 demonstrates the existence of this clustering in countries across the globe, and we see some evidence of its impact on incidence during the past two decades. In short, current national vaccination targets are not enough to eliminate measles, particularly because vaccination distribution is highly heterogeneous and susceptible individuals are often clustered.

While theoretical calculations from our work indicate that continued increases in national vaccination targets could enable these countries to eventually interrupt measles virus transmission and establish elimination, a more practical and effective strategy may be to establish local, rather than national, vaccination targets of 95% coverage. These local targets would directly counter the existence of susceptibility clustering. With local targets and local shortcomings, tactics for achieving and maintaining local vaccination targets would become more appropriate to the particulars facing those areas, thus becoming more effective than national public health efforts.
In Chapter 4, we developed a model that effectively quantifies the risk of importations into the U.S. Despite having been eliminated in the U.S. for almost two decades, measles virus continues to cause outbreaks there. We demonstrate with this third paper that risk of measles virus importation is determined by both connectivity to sources, in the form of air travel, and measles incidence in those source locations. As long as incidence continues elsewhere in the world, the risk for measles virus entering the U.S. remains substantial because of the high volume of international travel. This is a challenge that has become more pronounced since the smallpox eradication campaign, with global travel increasing eightfold in the past four decades. This global connectivity allows rapid transmission across thousands of miles, from harbors of disease to seemingly protected populations. To limit the ability for importations to result in outbreaks, efforts are needed at the local level to reduce the existence of pockets of susceptibility. These pockets, which often won’t lead to continuous onward transmission, can result in outbreaks of several hundred individuals, as was seen in the Disneyland and Ohio Amish outbreaks. Furthermore, localized efforts are also needed to reduce the probability of an importation. These efforts should include travel clinics and general practitioners targeting their patients who lack evidence of measles immunity, particularly those traveling to high risk countries. Through targeting ports of entry most at risk and travelers who are the connective links between eliminated and endemic populations, maintenance of measles interruption and elimination can be more feasible.

As we globally approach higher measles vaccination coverage a shift from national to local vaccination targets and strategies is likely needed to achieve measles eradication. Susceptible individuals will always be present in populations – they are being born daily – however, if we can limit transmission of measles virus between these susceptibles by limiting the spatial and social clustering of these individuals through vaccination, we can likely prevent continued transmission of measles virus. The success experienced with smallpox provides hope that this can be accomplished, though, just as with smallpox, more intensive and active response to measles cases may be needed.
References


5. Schuttenhelm, R. Take a look at this graph: Global air travel increased 8 fold in 4 decades – and it’s an accelerating trend. Yes we have a problem. *Bits Of Science*


PART V

APPENDICES
APPENDIX A

Supplement to the challenge of spatial clustering of susceptibility for measles elimination

A.1 DERIVATION OF SPATIALLY-STRUCTURED EFFECTIVE REPRODUCTIVE NUMBER ESTIMATE

The standard equation for $R$ effective, assuming homogeneous mixing:

$$ R = R_0(1 - v) $$

where $v$ is the overall proportion vaccinated in the population. We assume that $v$ is equal to the proportion successfully vaccinated, and to account for a vaccine efficacy below 100%, $v$ is simply adjusted (i.e. proportion vaccinated x vaccine efficacy). To calculate an $R$ that accounts for the increased probability of contact between susceptible individuals due to spatial clustering, we modify the standard $R$ to include what is essentially a measure of the relative risk of contact between susceptible individuals due to clustering as compared with an unclustered, homogenous population. The derivation is as follows:

$R_0$ assumes that every contact is susceptible, or:

$$ \Pr(\text{susceptible} = 1) = \frac{N_{\text{suscept}}}{N_{\text{total}}} = 1 $$

For a population with some level of immunity, $\Pr(\text{susceptible contact}) \neq 1$ for the full population. Therefore, we calculate $R$ as:

$$ R = R_0 \times \Pr(\text{susceptible} = 1) $$

Under the assumption of homogeneous contact probability:
Appendix A: Supplement to the challenge of spatial clustering of susceptibility for measles elimination

\[ \Pr(\text{susceptible} = 1) = 1 - \Pr(\text{susceptible} = 0) = 1 - \frac{N_{\text{vaccinated}}}{N_{\text{total}}} = 1 - v \]

thus we arrive at our traditional calculation for \( R_{\text{eff}} \):

\[ R = R_0 (1 - v) \]

To incorporate heterogeneity, we can modify this equation by modifying the probability of a contact being susceptible. First, we can define for each distance \( x \) from another susceptible individual:

\[ \Pr(\text{susceptible} = 1 \cap X = x) = \Pr(X = x) \Pr(\text{susceptible} = 1 \mid X = x), \]

so for all distances (the whole population):

\[ \Pr(\text{susceptible} = 1) = \int \Pr(X = x) \Pr(\text{susceptible} = 1 \mid X = x) \, dx \]

To quantify spatial clustering, we use the \( \tau \) statistic:

\[ \tau(x) = \frac{1 - v(x)}{1 - v(\infty)} = \frac{\Pr(\text{susceptible} = 1 \mid X = x)}{\Pr(\text{susceptible} = 1 \mid X = \infty)} \]

Using \( g(x) \) to represent \( \Pr(X = x) \), we incorporate \( \tau(x) \) and \( g(x) \) and get:

\[ \Pr(\text{susceptible} = 1) = \int g(x) \Pr(\text{susceptible} = 1 \mid X = x) \, dx \]

\[ = \int g(x) \tau(x) \, dx \]

Since \( \Pr(\text{susceptible} = 1 \mid X = \infty) = 1 - v(\infty) = 1 - v \),

\[ \Pr(\text{susceptible} = 1) = (1 - v) \int g(x) \tau(x) \, dx \]
Finally, incorporating this we get our final model:

\[ R = R_0 * \Pr(\text{susceptible} = 1) \]

\[ R_{\text{eff}} = R_0 (1 - v) \int g(x) \tau(x) \, dx \]

### A.2 Weighted Cluster Survey \( \tau(r) \) Calculation

To calculate an accurate \( \tau(r) \) using clustered survey data with sampling weights for the clusters, we use a revised calculation to account for the structure of the data and sampling weights.

The standard equation for \( \tau(r) \) is:

\[ \tau(r_1, r_2) = \frac{\Pr(z_i = z_j \mid j \in \Omega_i(r_1, r_2))}{\Pr(z_i = z_j \mid j \in \Omega_i(.))} \]

See Salje et al. 2014 and Lessler et al. 2016 for more details.\(^{1,2}\)

We designed these methods for use with both clustered survey and discrete data. As such, to use with DHS data, we aggregated the individual-level DHS data to cluster-level data, with cluster ID, geospatial coordinates, number of individuals, sampling weights, and proportion susceptible for each cluster. These methods can be directly applied to discrete data such as counties or districts.

First, we define the weighted probability of a susceptible individual within a distance range \((r_1, r_2)\) from cluster \(i\):

\[ \pi_i(r_1, r_2) = \frac{\sum_{j \in \Omega_i(r_1, r_2)} s_j w_j p_j}{\sum_{j \in \Omega_i(r_1, r_2)} s_j w_j} \]
where:

\[ n_i = \text{number of clusters at a distance } (r_1, r_2) \text{ from cluster } i \]

\[ s_j = \text{number of individuals in cluster } j \]

\[ w_j = \text{cluster sampling weight of cluster } j \]

\[ p_j = \text{proportion susceptible in cluster } j \]

From this we can calculate the weighted probability of a susceptible within a distance range \((r_1, r_2)\), from all clusters:

\[
\Pi(r_1, r_2) = \frac{1}{\sum_{i=1}^{N} s_i w_i p_i} \sum_{i=1}^{N} \left\{ \frac{\sum_{j \in \Omega_i(d_1, d_2)} s_j w_j p_j}{\sum_{j \in \Omega_i(d_1, d_2)} s_j w_j} \right\}
\]

where:

\[ N = \text{number of clusters in the survey data} \]

\[ s_i = \text{number of individuals in cluster } i \]

\[ w_i = \text{cluster sampling weight of cluster } i \]

\[ p_i = \text{proportion susceptible in cluster } i \]

Finally, applying this to the original format of \(\tau(r)\):

\[
\tau(r_1, r_2) = \frac{\Pr(z_i = z_j \mid j \in \Omega_i(r_1, r_2))}{\Pr(z_i = z_j \mid j \in \Omega_i(.))}
\]

\[
\tau(r_1, r_2) = \frac{\Pi(r_1, r_2)}{\Pi(.)}
\]
Appendix A: Supplement to the challenge of spatial clustering of susceptibility for measles elimination

A.3 $\tau(r)$ DISTRIBUTIONS

We used an exponential function consistently characterize $\tau(r)$:

$$\tau(r) = \theta e^{-\lambda r} + \psi$$

Where $\theta$ defines the maximum probability ratio of being in contact with a susceptible at a distance of 0 ($r = 0$). $\lambda$ defines the rate of decay of the probability ratio with distance $r$. $\psi$ is a rescaling value, allowing the distribution to average to 1 across the entire range of $r$. Empirical $\tau(r)$ estimates were fit to this exponential function using non-linear least squares with inverse variance weighting.

A.4 $g(r)$ DISTRIBUTIONS

We included three contact distance probability distributions ($g(r)$) in this analysis to demonstrate the use of the approach within a range of settings and to examine the sensitivity of the model to variable distributions. For $g_d(r)$ we used a published inverse cumulative probability distribution from data collected through intensive interviews and contact diaries from Guangzhou, China. This population ranges from highly urban and dense to rural. Using GPS logger data from rural villages in Zambia, we constructed $g_d(r)$ using the assumption that time spent at distance $r$ from home approximated probability of contact at distance $r$. This distribution represents a rural population from small villages. The final distribution, $g_c(r)$, was constructed with cell phone data from the U.S. using Foursquare check-ins. Though these data do not necessarily capture all contacts and contact distances for this populations, this distribution provides us with a highly dispersed $g(r)$ that may represent the personal vehicle owning and commuting populations of the U.S. Gamma distributions were fit to each distribution through mathematical optimization using the Nelder-Mead method with the stats R package.
Appendix A: Supplement to the challenge of spatial clustering of susceptibility for measles elimination

Figure A.1. Contact distance probability distribution \( g(r) \) derived from contact diary and interview acquired contact data from China \( (g_A(r)) \), GPS logger data from Zambia \( (g_B(r)) \), and cell phone data from the U.S. \( (g_C(r)) \). From these data, gamma distributions were fit to produce parameterized approximations of each distribution.

A.5 Closed-form solution for \( \phi \)

Using the exponential decay and gamma distribution forms of \( \tau(r) \) and \( g(r) \), we have derived a closed-form solution to the integral \( \phi \). This derivation is as follows:

If \( \tau(r) = \theta e^{-\lambda r} + b \) and \( g(r) = f(r, \alpha, \beta) \sim \text{gamma}(\alpha, \beta) \), then:

\[
\phi = \int_{0}^{\infty} \tau(r)g(r)dr
\]

\[
= \int_{0}^{\infty} \left( \theta e^{-\lambda r} + b \right) f(r, \alpha, \beta)dr
\]

\[
= b \int_{0}^{\infty} f(r, \alpha, \beta)dr + \theta \int_{0}^{\infty} e^{-\lambda r} \frac{\beta^\alpha}{\Gamma(\alpha)} r^{\alpha-1} e^{-\beta r} dr
\]
Appendix A: Supplement to the challenge of spatial clustering of susceptibility for measles elimination

\[ b + \theta \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} r^{\alpha-1} e^{-(\beta+\lambda)r} dr \]

\[ = b + \theta \frac{\beta^\alpha}{(\beta + \lambda)^\alpha} \int_0^\infty \frac{(\beta + \lambda)^\alpha}{\Gamma(\alpha)} r^{\alpha-1} e^{-(\beta+\lambda)r} dr \]

\[ = b + \theta \frac{\beta^\alpha}{(\beta + \lambda)^\alpha} \int_0^\infty f(r, \alpha, \beta + \lambda) dr \]

\[ = b + \frac{\beta^\alpha}{(\beta + \lambda)^\alpha} \theta \]

A.6 SYNTHETIC POPULATIONS

We validated these methods through simulation studies using spatially-explicit, synthetic populations with vaccination coverage of 85, 90, and 95% and with four defined levels of clustering of non-vaccination (none, low, medium, and high). For each vaccination level, 5 initial base populations with no susceptible clustering were generated. Each population consisted of 100,000 individuals randomly distributed within a 30x30km space. Additional sizes and densities were examined in the sensitivity analyses. Individuals were randomly selected to be non-vaccinated according to the defined level of vaccination coverage. For each level of clustering, we took base populations and stochastically clustered the vaccination status of individuals to create populations with low, medium, and high clustering. This heterogeneous mixing, or clustering, was performed through an iterative process of randomly swapping the vaccinated/unvaccinated statuses of randomly selected individuals until the empirical \( \tau(r) \) matched the pre-defined parametric distributions (Figure A.2).
## A.7 SENSITIVITY ANALYSES

### A.7.1 Population Variability

Sensitivity analyses were conducted using additional synthetic populations of different sizes (N=10,000, N=50,000) and different geographic sizes (10x10km, 50x50km, etc.). These were performed to determine applicability of the methods to real populations. Our methods were found to be largely robust to both size and density. However, our methods failed in simulation to produce the expected proportions of outbreaks at lower densities, though this was expected. This resulted in localized outbreak extinctions due to rapid depletion of susceptibles in low density populations.

### A.7.2 Distributions

Analyses examining the impact of the $\tau(d)$ shape demonstrated that the $R$ estimate is relatively robust to spread ($\lambda_d$; Figure A.3). The rate parameter of the $\tau(r)$ distribution, $\lambda_r$, also impacts the estimates produced by our approach. While $\theta$, which is determined by the clustering level, $\lambda_r$, is determined by spread of clustering, and can represent the household-centricity of non-vaccination clustering. As $\lambda_r$

---

**Figure A.2.** Example of pre-clustering $\tau(r)$ in a synthetic population and the resulting $\tau(r)$ after the vaccination clustering procedure. Red line is the pre-defined parametric $\tau(r)$ and the black line is the fitted form.
Appendix A: Supplement to the challenge of spatial clustering of susceptibility for measles elimination

increases, the $\tau(r)$ distribution becomes narrower, and the expected $R^*$ produced asymptotically approaches that of the unclustered estimate.

![Figure A.3. Effect of increasing spread of the $\tau(r)$ when $\theta$ is held constant. (A) Parametric $\tau(r)$ distributions with $\theta=1.5$. (B) Associated resulting $R^*$ as $\lambda$ increases. Assumes a constant stable $g(r)$.]

Analysis of the impact of $g(r)$ shape indicated that the broader the contact distribution, the lower the resulting $R^*$ (Figure A.4). As $g(r)$ follows a true probability distribution function, as the dispersion of the distribution increases, the peak of the distribution decreases; this is contrary to $\tau(r)$ which consist of related but independent ratios at each value of $r$. Through analyses with both the gamma and exponential distributions of $g(r)$, we see that as contact localness increases ($g(r)$ becomes steeper), departing further from homogeneous contact, $R^*$ increases, asymptotically approaching a clustering-dependent maximum (Figure A.4).
Figure A.4. $g(r)$ decay rate with exponential distribution. (A) Exponential PDF forms of $g(r)$ with rate parameters 0.1, 0.2, 0.5, 1.0. (B) Exponential rate versus $R$ for varying levels of spatial clustering.

This clustering-dependent maximum is beneficial as it limits the impact of misspecification of the exponential rate ($\lambda$) with increasing localness. For example, misspecification of $\lambda=0.8$ instead of 1.0 results in an underestimate of $R$ by 1.1\% ($R=1.73$ vs. 1.75) at low clustering and by 3.1 \% ($R=2.42$ vs. 2.50) at high clustering and 90\% vaccination, translating to absolute $V_c$ underestimates of 0.06\% (94.22 vs. 94.28\%) and 0.13\% (95.87 vs. 96.00\%).

A.7.3 Household-based clustering

The flexibility of our approach and the parametric distributions also allow us to experiment with real-world scenarios. For example, to verify that the model does not overestimate the risk among populations with largely household-based clustering, we can adjust the $\tau(r)$ rate parameter, $\lambda$. In such populations, we would expect a reduced risk of outbreaks compared with populations with equivalent $\theta$ values (i.e., the maximum value of $\tau(r)$), but clustering primarily outside of the household. Increasing $\lambda$, while holding $\theta$ constant, corresponds to increasing proportion of clustering contained within households, with the $\tau(r)$
distribution becomes narrower, and results in $\phi$ approaching 1, thus decreasing effect of clustering, and $R^*$ approaching $R$ (Figure A.5).

![Figure A.5. Impact of $\tau(r)$ rate parameter, $\lambda_\tau$, on the clustering coefficient, $\phi$. As $\lambda_\tau$ increases and the $\tau(r)$ function becomes steeper, resembling clustering centered within households, $\phi$ approaches 1 and the $R^*$ estimate will approach the unclustered $R$ estimate.](image)

**A.8 EMPIRICAL $\tau(r)$ CALCULATIONS WITH DHS CLUSTER DATA**

**A.8.1 Moving Window $\tau(r)$ Estimation**

To approximate a continuous $\tau(r)$ function from discrete cluster locations, we used a moving window approach, rather than a static set of discrete distance windows. For example, instead of using 2km discrete windows, such as 0-2km, 2-4km, 4-6km, etc. We used overlapping windows, such as 0-2km, 1-3km, 2-4km, 3-5km, 4-6km, etc. This provided a much smoother empirical estimate, particularly for distances with reduced numbers of cluster pairs.

**A.8.2 Jittered Bootstrap Estimates**

To protect the identities of surveys subjects, DHS randomly jitters the location of each cluster, with urban clusters jittered up to 2km in any direction and rural clusters jittered up to 5km. To account for this potential variability in the true locations of the clusters, we incorporated a random jitter of cluster
locations into a bootstrap framework of 10,000 iterations. This improved the smoothness of the curve, particularly for distances with low numbers of cluster pairs.

**A.9 TECHNICAL CHALLENGES AND LIMITATION**

Our results demonstrate some of the technical challenges of this analysis. Dissimilarity between spatial and SIR simulation final sizes, the result of local depletion of susceptibles, makes using final size estimates to compare between models invalid (Figure A.6). In this case, while probability of contact and ultimately infection of another within the SIR model is wholly dependent upon the number of susceptibles remaining, the spatial simulations employed a situation where probability of contact is dependent upon distance of susceptibles, thus when all susceptibles nearby have already been infected, the probability of a contact drops well below what it would be if numbers alone mattered.

![Figure A.6](image)

Figure A.6. Empirical CDFs of spatial SIR simulation final sizes compared to the mean ECDF of homogeneous SIR simulations (A), and homogeneous final size ECDFs vs mean ECDF of spatial simulations. Spatial SIRs result in a mean final size around half that of homogeneous SIRs, but a similar, though slightly lower probability of outbreak.
Due to network properties, however, as $R$ increases from 1, the expected probability of an outbreak diverges in the spatial simulations from that of the SIR simulations. This divergence can be explained by the distribution of the initial individual $R$, or $R_i$, in the spatial populations. While $E[R_i]$ is equal to the analytic value from Eq. 5, the distribution of $R_i$ is substantially right-skewed, with a large portion of the population being expected to infect several times more individuals than that of $E[R_i]$ (Figure A.7). Furthermore, resulting from the spatial constraints of the transmission model, $R_i$ of subsequently infected individuals is dependent on $R_i$ of the infector, thus, the higher the $R_i$ of the infector, the higher the $R_i$ of the infected. This results in reduced probability of stochastic extinction and higher probability that outbreaks will propagate.

Figure A.7. Individual initial $R$ for a synthetic population with medium clustering and 95% vaccination coverage. (A) Clustering of unvaccinated individuals in the population space (vaccinated individuals are not shown; individuals are distributed evenly in the space). Unvaccinated individuals are colored by initial $R$. (B) Distribution of initial $R$ estimates for individuals in the population.
A.10 Validation Challenges

To validate the results of our approach, we conducted individual-based spatially explicit simulations of transmission. Because the of contact distribution, we needed a population with both a high density and large area.

A high density population was required because of the high probability of contacts within a short distance. As individuals are infected, they are removed from the potential pool of susceptibles. If the probability of a short contact distance is high, then the pool of susceptibles within a short distance of new and current infectious individuals becomes depleted, reducing probability of additional infections. Furthermore, if susceptible individuals are clustered, as we are proposing, this further complicates the depletion of susceptibles, as the distance to the next cluster might be relatively high, thus the probability of a transmission to that cluster is low. In this case, as soon as the cluster is depleted, the outbreak dies out, potentially below the threshold of an outbreak. However, by increasing the density of the population overall, the numbers of individuals, and thus chance for longer distance transmission increases.

Additionally, as the contact distance distribution, $g(r)$, continues to infer some probability of transmission at higher distance, we need to make the population area large enough in order to not artificially truncate $g(r)$ and bias the distance of contacts toward shorter distances.

Complicating these validation procedures, the higher the density and area of the population, the larger the population size, and the larger the number of susceptibles. For our validations, we are introducing infection by randomly selecting a single susceptible individual to be infectious, and simulated the transmission that results. Furthermore, once selected, the process of infectious transmission is stochastic, thus 100 simulations with the same initial infectious individual could all produce different results. When we have base populations of 500,000 with 90% vaccination, thus 50,000 susceptible individuals, running enough simulations to simply capture an iteration where each susceptible is the initial infectious
individuals is computationally intensive. Doing this for multiple populations for each of the 12 vaccination coverage/clustering level combinations requires high computational capacity.
A.11 Additional Supplemental Tables and Figures

Table A.1. Non-vaccination clustering adjustment factor estimates for the three contact distance distributions and four levels of clustering of non-vaccination.

<table>
<thead>
<tr>
<th>Clustering Level</th>
<th>Clustering Adjustment Factor ($\phi$)</th>
<th>Read et al. 2014, $g_A(r)$</th>
<th>Rural Zambia, $g_B(r)$</th>
<th>Noulas et al. 2012, $g_C(r)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1.18</td>
<td>1.20</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>1.36</td>
<td>1.40</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>1.71</td>
<td>1.80</td>
<td>1.24</td>
<td></td>
</tr>
</tbody>
</table>

Table A.2. Non-vaccination clustering-adjusted effective reproductive number estimates for 85, 90, and 95% effective vaccination coverage at the three contact distance distributions and four levels of non-vaccination clustering. Assumes $R_0=15$.

<table>
<thead>
<tr>
<th>Effective Vaccination Coverage</th>
<th>Clustering</th>
<th>$R^*$† $g_A(r)$</th>
<th>$g_B(r)$</th>
<th>$g_C(r)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%</td>
<td>None</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>0.89</td>
<td>0.90</td>
<td>0.80</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td>1.03</td>
<td>1.06</td>
<td>0.85</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>1.31</td>
<td>1.36</td>
<td>0.94</td>
</tr>
<tr>
<td>90%</td>
<td>None</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>1.78</td>
<td>1.81</td>
<td>1.60</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td>2.06</td>
<td>2.11</td>
<td>1.69</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>2.61</td>
<td>2.72</td>
<td>1.89</td>
</tr>
<tr>
<td>85%</td>
<td>None</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>2.67</td>
<td>2.71</td>
<td>2.39</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td>3.08</td>
<td>3.17</td>
<td>2.54</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>3.92</td>
<td>4.08</td>
<td>2.83</td>
</tr>
</tbody>
</table>

† Assumes $R_0=15$ and $\lambda=0.5$
Appendix A: Supplement to the challenge of spatial clustering of susceptibility for measles elimination

Table A.3. Non-vaccination clustering-adjusted critical vaccination thresholds at none, low, medium, and high clustering with vaccine efficacy (VE) of 100% and 95%, with \( \lambda = 0.5 \).

<table>
<thead>
<tr>
<th>Vaccine Efficacy</th>
<th>Clustering Level</th>
<th>( V_c )</th>
<th>Read et al. 2014, ( g_a(r) )</th>
<th>Rural Zambia, ( g_a(r) )</th>
<th>Noulas et al. 2012, ( g_a(r) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>None</td>
<td>93.3%</td>
<td>93.3%</td>
<td>93.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>94.3%</td>
<td>94.4%</td>
<td>93.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>95.1%</td>
<td>95.2%</td>
<td>94.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>96.1%</td>
<td>96.3%</td>
<td>94.6%</td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>None</td>
<td>98.2%</td>
<td>98.2%</td>
<td>98.2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>99.3%</td>
<td>99.4%</td>
<td>98.6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>&gt;100%</td>
<td>&gt;100%</td>
<td>99.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>&gt;100%</td>
<td>&gt;100%</td>
<td>99.6%</td>
<td></td>
</tr>
</tbody>
</table>

Table A.4. Non-vaccination clustering-adjusted critical vaccination thresholds at none, low, medium, and high clustering with vaccine efficacy (VE) of 100% and 95%, with \( \lambda = 0.5 \).

<table>
<thead>
<tr>
<th>Vaccination Coverage</th>
<th>Clustering Level</th>
<th>( R^* )</th>
<th>Probability of Outbreak†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spatial Simulations, mean (95% CI)</td>
</tr>
<tr>
<td>95%</td>
<td>None</td>
<td>0.75</td>
<td>0 (0-0.01)</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0.89</td>
<td>0.01 (0.00-0.02)</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>1.03</td>
<td>0.07 (0.06-0.08)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>1.24</td>
<td>0.23 (0.22-0.24)</td>
</tr>
<tr>
<td>90%</td>
<td>None</td>
<td>1.45</td>
<td>0.32 (0.28-0.36)</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>1.78</td>
<td>0.43 (0.38-0.47)</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>2.01</td>
<td>0.47 (0.44-0.50)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>2.31</td>
<td>0.52 (0.48-0.57)</td>
</tr>
<tr>
<td>85%</td>
<td>None</td>
<td>2.24</td>
<td>0.58 (0.51-0.66)</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>2.61</td>
<td>0.64 (0.52-0.76)</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>3.08</td>
<td>0.70 (0.58-0.82)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>3.92</td>
<td>0.75 (0.61-0.89)</td>
</tr>
</tbody>
</table>

† Outbreak is defined by \( \geq 5\% \) of the susceptible individuals becoming infected. Assumes introduction happens at equal rates randomly among susceptibles.
‡ \( R^* \) calculated analytically using the empirical \( g(r) \) from spatial simulations and defined \( \tau(r) \) distributions.
¥ Relative risk of outbreak occurring in spatial simulations, where “None” is the reference for each vaccination coverage level. Reference is no clustering. RR for SIR simulation results.
Table A.5. Incidence and estimated effective reproductive numbers for Tanzania from 1998 to 2014.

<table>
<thead>
<tr>
<th>Year</th>
<th>Routine Vaccination</th>
<th>Incidence (per 100,000)</th>
<th>(R_{\text{unadjusted}}^{**})</th>
<th>(R_{\text{adjusted}}^{***})</th>
<th>(V_c, \text{ adj.})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>72%</td>
<td>31.0</td>
<td>4.2</td>
<td>5.2</td>
<td>94.6%</td>
</tr>
<tr>
<td>1999*</td>
<td>72%</td>
<td>17.8</td>
<td>4.2</td>
<td>5.3</td>
<td>94.7%</td>
</tr>
<tr>
<td>2000</td>
<td>78%</td>
<td>43.1</td>
<td>3.3</td>
<td>4.3</td>
<td>94.8%</td>
</tr>
<tr>
<td>2001</td>
<td>86%</td>
<td>34.0</td>
<td>2.1</td>
<td>2.8</td>
<td>94.9%</td>
</tr>
<tr>
<td>2002</td>
<td>89%</td>
<td>14.3</td>
<td>1.6</td>
<td>2.2</td>
<td>95.0%</td>
</tr>
<tr>
<td>2003</td>
<td>97%</td>
<td>4.5</td>
<td>0.5</td>
<td>0.6</td>
<td>95.1%</td>
</tr>
<tr>
<td>2004</td>
<td>94%</td>
<td>3.7</td>
<td>0.9</td>
<td>1.3</td>
<td>95.2%</td>
</tr>
<tr>
<td>2005</td>
<td>91%</td>
<td>0.1</td>
<td>1.3</td>
<td>1.9</td>
<td>95.3%</td>
</tr>
<tr>
<td>2006</td>
<td>93%</td>
<td>5.9</td>
<td>1.0</td>
<td>1.5</td>
<td>95.4%</td>
</tr>
<tr>
<td>2007</td>
<td>90%</td>
<td>18.6</td>
<td>1.5</td>
<td>2.2</td>
<td>95.5%</td>
</tr>
<tr>
<td>2008</td>
<td>88%</td>
<td>8.0</td>
<td>1.8</td>
<td>2.7</td>
<td>95.6%</td>
</tr>
<tr>
<td>2009</td>
<td>91%</td>
<td>3.6</td>
<td>1.3</td>
<td>2.1</td>
<td>95.6%</td>
</tr>
<tr>
<td>2010*</td>
<td>92%</td>
<td>0.4</td>
<td>1.2</td>
<td>1.9</td>
<td>95.7%</td>
</tr>
<tr>
<td>2011</td>
<td>93%</td>
<td>3.4</td>
<td>1.0</td>
<td>1.7</td>
<td>95.8%</td>
</tr>
<tr>
<td>2012</td>
<td>97%</td>
<td>3.4</td>
<td>0.5</td>
<td>0.7</td>
<td>95.9%</td>
</tr>
<tr>
<td>2013</td>
<td>99%</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
<td>95.9%</td>
</tr>
<tr>
<td>2014</td>
<td>99%</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>96.0%</td>
</tr>
</tbody>
</table>

* Years for which DHS data is available; from these the \(\tau(r)\) functions for Tanzania were calculated to estimate clustering of susceptible individuals.

** Unadjusted effective reproductive number; calculated as \(X=X_0(1-v)\).

*** Adjusted effective reproductive number; calculated as \(X=X_0(1-v)\phi\), where \(\phi\) is assumed to change linearly between 1999 and 2010.

Figure A.8. (A) Probability of an outbreak with increasing clustering (\(\phi\)) and effective reproductive number (\(R\)). This assumes an introduction has occurred and an outbreak is defined as 5% of susceptible individuals becoming infected. (B) Probability ratio (PrR) of an outbreak with increasing clustering (\(\phi\)) compared with homogeneous (\(\phi=1\)).
Figure A.9. Correlation between $\phi$ and AUC of $1 - G(r)$ versus $\tau(r)$ curves. All three levels of clustering ($\theta=0.25, 0.5, 1.0$) have perfect correlation with AUC (Pearson’s $r=1.0$). For each combination of $\tau(r)$ and $g(r)$, $\phi = \text{AUC} + 1$.

Figure A.10. The association between $R_0$ and $V_c^*$ at the four defined levels of clustering and three $g(r)$ distributions. As clustering increases, the required vaccination coverage to maintain $R=1$ increases. At low $R_0$, this increase due to clustering is much greater than at high $R_0$. (A) Linear $V_c^*$ scale. (B) Logarithmic $V_c^*$ scale.
Appendix A: Supplement to the challenge of spatial clustering of susceptibility for measles elimination

Figure A.11. Association between $R_0$, clustering maximum ($\theta$ parameter of $\tau(r)$), and the critical vaccination threshold ($V_c$).

Figure A.12. Sensitivity analyses of the varied impact of $\tau(x)$ parameters ($\theta$ and $\lambda$) and $g(x)$ parameters (shape and scale) on the resulting $R_{eff}$ estimate. $\tau(x)$ follows an exponential distribution characterized as $\tau(x) = \theta e^{-\lambda x} + \psi$, while $g(x)$ follows a standard gamma distribution parameterized by shape and scale parameters. Regardless of the $g(x)$ parameters, as $\theta$ increases, $R_{eff}$ increases (A and B).
Figure A.13. Relationship between $R_0$, $\phi$, and $V_c$.

Figure A.14. Analytic $R^*$ estimates for measles assuming $R_0=15$ and using contact distributions derived from Read et al. 2014 ($g_a(r)$) and Noulas et al. 2012 ($g_c(r)$). As clustering increases, $R$ increases; however, this difference decreases as vaccination coverage increases.
Appendix A: Supplement to the challenge of spatial clustering of susceptibility for measles elimination

References


APPENDIX B

Supplement to The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

B.1 EFFECTIVE REPRODUCTIVE NUMBER (R) AND CRITICAL VACCINATION THRESHOLD (V_c) CALCULATIONS

As detailed in Chapter 2, we previously developed and validated modified calculations to estimate the $R$ and $V_c$ values adjusted for spatial clustering of non-vaccination. These calculations assume the unadjusted forms of $R = R_0(1 - v)$ and $V_c = 1 - \frac{1}{R_0}$, where $R_0$ is the basic reproductive number and $v$ is the proportion of the population successfully vaccinated for an infectious disease. We use the clustering adjustment factor ($\phi$) to adjust these, where $\phi$ is the relative probability that an infected individual will be in contact with an unvaccinated individual compared with completely random. $\phi$ is equivalent to $R^*/R$.

The adjust calculations are as follows:

$$R^* = R_0(1 - v) \int_0^\infty g(r)\tau(r) \, dr$$  \[1\]

$$V_c^* = 1 - \frac{1}{R_0\phi}$$  \[2\]

where

$$\phi = \int_0^\infty g(r)\tau(r) \, dr$$  \[3\]

Using the gamma PDF and exponential decay functions of $g(r)$ and $\tau(r)$, we get the closed-form solution for $\phi$ of:

$$\phi = b + \frac{\beta^\alpha}{(\beta + \lambda)^\alpha} \theta$$  \[4\]
Appendix B: Supplement to The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

B.2 Weighted Cluster Survey $\tau(r)$ Calculation

To calculate an accurate $\tau(r)$ using clustered survey data with sampling weights for the clusters, we use a revised calculation to account for the structure of the data and sampling weights.

The standard equation for $\tau(r)$ is:

$$\tau(r_1, r_2) = \frac{\Pr(z_i = z_j | j \in \Omega_i(r_1, r_2))}{\Pr(z_i = z_j | j \in \Omega_i(.))}$$

$$\tau(r_1, r_2) = \frac{\Pr(z_i = z_j = Z | j \in \Omega_i(r_1, r_2))}{\Pr(z_j = Z | j \in \Omega_i(r_1, r_2))}$$

Where $Z$ is the type. In this case $Z =$ unvaccinated status.

See Salje et al. 2014 and Lessler et al. 2016 for more details.1,2

We designed these methods for use with both clustered survey and discrete data. As such, to use with DHS data, we aggregated the individual-level DHS data to cluster-level data, with cluster ID, geospatial coordinates, number of individuals, sampling weights, and proportion susceptible for each cluster. These methods can be directly applied to discrete data such as counties or districts.

First, we define the weighted probability of a susceptible individual within a distance range $(r_1, r_2)$ from cluster $i$:

$$\pi_i(r_1, r_2) = \frac{\sum_{j \in \Omega_i(r_1, r_2)} s_j w_j p_j}{\sum_{j \in \Omega_i(r_1, r_2)} s_j w_j}$$

where:
Appendix B: Supplement to The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

\[ n_i = \text{number of clusters at a distance } (r_1, r_2) \text{ from cluster } i \]
\[ s_j = \text{number of individuals in cluster } j \]
\[ w_j = \text{cluster sampling weight of cluster } j \]
\[ p_j = \text{proportion susceptible in cluster } j \]

From this we can calculate the weighted probability of a susceptible within a distance range \((r_1, r_2)\) from another susceptible in all clusters:

\[
\Pi_\alpha(r_1, r_2) = \frac{1}{\sum_{i=1}^{N} s_i w_i p_i} \sum_{i=1}^{N} \left\{ \frac{\sum_{j \in \Omega_i(d_1, d_2)} s_j w_j p_j}{\sum_{j \in \Omega_i(d_1, d_2)} s_j w_j} s_i w_i p_i \right\}
\]

where:

\[ N = \text{number of clusters in the survey data} \]
\[ s_i = \text{number of individuals in cluster } i \]
\[ w_i = \text{cluster sampling weight of cluster } i \]
\[ p_i = \text{proportion susceptible in cluster } i \]

It follows that we can similarly calculate the weighted probability of a susceptible within a distance range \((r_1, r_2)\) from any individual in all clusters:

\[
\Pi_\beta(r_1, r_2) = \frac{1}{\sum_{i=1}^{N} s_i w_i} \sum_{i=1}^{N} \left\{ \frac{\sum_{j \in \Omega_i(d_1, d_2)} s_j w_j p_j}{\sum_{j \in \Omega_i(d_1, d_2)} s_j w_j} s_i w_i \right\}
\]

Finally, applying this to the original format of \(\tau(r)\):

\[
\tau(r_1, r_2) = \frac{\Pr(z_i = z_j = Z | j \in \Omega_i(r_1, r_2))}{\Pr(z_j = Z | j \in \Omega_i(r_1, r_2))}
\]
\[ \tau(r_1, r_2) = \frac{\Pi_a(r_1, r_2)}{\Pi_\beta(r_1, r_2)} \]

### B.3 \( \tau(r) \) Distributions

We used an exponential function consistently characterize \( \tau(r) \):

\[ \tau(r) = \theta e^{-\lambda r} + \psi \]

Where \( \theta \) defines the maximum probability ratio of being in contact with a susceptible at a distance of 0 \( (r = 0) \). \( \lambda \) defines the rate of decay of the probability ratio with distance \( r \). \( \psi \) is a rescaling value, allowing the distribution to average to 1 across the entire range of \( r \). Empirical \( \tau(r) \) estimates were fit to this exponential function using non-linear least squares with inverse variance weighting.

### B.4 \( g(r) \) Distributions

We used the same three contact distance distributions as previously: \( g_A(r) \) from China, \( g_B(r) \) from Zambia, and \( g_C(r) \) from the U.S. Once again, for efficient calculation, we assume a gamma distribution for \( g(r) \), where \( g(r) \sim \Gamma(\alpha, \beta) \). This 2-parameter distribution provides high flexibility to both fit empirical distributions, and explore the effects of the shape of the contact distribution on the resulting \( \phi, R^* \), and \( V_c^* \). Fitting of gamma \( g(r) \) distributions to empirical data was done through mathematical optimization using the Nelder-Mead method, through the *stats R* package.

### B.5 METHODS FOR APPLICATION TO CLUSTER SURVEY DATA

#### B.5.1 Moving Window \( \tau(r) \) Estimation

To approximate a continuous \( \tau(r) \) function from discrete cluster locations, we used a moving window approach, rather than a static set of discrete distance windows. For example, instead of using 2km discrete
windows, such as 0-2km, 2-4km, 4-6km, etc. We used overlapping windows, such as 0-2km, 1-3km, 2-4km, 3-5km, 4-6km, etc. This provided a much smoother empirical estimate, particularly for distances with reduced numbers of cluster pairs.

**B.5.2 Jittered Bootstrap Estimates**

To protect the identities of surveys subjects, DHS randomly jitters the location of each cluster, with urban clusters jittered up to 2km in any direction and rural clusters jittered up to 5km. To account for this potential variability in the true locations of the clusters, we incorporated a random jitter of cluster locations into a bootstrap framework of 10,000 iterations. This improved the smoothness of the curve, particularly for distances with low numbers of cluster pairs.

**B.5.3 Identification of Outlying Clusters**

Outlying clusters were identified through a jackknife approach whereby for each iteration, a sampling cluster was removed from the calculation of $\tau(r)$. At each $r$ iteration values are compared. If value from an iteration is $\geq 3$ standard deviations from the mean at that $r$, it is considered an outlying value. Clusters were considered outliers if they produced outlying values at $\geq 10\%$ of $r$ values.

**B.6 COMPARISON OF CLUSTERING**

**B.6.1 Outbreak probability**

We estimated the probability of an outbreak occurring, assuming an introduction has happened, analytically and through simulations. Analytically, this can be estimated as $1 - \frac{1}{R_0}$. For simulated estimations, we conducted SIR simulations using the $R^*$ estimates and estimated the probability of an outbreak as the proportion of introductions that resulted in outbreaks. When $R$ is high enough ($\geq 1$), we begin to see bimodal distributions of outbreak final sizes, where there is a clear distinction between introductions that resulted in outbreaks and those that died off quickly. However, because with a lower $R$
we see less distinction, we designated a threshold of infection of ≥5% of non-vaccinated individuals as being an outbreak.

**B.6.2 Regression analyses**

We used multiple linear regression models to examine the association between effective reproductive numbers ($R$ and $R^*$) and probability of outbreak given introduction, and measles population incidence. Additional variables were examined for association with incidence including region and being an island nation (Philippines, Indonesia, Madagascar, and Comoros) as these are indicators of different risks of introduction. None of these additional variables were significant in the multivariable model. Model fits were compared for unclustered versus clustered estimates of the effective reproductive number and outbreak probability. Measles incidence outcome variables included both reported and estimated true incidence, which were estimated using methods from Chen et al. 2012. To capture the overall effect of the $R$ estimates over time, we used mean and median incidence and variables, with the mean or median taken for plus and minus 2 years from the date of the DHS survey.

**B.7 Sensitivity Analyses**

**B.7.1 Household Clustering Exclusion and Sample Weighting**

Inclusion and exclusion of household clustering and sample weighting using DHS-provided weights were examined (Figure B.1). While household clustering inclusion did not largely affect the results, exclusion was chosen. Sample weighting was not done for the final fitting as it was determined that use of the sample weights from DHS were be potentially invalid in this application.
Figure B.1. Empirical and fitted $\tau(r)$ functions including and excluding household-based clustering and sample weighting.
B.7.2 Exclusion of Outliers

Figure B.2. Empirical $\tau(r)$ results with and excluding four outlying clusters. No obvious difference was found.

B.7.3 Impact of Urban-focused Vaccination Efforts

Because the $\tau$-statistic requires pairs of individuals or pairs of locations, bias may exist at particular distances because those pairwise distances are only being contributed to by pairs with particular characteristics, such as urban location. The majority of short distance pairs comes from urban areas, including cities and towns, where multiple clusters are sampled within a short distance because of the high density of people. We consistently observed an initial dip in empirical $\tau(r)$ functions for several countries. To see if we could replicate this effect, we used simulation to examine the impact of prioritizing vaccination in rural or urban areas. By prioritizing vaccination we produce a $\tau(r)$ with the initial dip, similar (though exaggerated) to the original survey results (Figure B.3).
Appendix B: Supplement to The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

A. Original Clusters

B. Increased Urban Vaccination (urban mean=95%, rural mean=70%)

C. Increased Rural Vaccination (urban mean=70%, rural mean=95%)

Figure B.3. Results of simulated urban rural disparities in vaccination coverage.
B.7.4 Validation of Consistency of Cluster Survey Results

To validate that clusters consistently will produce similar $\tau(r)$ functions, we generated synthetic populations and simulated cluster surveys using the DHS sampling methods. Resulting $\tau(r)$ functions were similar.

![Simulated Clusters and DHS Clusters](image)

Figure B.4. Cluster locations of simulated clusters and true sampling clusters from Zambia.

B.7.5 Generalized Additive Models for Modeled Surface Estimation

Using a generalized additive model, we produced smoothed surfaces of vaccination coverage from DHS survey data. While this produced $\tau(r)$ functions that were smoother, they also washed out some of the initial peak of the function.
Figure BS.5. Results from GAM smoothing.

### B.8 Extrapolation of Methods to Other Forms of Data

#### B.8.1 Measuring spatial clustering using modeled surfaces

Production of modeled surfaces demonstrating characteristics geospatially has become common across the disease spectrum, including measles vaccination coverage. The benefit of these methods is they account for the entire surface of a locality through use of covariates, even when measles vaccination data
Appendix B: Supplement to The Existence of Local Clustering of Non-Vaccination and Its Impact on Outbreak Risk

does not exist. The result is smoother estimates of clustering, though potentially biased conservatively, as by nature of their production correlation is imposed for short distances. However, with continuing development in spatial statistics and modeling and new efforts to produce these maps, which include work by the DHS and WorldPop, these modeled surfaces provide an additional and potentially more useable form of data for estimation of susceptibility clustering.

Figure B.6 Modeled surface estimation of clustering, Togo 2013.
B.9 ADDITIONAL SUPPLEMENTAL TABLES

Table B.1 Exclusion criteria and numbers of surveys included/excluded with each criterium.

<table>
<thead>
<tr>
<th>Exclusion Criteria</th>
<th>Included Surveys</th>
<th>Excluded Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Surveys</td>
<td>109</td>
<td></td>
</tr>
<tr>
<td>GPS data available</td>
<td>75</td>
<td>36</td>
</tr>
<tr>
<td>Years: 2000-2015</td>
<td>70</td>
<td>5</td>
</tr>
<tr>
<td>Vaccination &gt;50%*</td>
<td>66</td>
<td>4</td>
</tr>
<tr>
<td>Consistent tau(r) fitted</td>
<td>59</td>
<td>7</td>
</tr>
</tbody>
</table>
Table B.2. Estimated clustering adjusted critical vaccination thresholds ($V_c$) for the most recent survey from each of the 33 included countries. $R_0=15$.

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey Year</th>
<th>Median</th>
<th>95% CrI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albania</td>
<td>2008</td>
<td>98.0%</td>
<td>96.8 - 98.8%</td>
</tr>
<tr>
<td>Armenia</td>
<td>2010</td>
<td>95.3%</td>
<td>93.7 - 96.4%</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2011</td>
<td>94.4%</td>
<td>92.9 - 95.5%</td>
</tr>
<tr>
<td>Benin</td>
<td>2011</td>
<td>95.1%</td>
<td>93.7 - 96%</td>
</tr>
<tr>
<td>Cambodia</td>
<td>2014</td>
<td>95.0%</td>
<td>93.6 - 96%</td>
</tr>
<tr>
<td>Comoros</td>
<td>2012</td>
<td>94.3%</td>
<td>92.7 - 95.4%</td>
</tr>
<tr>
<td>Cote d'Ivoire</td>
<td>2011</td>
<td>94.7%</td>
<td>93.2 - 95.7%</td>
</tr>
<tr>
<td>Egypt</td>
<td>2014</td>
<td>94.1%</td>
<td>92.6 - 95.2%</td>
</tr>
<tr>
<td>Gabon</td>
<td>2012</td>
<td>94.9%</td>
<td>93.4 - 95.9%</td>
</tr>
<tr>
<td>Ghana</td>
<td>2014</td>
<td>94.3%</td>
<td>92.8 - 95.4%</td>
</tr>
<tr>
<td>Guinea</td>
<td>2005</td>
<td>94.7%</td>
<td>93.2 - 95.7%</td>
</tr>
<tr>
<td>Indonesia</td>
<td>2002</td>
<td>95.1%</td>
<td>93.7 - 96%</td>
</tr>
<tr>
<td>Jordan</td>
<td>2012</td>
<td>94.4%</td>
<td>92.9 - 95.4%</td>
</tr>
<tr>
<td>Kenya</td>
<td>2014</td>
<td>95.3%</td>
<td>93.9 - 96.2%</td>
</tr>
<tr>
<td>Kenya</td>
<td>2008</td>
<td>95.0%</td>
<td>93.5 - 96%</td>
</tr>
<tr>
<td>Lesotho</td>
<td>2014</td>
<td>94.7%</td>
<td>93.2 - 95.7%</td>
</tr>
<tr>
<td>Madagascar</td>
<td>2008</td>
<td>95.2%</td>
<td>93.8 - 96.1%</td>
</tr>
<tr>
<td>Malawi</td>
<td>2010</td>
<td>95.7%</td>
<td>94.4 - 96.5%</td>
</tr>
<tr>
<td>Mali</td>
<td>2012</td>
<td>94.7%</td>
<td>93.2 - 95.6%</td>
</tr>
<tr>
<td>Morocco</td>
<td>2003</td>
<td>95.3%</td>
<td>93.9 - 96.3%</td>
</tr>
<tr>
<td>Mozambique</td>
<td>2011</td>
<td>94.7%</td>
<td>93.1 - 95.7%</td>
</tr>
<tr>
<td>Namibia</td>
<td>2013</td>
<td>95.3%</td>
<td>93.9 - 96.3%</td>
</tr>
<tr>
<td>Nepal</td>
<td>2006</td>
<td>95.7%</td>
<td>94.2 - 96.8%</td>
</tr>
<tr>
<td>Pakistan</td>
<td>2006</td>
<td>94.7%</td>
<td>93.3 - 95.7%</td>
</tr>
<tr>
<td>Philippines</td>
<td>2008</td>
<td>94.8%</td>
<td>93.4 - 95.8%</td>
</tr>
<tr>
<td>Rwanda</td>
<td>2005</td>
<td>96.4%</td>
<td>95 - 97.3%</td>
</tr>
<tr>
<td>Senegal</td>
<td>2005</td>
<td>94.9%</td>
<td>93.4 - 95.9%</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>2013</td>
<td>94.3%</td>
<td>92.8 - 95.4%</td>
</tr>
<tr>
<td>Swaziland</td>
<td>2006</td>
<td>94.9%</td>
<td>93.4 - 96%</td>
</tr>
<tr>
<td>Tanzania</td>
<td>2015</td>
<td>94.5%</td>
<td>93 - 95.5%</td>
</tr>
<tr>
<td>Timor-Leste</td>
<td>2009</td>
<td>94.5%</td>
<td>93 - 95.5%</td>
</tr>
<tr>
<td>Uganda</td>
<td>2006</td>
<td>94.7%</td>
<td>93.1 - 95.7%</td>
</tr>
<tr>
<td>Uganda</td>
<td>2011</td>
<td>94.1%</td>
<td>92.5 - 95.2%</td>
</tr>
<tr>
<td>Zambia</td>
<td>2013</td>
<td>95.0%</td>
<td>93.6 - 96%</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>2010</td>
<td>94.7%</td>
<td>93.1 - 95.7%</td>
</tr>
</tbody>
</table>
*Table B.3 Estimated effective reproductive numbers ($R^*$) for the most recent survey from each of the 33 included countries. $R_0=15.*

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey Year</th>
<th>$R^*$ Median</th>
<th>95% CrI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albania</td>
<td>2008</td>
<td>1.0</td>
<td>0.6 - 1.7</td>
</tr>
<tr>
<td>Armenia</td>
<td>2010</td>
<td>0.6</td>
<td>0.5 - 0.8</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2014</td>
<td>2.2</td>
<td>1.7 - 2.7</td>
</tr>
<tr>
<td>Benin</td>
<td>2011</td>
<td>6.1</td>
<td>4.8 - 7.5</td>
</tr>
<tr>
<td>Cambodia</td>
<td>2014</td>
<td>1.2</td>
<td>0.9 - 1.5</td>
</tr>
<tr>
<td>Comoros</td>
<td>2012</td>
<td>2.6</td>
<td>2.1 - 3.2</td>
</tr>
<tr>
<td>Cote d’Ivoire</td>
<td>2011</td>
<td>9.7</td>
<td>7.5 - 12</td>
</tr>
<tr>
<td>Egypt</td>
<td>2014</td>
<td>1.2</td>
<td>0.9 - 1.4</td>
</tr>
<tr>
<td>Gabon</td>
<td>2012</td>
<td>5.7</td>
<td>4.4 - 7.1</td>
</tr>
<tr>
<td>Ghana</td>
<td>2014</td>
<td>1.4</td>
<td>1.1 - 1.7</td>
</tr>
<tr>
<td>Guinea</td>
<td>2005</td>
<td>9.3</td>
<td>7.2 - 11.5</td>
</tr>
<tr>
<td>Indonesia</td>
<td>2002</td>
<td>5.7</td>
<td>4.5 - 7</td>
</tr>
<tr>
<td>Jordan</td>
<td>2012</td>
<td>0.4</td>
<td>0.3 - 0.4</td>
</tr>
<tr>
<td>Kenya</td>
<td>2014</td>
<td>4.4</td>
<td>3.5 - 5.5</td>
</tr>
<tr>
<td>Lesotho</td>
<td>2014</td>
<td>1.5</td>
<td>1.2 - 1.9</td>
</tr>
<tr>
<td>Madagascar</td>
<td>2008</td>
<td>6.3</td>
<td>4.9 - 7.8</td>
</tr>
<tr>
<td>Malawi</td>
<td>2010</td>
<td>1.6</td>
<td>1.2 - 2</td>
</tr>
<tr>
<td>Mali</td>
<td>2012</td>
<td>4.7</td>
<td>3.7 - 5.7</td>
</tr>
<tr>
<td>Morocco</td>
<td>2003</td>
<td>2.1</td>
<td>1.6 - 2.7</td>
</tr>
<tr>
<td>Mozambique</td>
<td>2011</td>
<td>3.4</td>
<td>2.6 - 4.2</td>
</tr>
<tr>
<td>Namibia</td>
<td>2013</td>
<td>3.8</td>
<td>3 - 4.8</td>
</tr>
<tr>
<td>Nepal</td>
<td>2006</td>
<td>3.5</td>
<td>2.6 - 4.6</td>
</tr>
<tr>
<td>Pakistan</td>
<td>2006</td>
<td>5.5</td>
<td>4.3 - 6.7</td>
</tr>
<tr>
<td>Philippines</td>
<td>2008</td>
<td>1.5</td>
<td>1.2 - 1.9</td>
</tr>
<tr>
<td>Rwanda</td>
<td>2005</td>
<td>3.0</td>
<td>2.2 - 4.1</td>
</tr>
<tr>
<td>Senegal</td>
<td>2005</td>
<td>5.1</td>
<td>3.9 - 6.3</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>2013</td>
<td>3.0</td>
<td>2.3 - 3.7</td>
</tr>
<tr>
<td>Swaziland</td>
<td>2006</td>
<td>1.4</td>
<td>1.1 - 1.7</td>
</tr>
<tr>
<td>Tanzania</td>
<td>2010</td>
<td>2.3</td>
<td>1.7 - 3.1</td>
</tr>
<tr>
<td>Timor-Leste</td>
<td>2009</td>
<td>5.4</td>
<td>4.3 - 6.7</td>
</tr>
<tr>
<td>Uganda</td>
<td>2011</td>
<td>4.2</td>
<td>3.3 - 5.2</td>
</tr>
<tr>
<td>Zambia</td>
<td>2013</td>
<td>4.0</td>
<td>3.1 - 4.9</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>2010</td>
<td>1.9</td>
<td>1.5 - 2.3</td>
</tr>
<tr>
<td>Tanzania</td>
<td>2015</td>
<td>0.2</td>
<td>0.1 - 0.2</td>
</tr>
</tbody>
</table>
Table B.4. Regression model likelihood estimates and likelihood ratio comparing clustering adjusted and unadjusted models.

<table>
<thead>
<tr>
<th>Measles Incidence Variable</th>
<th>Log Likelihood</th>
<th>Likelihood Ratio (Clustered vs. Unclustered)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted</td>
<td>Clustering Adjusted</td>
</tr>
<tr>
<td><strong>Outcome Variable: Effective Reproductive Number (R)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reported to WHO, year of survey</td>
<td>-147.42</td>
<td>-147.63</td>
</tr>
<tr>
<td>Reported to WHO, 5yr mean</td>
<td>-119.8</td>
<td>-119.1</td>
</tr>
<tr>
<td>Reported to WHO, 5yr median</td>
<td>-131.4</td>
<td>-131.5</td>
</tr>
<tr>
<td>Estimated true incidence, 5yr mean</td>
<td>-132.6</td>
<td>-132.9</td>
</tr>
<tr>
<td>Estimated true incidence, 5yr median</td>
<td>-149.6</td>
<td>-149.7</td>
</tr>
<tr>
<td>**Outcome Variable: Pr(Outbreak</td>
<td>Introduction)**</td>
<td></td>
</tr>
<tr>
<td>Reported to WHO, year of survey</td>
<td>-148.31</td>
<td>-147.09</td>
</tr>
<tr>
<td>Reported to WHO, 5yr mean</td>
<td>-121.0</td>
<td>-118.8</td>
</tr>
<tr>
<td>Reported to WHO, 5yr median</td>
<td>-132.1</td>
<td>-131.2</td>
</tr>
<tr>
<td>Estimated true incidence, 5yr mean</td>
<td>-132.5</td>
<td>-132.0</td>
</tr>
<tr>
<td>Estimated true incidence, 5yr median</td>
<td>-149.4</td>
<td>-149.0</td>
</tr>
</tbody>
</table>
Figure B.7. Case study countries.
Figure B.8. (a) Unadjusted and (b) spatially clustering adjusted effective reproductive estimates, using Demographics and Health Survey data and WHO/UNICEF vaccination coverage estimates.
Figure B.9 Relative probability of outbreak, clustering adjusted probability versus unadjusted probability estimates.
References


APPENDIX C

Supplement to Biosecurity in a Globalized World:
The Case of Measles

C.1 MATHEMATICAL MODEL FOR THE IMPORTATION OF MEASLES-INFECTED AIR TRAVELERS INTO THE U.S.

Here, we give a detailed explanation of our model for the importation of measles-infected individuals via air travel into the U.S. For each source $s$, we collect monthly data on measles incidence at month $m$ ($I_{s,m}$) and the number of monthly air trips to every U.S. destination $d$, denoted by $T_{d,s,m}$. We consider that each of these infected individuals contributes a certain number of ‘infectious person days’ during which travel to the U.S. might occur; this time span is a random variable denoted by $D$. On each of these days, we consider that these individuals have a daily probability of travel from source $s$ to destination $d$ at month $m$, given by the number of monthly air trips divided by the total number of person days in that month:

$$p_{d,s,m} = \frac{T_{d,s,m}}{N_s \delta_m}$$

where $N_{s,m}$ is the total population at source $s$ on month $m$ and $\delta_m$ is the number of days in month $m$. Putting these pieces together, our framework posits that the number of measles-infected individuals introduced from source $s$ to destination $d$ on month $m$ is given by the binomial variable:

$$n_{d,s,m} = Bin(p_{d,s,m}, I_{s,m}D)$$
This framework allows for simple expansion if more data are available, including reporting rates or knowledge of the likelihood of traveling while infected. However, for this initial application we maintained the simplified form.

C.1.1 Monthly Measles Incidence

Monthly measles cases are obtained from WHO (ref) from January 2002 to December 2016. Due to incomplete surveillance and reporting, this was truncated to January 2006 through December 2016.

C.1.2 Monthly Air Trips from Source Locations to U.S. States

We collected itinerary data from all source locations to all U.S. airports during the span of Jan 2005 to Dec 2015 from the International Air Transport Association (IATA).

C.1.3 Infectious Person Days

This random variable represents the number of days during which a person infected with the measles virus may travel to the U.S. We consider that it includes the whole incubation period and the infectious days before rash starts, as air travel is unlikely after the onset of rash combined with other symptoms.

The incubation period of measles has been estimated to be 12.5 (95% CI: 11.8–13.3); while the rash appears 2–3 days after incubation concludes.¹ We thus take the ‘infectious person days’ \((D)\) to be the sum of a normal random variable with a mean of 12.5 and 95% CI of 11.8–13.3, representing the incubation period, and a uniform random variable between 2 and 3 days, representing rash onset after incubation; the result of the sum is rounded in order to obtain an integer number of days.
Table C.1. Model Parameters

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I_{s,m} )</td>
<td>Monthly measles cases at source ( s ) and month ( m )</td>
</tr>
<tr>
<td>( D )</td>
<td>Number of days of incubation during which an infected individual might travel</td>
</tr>
<tr>
<td>( T_{s,d,m} )</td>
<td>Number of monthly air trips from source ( s ) to destination ( d ) on month ( m )</td>
</tr>
<tr>
<td>( P_{l,m} )</td>
<td>Population of sources on month ( m )</td>
</tr>
<tr>
<td>( D_{m} )</td>
<td>Number of days in month ( m )</td>
</tr>
<tr>
<td>( n_{i,d,m} )</td>
<td>Number of monthly imported measles cases from source ( s ) to destination ( d ) on month ( m )</td>
</tr>
</tbody>
</table>

C.2 MODEL VALIDATION

Model validation was performed through probability models which were fit to reported importation cases. We estimated the relative likelihoods of each validation model as compared with the null validation model (VM\(_0\)), given the reported importation cases from the CDC. These models were as follows:

Validation Models:

\[ VM_0 \text{ (Null): } \rho_{i,s,m} = \frac{1}{M+S+t} \]

\[ VM_1 \text{ (Incidence only): } \rho_{i,s,m} = \frac{w_{m,s}}{\sum_{i=1}^{M} \sum_{t=0}^{S} w_{m,i}} \]

\[ VM_2 \text{ (Travel only): } \rho_{i,s,m} = \frac{w_{M,s,i}}{\sum_{i=1}^{M} \sum_{t=0}^{S} \sum_{l=1}^{I} w_{M,i}} \]

\[ VM_3 \text{ (Full): } \rho_{i,s,m} = \frac{w_{F,m,s,i}}{\sum_{i=1}^{M} \sum_{t=0}^{S} \sum_{l=1}^{I} w_{F,m,i}} \]

Table C.2. Validation model likelihoods and likelihood ratio compared to the null model.

<table>
<thead>
<tr>
<th>Model</th>
<th>log(Likelihood Ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null (reference)</td>
<td></td>
</tr>
<tr>
<td>Incidence only</td>
<td>53.09</td>
</tr>
<tr>
<td>Travel only</td>
<td>74.92</td>
</tr>
<tr>
<td>Full</td>
<td>96.86</td>
</tr>
</tbody>
</table>
C.3 **Estimated Numbers of Importations and Exportations**

The model produces actual count of importations and exportations. We chose to estimate the number, rather than the probability, of importations to better correlate with true reported importation counts and because in some instances, the numbers provide more information on risk than the probability. For example, there are several months, specifically during 2011, when New York/New Jersey/Connecticut had 100% probability of importation (i.e. each simulation produced at least one importation). However, the difference between 1 importation and 10 importations is substantial, especially when each importation can produce an outbreak. The problem with exact counts, though, is that there are multiple factors, both unmeasured and random, that affect the exact number, thus we opted to use relative risks to avoid incorrect exact counts, and rather focus on the risk between states, countries, and times.

![Reported Measles Cases by Country](image_url)

Figure C.1. Reported measles incidence, Europe, June 1999 to April 2016.
Figure C.2. Estimated measles cases imported to the United States from Europe, by country of origin, June 1999 to April 2016.
C.4 ESTIMATING RELATIVE RISK

To limit the bias and avoid inaccurate reporting of importation counts, we used relative risks, calculated from the simulation results of estimated counts of reported importations and exportations.

C.4.1 Relative risk versus the median risk across all time and locations

The primary results presented were calculated with the reference group (i.e. denominator) being the median count of importation across time and states or countries. For example, for monthly risk calculations, for each simulation (n=10,000) a median importation count was estimated across the 120 months and either 43 destination states/state groups/territories or 28 source countries. Thus, for importation risk the reference for each simulation was the median of 5,160 values, and for exportation
risk the median of 3,360 values. With these medians, we calculated a relative risk (RR) for each combination of time and location for each of the 10,000 simulations. We took the log-adjusted means across the 10,000 simulations for each location to estimate a RR for each time period.

For aggregate time periods, such as year, outbreak period, or overall (across the 10 years), we first summed the estimated importation counts within each simulation, then estimated RR.

Relative risk estimates were also adjusted for population and traveler volume. For this the estimated importation counts were first adjusted, then the RRs were calculated the same as before. The full results for overall RR of importation of measles into U.S. destinations and exportation of measles from European sources is shown in Tables C.2 and C.3.
## Appendix C: Supplement to Biosecurity in a Globalized World: The Case of Measles

Table C.2. Overall estimated relative risks of measles importation by state during January 2006 – December 2015, sorted by state mean mean population. Relative risk uses the median risk across all states as the reference.

<table>
<thead>
<tr>
<th>Destination State</th>
<th>Mean Population</th>
<th>Overall</th>
<th>Per Capita</th>
<th>Per Traveler</th>
</tr>
</thead>
<tbody>
<tr>
<td>NY/NJ/CT</td>
<td>31,859,927</td>
<td>40 (29-54)</td>
<td>5.2 (3.3-9)</td>
<td>1 (0.74-1.4)</td>
</tr>
<tr>
<td>California</td>
<td>37,647,061</td>
<td>20 (13-30)</td>
<td>2.2 (1.3-4)</td>
<td>1 (0.66-1.5)</td>
</tr>
<tr>
<td>Florida</td>
<td>19,212,100</td>
<td>18 (11-26)</td>
<td>3.8 (2.1-7)</td>
<td>1 (0.62-1.5)</td>
</tr>
<tr>
<td>IL/WI/IN</td>
<td>25,006,142</td>
<td>9.3 (4.4-16)</td>
<td>1.5 (0.76-3)</td>
<td>1.1 (0.55-1.8)</td>
</tr>
<tr>
<td>DC/MD/VA</td>
<td>14,530,686</td>
<td>8.6 (4.4-15)</td>
<td>2.4 (1.1-4.8)</td>
<td>1.1 (0.57-1.9)</td>
</tr>
<tr>
<td>MA/ME/RI</td>
<td>8,994,548</td>
<td>8.4 (4.4-14)</td>
<td>3.8 (1.7-7.6)</td>
<td>1.1 (0.54-1.9)</td>
</tr>
<tr>
<td>Texas</td>
<td>25,634,112</td>
<td>7.4 (3.5-13)</td>
<td>1.2 (0.55-2.3)</td>
<td>1.1 (0.53-2)</td>
</tr>
<tr>
<td>PA/DE</td>
<td>13,610,750</td>
<td>4.6 (1.8-8.8)</td>
<td>1.4 (0.52-3.1)</td>
<td>1.2 (0.45-2.4)</td>
</tr>
<tr>
<td>Georgia</td>
<td>9,786,657</td>
<td>3.8 (1.7-7.9)</td>
<td>1.6 (0.56-3.7)</td>
<td>1.2 (0.47-2.5)</td>
</tr>
<tr>
<td>Nevada</td>
<td>2,734,504</td>
<td>3.7 (1.7-7.9)</td>
<td>5.6 (1.9-13)</td>
<td>1.1 (0.43-2.4)</td>
</tr>
<tr>
<td>North Carolina</td>
<td>9,609,849</td>
<td>2.9 (0.83-6.1)</td>
<td>1.3 (0.36-3.1)</td>
<td>1.2 (0.36-2.8)</td>
</tr>
<tr>
<td>Washington</td>
<td>6,817,587</td>
<td>2.9 (0.82-6.1)</td>
<td>1.7 (0.49-4.4)</td>
<td>1.2 (0.36-2.7)</td>
</tr>
<tr>
<td>Michigan</td>
<td>9,926,141</td>
<td>2.8 (0.82-6.1)</td>
<td>1.2 (0.33-2.8)</td>
<td>1.2 (0.36-2.8)</td>
</tr>
<tr>
<td>Colorado</td>
<td>5,122,700</td>
<td>2.5 (0.79-6.1)</td>
<td>2 (0.62-5.2)</td>
<td>1.2 (0.36-2.8)</td>
</tr>
<tr>
<td>Arizona</td>
<td>6,485,336</td>
<td>2 (0.73-5.1)</td>
<td>1.2 (0.43-4.3)</td>
<td>1.3 (0.47-3.3)</td>
</tr>
<tr>
<td>Minnesota</td>
<td>5,346,635</td>
<td>1.9 (0.72-5.1)</td>
<td>1.5 (0.49-4.1)</td>
<td>1.3 (0.46-3.3)</td>
</tr>
<tr>
<td>Ohio</td>
<td>5,114,682</td>
<td>1.5 (0.65-4.2)</td>
<td>0.55 (0.18-1.5)</td>
<td>1.3 (0.52-3.7)</td>
</tr>
<tr>
<td>Missouri</td>
<td>5,992,765</td>
<td>1.5 (0.64-4.2)</td>
<td>1 (0.36-3)</td>
<td>1.4 (0.54-3.8)</td>
</tr>
<tr>
<td>Tennessee</td>
<td>6,392,024</td>
<td>1.3 (0.6-4.1)</td>
<td>0.86 (0.35-2.5)</td>
<td>1.4 (0.58-4)</td>
</tr>
<tr>
<td>Kentucky</td>
<td>4,350,982</td>
<td>1.3 (0.57-3.2)</td>
<td>1.2 (0.44-3.6)</td>
<td>1.4 (0.59-3.8)</td>
</tr>
<tr>
<td>Oregon</td>
<td>3,871,642</td>
<td>1.3 (0.58-4.1)</td>
<td>1.4 (0.54-3.1)</td>
<td>1.4 (0.59-4.1)</td>
</tr>
<tr>
<td>Louisiana</td>
<td>4,541,292</td>
<td>0.95 (0.48-3.1)</td>
<td>0.86 (0.32-2.9)</td>
<td>1.4 (0.64-5)</td>
</tr>
<tr>
<td>Utah</td>
<td>2,803,935</td>
<td>0.87 (0.45-3.1)</td>
<td>1.3 (0.52-4.5)</td>
<td>1.4 (0.66-4.8)</td>
</tr>
<tr>
<td>South Carolina</td>
<td>4,672,949</td>
<td>0.74 (0.4-3)</td>
<td>0.66 (0.26-2.4)</td>
<td>1.4 (0.68-5.1)</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>3,648,788</td>
<td>0.66 (0.37-2.1)</td>
<td>0.74 (0.31-2.6)</td>
<td>1.3 (0.64-4.4)</td>
</tr>
<tr>
<td>Hawaii</td>
<td>1,374,046</td>
<td>0.55 (0.32-2.1)</td>
<td>1.6 (0.71-6.9)</td>
<td>1.1 (0.59-4.4)</td>
</tr>
<tr>
<td>Alabama</td>
<td>4,779,018</td>
<td>0.49 (0.29-2.1)</td>
<td>0.42 (0.18-1.8)</td>
<td>1.4 (0.76-6)</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>3,776,255</td>
<td>0.47 (0.29-2)</td>
<td>0.51 (0.23-2.3)</td>
<td>1.4 (0.76-6.3)</td>
</tr>
<tr>
<td>New Mexico</td>
<td>2,050,502</td>
<td>0.35 (0.23-2)</td>
<td>0.7 (0.33-3.6)</td>
<td>1.4 (0.78-7.5)</td>
</tr>
<tr>
<td>Iowa</td>
<td>3,061,853</td>
<td>0.32 (0.21-2)</td>
<td>0.43 (0.21-2.3)</td>
<td>1.4 (0.81-8)</td>
</tr>
<tr>
<td>Nebraska</td>
<td>1,840,349</td>
<td>0.28 (0.19-2)</td>
<td>0.63 (0.31-3.6)</td>
<td>1.4 (0.81-8.2)</td>
</tr>
<tr>
<td>Arkansas</td>
<td>2,922,332</td>
<td>0.26 (0.18-1.1)</td>
<td>0.36 (0.18-2.1)</td>
<td>1.4 (0.85-6.3)</td>
</tr>
<tr>
<td>Alaska</td>
<td>714,816</td>
<td>0.22 (0.16-1)</td>
<td>1.3 (0.66-8.5)</td>
<td>1.2 (0.75-6.1)</td>
</tr>
<tr>
<td>Montana</td>
<td>998,475</td>
<td>0.17 (0.12-1)</td>
<td>0.69 (0.37-4.6)</td>
<td>1.2 (0.77-7.5)</td>
</tr>
<tr>
<td>Idaho</td>
<td>1,581,373</td>
<td>0.16 (0.12-1)</td>
<td>0.42 (0.22-2.9)</td>
<td>1.2 (0.81-8.1)</td>
</tr>
<tr>
<td>Kansas</td>
<td>2,855,209</td>
<td>0.15 (0.12-1)</td>
<td>0.22 (0.12-1.6)</td>
<td>1.3 (0.85-9)</td>
</tr>
<tr>
<td>Mississippi</td>
<td>2,966,758</td>
<td>0.14 (0.11-0.99)</td>
<td>0.2 (0.11-1.5)</td>
<td>1.3 (0.86-9.5)</td>
</tr>
<tr>
<td>Vermont</td>
<td>625,391</td>
<td>0.1 (0.08-0.96)</td>
<td>0.66 (0.37-6.5)</td>
<td>1.3 (0.89-13)</td>
</tr>
<tr>
<td>North Dakota</td>
<td>696,871</td>
<td>0.096 (0.077-0.96)</td>
<td>0.57 (0.33-5.8)</td>
<td>1.2 (0.8-12)</td>
</tr>
<tr>
<td>South Dakota</td>
<td>825,170</td>
<td>0.095 (0.076-0.96)</td>
<td>0.47 (0.27-4.9)</td>
<td>1.2 (0.81-12)</td>
</tr>
<tr>
<td>Wyoming</td>
<td>564,620</td>
<td>0.055 (0.047-0.93)</td>
<td>0.4 (0.25-6.5)</td>
<td>1 (0.74-17)</td>
</tr>
<tr>
<td>West Virginia</td>
<td>1,844,519</td>
<td>0.05 (0.043-0.92)</td>
<td>0.11 (0.069-1.9)</td>
<td>1.1 (0.79-20)</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>1,320,527</td>
<td>0.027 (0.024-0.9)</td>
<td>0.082 (0.053-2)</td>
<td>0.9 (0.69-26)</td>
</tr>
</tbody>
</table>
Table C.3. Overall estimated relative risk of exporting measles by European countries, during January 2006 – December 2015, sorted by country mean population. Relative risk uses the median risk across all countries as the reference.

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Measles Incidence</th>
<th>Overall RR</th>
<th>Per Capita RR</th>
<th>Per Traveler RR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Population</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>UK</td>
<td>63,225,965</td>
<td>87 (31-170)</td>
<td>8.9 (3.7-20)</td>
<td>3.7 (1.6-8.4)</td>
</tr>
<tr>
<td>France</td>
<td>65,327,311</td>
<td>68 (24-140)</td>
<td>6.8 (2.7-15)</td>
<td>7.7 (3.3-18)</td>
</tr>
<tr>
<td>Italy</td>
<td>59,550,475</td>
<td>47 (16-98)</td>
<td>5.1 (2-12)</td>
<td>5.5 (2.3-13)</td>
</tr>
<tr>
<td>Germany</td>
<td>81,555,942</td>
<td>30 (9.7-66)</td>
<td>2.4 (0.97-5.8)</td>
<td>2.7 (1.1-6.8)</td>
</tr>
<tr>
<td>Ireland</td>
<td>4,553,341</td>
<td>21 (6.4-49)</td>
<td>10 (7-8)</td>
<td>5.7 (2-16)</td>
</tr>
<tr>
<td>Spain</td>
<td>46,184,219</td>
<td>16 (4.7-40)</td>
<td>2.3 (0.89-6.1)</td>
<td>3.2 (1.1-8.8)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>16,672,271</td>
<td>15 (4.4-37)</td>
<td>5.9 (1.9-17)</td>
<td>4.2 (1.3-12)</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>7,357,095</td>
<td>13 (3.5-33)</td>
<td>11 (3.4-32)</td>
<td>52 (15-150)</td>
</tr>
<tr>
<td>Romania</td>
<td>20,258,908</td>
<td>9.1 (2.4-26)</td>
<td>2.9 (0.95-9.2)</td>
<td>18 (4.7-58)</td>
</tr>
<tr>
<td>Belgium</td>
<td>10,978,721</td>
<td>7.2 (1.6-21)</td>
<td>4.2 (1.2-13)</td>
<td>4.3 (1.3-14)</td>
</tr>
<tr>
<td>Austria</td>
<td>8,437,769</td>
<td>4.6 (1.3-15)</td>
<td>3.5 (0.92-13)</td>
<td>4.3 (1-16)</td>
</tr>
<tr>
<td>Denmark</td>
<td>5,572,569</td>
<td>1.6 (0.39-7.4)</td>
<td>1.8 (0.58-8.6)</td>
<td>0.96 (0.29-3.4)</td>
</tr>
<tr>
<td>Greece</td>
<td>10,995,472</td>
<td>1.2 (0.31-6.1)</td>
<td>0.69 (0.21-2.5)</td>
<td>0.84 (0.27-3.3)</td>
</tr>
<tr>
<td>Sweden</td>
<td>9,462,933</td>
<td>1.1 (0.28-4.2)</td>
<td>0.72 (0.22-2.7)</td>
<td>0.67 (0.21-2.9)</td>
</tr>
<tr>
<td>Poland</td>
<td>38,063,160</td>
<td>1 (0.27-4.3)</td>
<td>0.18 (0.05-0.86)</td>
<td>0.73 (0.23-3.2)</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>10,454,079</td>
<td>0.7 (0.19-3.6)</td>
<td>0.43 (0.13-1.9)</td>
<td>0.9 (0.33-5.7)</td>
</tr>
<tr>
<td>Norway</td>
<td>4,951,461</td>
<td>0.6 (0.17-3.5)</td>
<td>0.79 (0.27-4)</td>
<td>0.57 (0.21-3.9)</td>
</tr>
<tr>
<td>Iceland</td>
<td>320,477</td>
<td>0.22 (0.072-2.7)</td>
<td>4.5 (1.56-56)</td>
<td>0.49 (0.19-4.6)</td>
</tr>
<tr>
<td>Lithuania</td>
<td>3,058,435</td>
<td>0.2 (0.066-2.7)</td>
<td>0.43 (0.16-3.5)</td>
<td>1.8 (0.71-32)</td>
</tr>
<tr>
<td>Finland</td>
<td>5,386,179</td>
<td>0.11 (0.04-1.4)</td>
<td>0.14 (0.053-1.6)</td>
<td>0.24 (0.1-2.7)</td>
</tr>
<tr>
<td>Latvia</td>
<td>2,079,421</td>
<td>0.067 (0.024-0.89)</td>
<td>0.21 (0.083-2.9)</td>
<td>0.54 (0.24-8.2)</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>522,506</td>
<td>0.045 (0.017-0.066)</td>
<td>0.57 (0.23-0.99)</td>
<td>0.44 (0.2-0.8)</td>
</tr>
<tr>
<td>Cyprus</td>
<td>1,117,153</td>
<td>0.041 (0.015-0.06)</td>
<td>0.24 (0.098-0.42)</td>
<td>0.48 (0.22-0.87)</td>
</tr>
<tr>
<td>Portugal</td>
<td>10,488,929</td>
<td>0.038 (0.014-0.055)</td>
<td>0.023 (0.0094-0.041)</td>
<td>0.038 (0.017-0.07)</td>
</tr>
<tr>
<td>Estonia</td>
<td>1,327,740</td>
<td>0.028 (0.01-0.041)</td>
<td>0.13 (0.055-0.24)</td>
<td>0.45 (0.21-0.84)</td>
</tr>
<tr>
<td>Malta</td>
<td>418,516</td>
<td>0.019 (0.0071-0.028)</td>
<td>0.29 (0.12-0.52)</td>
<td>0.47 (0.22-0.87)</td>
</tr>
<tr>
<td>Hungary</td>
<td>9,954,603</td>
<td>0.007 (0.003-0.01)</td>
<td>0.004 (0.002-0.008)</td>
<td>0.008 (0.004-0.015)</td>
</tr>
<tr>
<td>Slovakia</td>
<td>5,399,568</td>
<td>0.0004 (0.0002-0.0006)</td>
<td>0.0005 (0.0002-0.001)</td>
<td>0.016 (0.008-0.032)</td>
</tr>
</tbody>
</table>

C.4.2 Relative risk versus median across time by location

While the primary method already does well emphasizing the variation in risk over time, to examine how this risk fluctuation varies across locations we also estimated the relative risk using the median across time by location. The methods were the same as above except each location had medians calculated for each simulation separately. This provides a well-defined comparison of how each epidemic period impacted each location. As we see in Figure C.4, time variations in risk were fairly consistent across states.
Appendix C: Supplement to Biosecurity in a Globalized World: The Case of Measles

Figure C.4. Relative risk of measles importation monthly by state. Here the relative risk is calculated as the estimated risk during each month divided by the median estimated risk of importation across the study period for that state. This figure highlights the periods of elevated and decreased risk of measles importation.

C.5 Incidence and Travel Trends

We examine the trends in incidence and travel over time to understand the consistency of these data and to determine whether any anomalies existed between years or countries. Monthly travel volume fraction was found to be highly consistent across years, but monthly incidence fraction, while generally demonstrating the same trends, several years demonstrated substantial variation, particularly 2005, 2007, 2017 (Figures C.5, C.6).

The consistencies of travel fraction by month within Europe should allow us to extrapolate these analyses to years when we do not have travel data available, particularly if we know the annual volume of travel, or the general annual travel trend.

The variations in incidence monthly fraction make extrapolating our model to years without incidence particularly challenging, particularly as we have found incidence to be critical to the risk, from both estimated and reported importations. However, as we have found, risk is relative within the U.S. – while there might be less absolute risk one year, the ratio of risk between states is mostly constant.
Figure C.5. Monthly travel fraction from all of Europe to the United States, 2005-2015.

Figure C.6. Total monthly measles incidence in Europe, 2005-2015. (A) absolute monthly incidence is highly variable from one year to the next, governed by the occurrence of outbreaks. However, (B) monthly incidence fraction is somewhat consistent, as Europe typically experiences seasonality in measles incidence, though substantial variation does occur (i.e. 2005, 2007, 2013).
References

Intended to be blank.
CURRICULUM VITAE
SHAUN TRUELOVE

114 S. Ann St., Baltimore, Maryland 21231
Phone: (410) 622-2595
Email: shauntruelove@jhu.edu

EDUCATION

Doctor of Philosophy
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
Epidemiology
Track: Infectious Disease
Advisor: Justin Lessler

Master of Public Health
University of Wisconsin-Madison
Madison, WI
Concentration: Epidemiology
Advisor: Lori Diprete-Brown

Certificate in Global Health
University of Wisconsin-Madison
Madison, WI

Bachelor of Science
University of Wisconsin-Madison
Madison, WI
Major: Biology
Minor (equivalent): Spanish

HONORS AND AWARDS

2016 Student Assembly Merit-Based Award
_Johns Hopkins Bloomberg School of Public Health, Baltimore, MD_

2015 Health Resources and Services Administration (HRSA) Trainee Fellow
_Johns Hopkins Bloomberg School of Public Health, Baltimore, MD_

2015 The Dorothy and Arthur Samet Student Support Fund in Epidemiology
_Johns Hopkins Bloomberg School of Public Health, Baltimore, MD_

2012-2017 Epidemiology Department Scholarship
_Johns Hopkins Bloomberg School of Public Health, Baltimore, MD_

2014 & 2015 Summer Institute in Statistics and Modeling in Infectious Diseases Scholarship
_University of Washington, Seattle, WA_
PROFESSIONAL EXPERIENCE

Temporary Advisor
World Health Organization;
Switzerland, Geneva
▪ Assisting with influenza surveillance priorities, particularly in the Eastern Mediterranean Region.
▪ Conduct training of local health staff
▪ Perform assessments of potential surveillance sites and infrastructure
▪ Produce reports and recommendations for implementation and maintenance of surveillance

Graduate Research Assistant
Johns Hopkins School of Public Health & Walter Reed Army Institute of Research;
Baltimore, MD, USA
▪ Conducting research study in collaboration between JHSPH and WRAIR to better understand influenza vaccine long-term immunogenicity and protection
▪ Utilizing U.S. Department of Defense’s Defense Medical Surveillance System and Serum Repository
▪ Developed protocol, managing the study, and leading the analysis

Graduate Research Assistant
Cutaneous Nerve Laboratory, Johns Hopkins Medicine;
Baltimore, MD, USA
▪ Conduct research under Michael Polydefkis, MD to better understand the association between intraepidermal nerve fiber density (IENFD) and peripheral neuropathy and other neurologic diseases
▪ Developing statistical methods to better use IENFD to provide clinical diagnosis of peripheral neuropathy
▪ Investigating risk factors associated with decreased IENFD measurements as well as factors for elevated measurements among healthy individuals
▪ Investigating potential for use of IENFD as a measure for other disorders and diseases

Graduate Research Assistant
Infectious Disease Dynamics group, Johns Hopkins School of Public Health;
Baltimore, MD, USA
▪ Conduct research under Justin Lessler and Derek Cummings, focused on response to and understanding of disease dynamics, emergence, and control for infectious diseases including measles, rubella, influenza, MERS-CoV, Zika, Ebola, and pertussis
▪ Developing novel methods for understanding transmission and outbreak potential of vaccine-preventable diseases.
▪ Developing statistical methods and tools for estimating population susceptibility to measles
▪ Investigating use of correlate measures for understanding vaccination activity efficacy
▪ Simulation of long-term population risk for measles infection to inform vaccination policy for the WHO’s Special Advisory Group of Experts on Measles
Conducted analysis on influenza detection, transmission, and immunity in China as part of the FluScape project

**Regional Epidemiologist** 2010-2012

CDC Global Disease Detection-Egypt, U.S. Naval Medical Research Unit No. 3; Cairo, Egypt

- Established the Eastern Mediterranean Acute Respiratory Infection Surveillance (EMARIS) network with 9 countries
- Lead epidemiologist for Egypt’s influenza-like illness sentinel surveillance network, H5N1 surveillance in Egypt, and all collaborative projects in Iraq
- Conducted epidemiologic analysis for various reports and publications
- Conducted numerous trainings in multiple countries on surveillance and study protocols
- Collaborated regularly with the U.S. CDC, WHO-Geneva, and WHO regional offices
- Collaborated regularly with ministries of health, national laboratories, and hospitals throughout the Middle East
- Organized a regional surveillance network meeting for over 90 participants from 14 countries
- Submitted and maintained IRB status and documentation for multiple projects
- Regularly communicated with funding agencies (U.S. DoS, DoD, USAID)
- Wrote and received funding for several grants

**Epidemiologist, Communicable Diseases** 2009-2010

Wisconsin Dept. of Health Services, Division of Public Health, Bureau of Comm. Diseases; Madison, WI, USA

- Coordinated state-wide hospital-based surveillance for influenza during the 2009 H1N1 pandemic
- Coordinated and performed data management, analysis, and reporting on pandemic H1N1 influenza
- Maintained communication networks with hospitals and local health departments during and after the pandemic
- Assisted with multiple outbreak investigations for legionellosis, foodborne and enteric diseases, etc.
- Conducted matched case-control studies during outbreaks

**Public Health Research Assistant** 2009

Robert Wood Johnson Foundation / University of Wisconsin School of Medicine and Public; Madison, WI, USA

- Assisted with development and execution of several asthma-related public health/epidemiological studies
- Coordinated studies and managed participants for principle investigators
- Recruited participants
- Designed and administered surveys
- Performed literature reviews
- Collected and managed data
- Analyzed asthma use location and time data in relation to asthma exacerbation events
PROFESSIONAL ACTIVITIES

Society Membership

American Society for Tropical Medicine & Hygiene
Society for Epidemiologic Research

Advisory Panels, Working Groups, Program Development

Infectious Disease Dynamics Group, Johns Hopkins Bloomberg School of Public Health
Working group, WHO Special Advisory Group of Experts (SAGE) for Measles
Surveillance and Outbreak Response Team (SORT), Johns Hopkins Bloomberg School of Public Health
Epidemiology Student Organization, Johns Hopkins Bloomberg School of Public Health

EDITORIAL ACTIVIES

Peer Review Activities
Reviewer, American Journal of Epidemiology
Reviewer, PLOS One
Reviewer, BMC Infectious Diseases
Reviewer, PeerJ
Reviewer, Expert Review of Vaccines

PUBLICATIONS


Curriculum Vitae

SHAUN TRUELOVE, MPH

PART II

TEACHING AND MENTORING

Teaching Assistant
Johns Hopkins School of Public Health

2016 Concepts and Methods of Infectious Disease Epidemiology
2016 Spatial Analysis III: Spatial Statistics
2014-16 Infectious Disease Dynamics: Theoretical and Computational Approaches
2015 Introduction to R for Public Health Researchers
2014 Epidemiology Methods 3
2013 Epidemiology of Infectious Diseases
2013 Principles of Epidemiology
2012 Epidemiology of Infectious Diseases

Mentoring

2014-2015 Master and Doctoral students in Epidemiology at JHSPH
   1 doctoral and 1 master student
2013-2014 Master’s students in Epidemiology at JHSPH
   2 master students

ACADEMIC SERVICE
Johns Hopkins School of Public Health

2014-2015 President, Epidemiology Student Organization
2013-2014 Coordinator, Infectious Disease Journal Club, Department of Epidemiology
2012-2014 Athletics Chair, Epidemiology Student Organization
PRESENTATIONS

Scientific Meetings


2016  Impact of Spatial Clustering of Non-Vaccination on Outbreak Potential (poster). American Society of Tropical Medicine & Hygiene, Atlanta, GA.

2015  Vaccinating for Elimination: Are our vaccination goals too low? (poster). Epidemics 5: 5th International Conference on Infectious Disease Dynamics, Clearwater, FL.

2014  The impact of measles susceptibility clustering on outbreak potential (poster). Vaccine Day, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

2014  Measles risk following disruption of routine immunization services during the 2014 Ebola outbreak (poster). Vaccine Day, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

COMMUNITY SERVICE

2008  Volunteer intern, La Senda Verde Ecological Refuge, Coroico, Bolivia. Worked for 3 months with the animal refuge to develop an improved animal resource database, website, and marketing materials, as well as helping to run the refuge and care for the animals.

2007-2008 Volunteer, Asociacion Senior de Huanca, Cusco, Peru. Worked for 6 months with this program for poor/street children, assisting with schoolwork, learning activities, and teaching English (4 hours/day, Monday-Friday).

2007-2008 Volunteer intern, Centre de Salud-Belen Pampa (Health Center of Belen Pampa), Cusco, Peru. Worked for 7 months with this medical clinic, assisting medical staff with patient triage and care, and learning medical Spanish (4 hours/day, Monday-Friday).
ADDITIONAL INFORMATION

Brief description of research interests

Mr. Truelove is an infectious disease epidemiologist whose research on measles and other infectious diseases focuses on the development of methods and tools for practical assessment of outbreak potential and improved vaccine distribution. Through use of the concepts from infectious disease dynamics and mathematical modeling, he is currently developing methods to better understand the risks of vaccine-preventable disease outbreak among populations with varying levels of vaccination coverage, from low coverage among select African countries, to high coverage in the Americas. Through better methods and better understanding of how diseases might enter and transmit through populations, Mr. Truelove is working to assist the WHO, through the Special Advisory Group of Experts (SAGE) on measles, to reach measles and rubella elimination goals worldwide. In addition, Mr. Truelove has a long-standing interest in influenza and works with the Fluscape project to better understand cross-protection immunity between different strains of influenza. Finally, Mr. Truelove has a strong interest in peripheral neuropathy and other neurologic diseases because of affected family members. He is currently working to achieve earlier detection and identification of risk factors for these conditions through examination of cutaneous nerve densities.