Using Mathematical Models to Understand Causal Mechanisms Underlying Counterintuitive Epidemiological Data

by

Joshua Havumaki

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in The University of Michigan
2019

Doctoral Committee:

Associate Professor Marisa C. Eisenberg, Chair
Associate Professor Veronica J. Berrocal
Professor Joseph N. S. Eisenberg
Associate Professor Rafael Meza
Assistant Professor Jon L. Zelner

Joshua Havumaki

joshsh@umich.edu

0000-0002-6749-1217

For my father, Bruce Havumaki.

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Marisa Eisenberg, for her support over the past three years as well as during my MPH degree (along with Rafael Meza) from 2011 to 2013. I am grateful for the fact that she has allowed me to make this dissertation my own, and to focus on projects that I find particularly interesting and relevant for my career goals. She has provided me with tools that are fundamental for successful research. Additionally, she has continuously offered invaluable help solving technical and non-technical problems in an efficient way, which has allowed me to both learn new things and quickly clear roadblocks that inevitably come up during research projects.

Next, I would like to thank my committee members. Thanks to Joseph Eisenberg who has provided me with useful feedback on the framing and public health impacts of my projects. This has improved my writing, and more importantly, has allowed me to keep perspective on the underlying goals of my research. Thanks to Jon Zelner who has helped me expand my set of technical skills in a way that I wouldn't have been able to do on my own. I believe that this has prepared me well for the next steps in my research career. Thanks to Veronica Berrocal and Rafael Meza for providing me with their time and very valuable feedback throughout each stage of this process.

I would also like to acknowledge a few other individuals. Thanks to Andrew

Brouwer for always being willing to talk through technical problems, and for providing very useful feedback on my dissertation talk. Thanks to Nancy Fleischer for her helpful advice on DAGs, and the very useful material she taught in EPID 824. I have referred to class notes from 824 many times, and I'm sure I will continue to do so in the future. And thanks to Michael Hayashi for our yearly chats at the MIDAS meeting, and for setting up and maintaining the *muppetlabs* server.

Finally, I'd like to thank my wife, Alice Pastorino, who has given my life never ending support, balance, and stability. She has taught me how to finish projects that I start. She encouraged me to get a PhD in the first place. And I am certain that successful completion of my PhD would not have been possible without her.

My work was funded by National Institute of General Medical Sciences Grant U01GM110712, and The Joint Initiative for Vaccine Economics, Phase 5, a cooperative agreement between the University of Michigan and the Centers for Disease Control and Prevention (U01IP000965).

PREFACE

Chapter 2 is in the process of being submitted for publication with Marisa C. Eisenberg. A version of chapter 3 will be submitted for publication with the following co-authors: Joel C. Miller, Ted Cohen, Chengwei Zhai, Marisa C. Eisenberg, and Jon L. Zelner. Chapter 4 is under CDC clearance review pending submission with the following co-authors: Joseph NS Eisenberg, Claire P. Mattison, Ben A. Lopman, Ismael R. Ortega-Sanchez, Aron J. Hall, David W. Hutton, and Marisa C. Eisenberg are all co-authors.

TABLE OF CONTENTS

CHAPTER

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

Analyzing epidemiological data (e.g., from observational studies or surveillance) can reveal results contrary to what might be expected given *a priori* knowledge about the study question. In these cases, a clear mechanistic understanding of why counterintuitive results are observed is critical to minimize bias in study designs and implement effective interventions targeting diseases. Mathematical modeling approaches provide a flexible way to connect mechanisms with real-world data. In this dissertation, we describe the use of mathematical models to explore 3 cases in which seemingly counterintuitive results have been observed. First, we examined the obesity paradox or the apparent protective effect of obesity on mortality among certain high-risk groups, e.g. diabetic ever-smokers. Second, we examined how to leverage spatial and contact heterogeneity to optimize tuberculosis screening interventions in a variety of settings including those with high incidence-levels where household-based interventions have unexpectedly limited population-level effects. Finally, we examined why norovirus outbreaks are explosive in nature, but result in relatively low attack rates (the percentage of individuals who become diseased) in school and daycare settings.

In Aim 1, we developed a method to simulate epidemiological studies using compartmental models (CMs) derived from directed acyclic graphs (DAGs). We illustrated our approach using the obesity paradox as a case study. Specifically, we examined how altering underlying causal mechanisms (i.e. CM structure), can cause spurious associations in the data. We found that incorporating study design bias

(e.g., including covariates in the causal mechanism and not adjusting for them), can lead to the obesity paradox. Overall, we showed how mathematical modeling of DAGs can be used to inform analyses, and explore underlying biases which may be helpful for designing sound observational studies and obtaining accurate measures of effect.

In Aim 2, we explored how variation in community contact and endemic incidence levels can affect the impact of household or community-targeted screening interventions using an individually-based network model. Overall, we found that the community drives transmission in high incidence settings. In general, more protection was conferred by targeted interventions and in lower incidence settings within networks that had fewer numbers of contacts, or shorter distance between contacts. Ultimately, these results may help identify the settings in which household or community targeted screening interventions will be effective.

In Aim 3, we explored mechanisms that underlie norovirus outbreak dynamics using a disease transmission model. We compared different scenarios, including a partially immune population, stochastic extinction, and individual exclusion, and calibrated our model to daycare and school outbreaks from surveillance data. We found that incorporating both a partially immune population and individual exclusion was sufficient to recreate explosive norovirus dynamics, more realistic outbreak durations (compared with immunity alone), and relatively low attack rates in school and daycare venues.

Ultimately, epidemiological findings only appear counterintuitive when there is a lack of understanding about the underlying mechanisms leading to what is observed in data. This dissertation highlights the importance of resolving this lack of understanding, and the use of models as a tool in this process. We used mathe-

matical models as *in silico* laboratories to compare competing causal mechanisms, understand transmission patterns across different settings, and reveal key features of the natural history of disease. Gaining insight into causal mechanisms underlying seemingly counterintuitive data is critical to be able to minimize bias in study designs and implement effective disease targeting interventions.

CHAPTER I

Introduction

Analyses of epidemiological data may reveal findings that are contrary to what one would expect given *a priori* knowledge about risk factors for disease. In these cases, a mechanistic understanding of why counterintuitive results are observed is necessary to minimize bias in study designs and implement effective interventions targeting the disease. Mathematical modeling approaches can provide a flexible framework that can be used to link mechanism with data. An example of counterintuitive study results often discussed in the chronic disease literature are survival paradoxes (or so-called 'reverse-epidemiology') wherein individuals in what are usually considered higher risk groups actually have lower mortality rates compared with exchangeable (with respect to other covariates) [6] individuals in lower risk groups [7]. For instance, the 'obesity paradox' was observed in a study conducted among a diabetic study population in which obese ever-smokers were found to have a lower risk of mortality than normal weight ever-smokers [2]. Given what is known about the relationship between obesity and mortality, it would be expected that obese individuals would always have higher mortality rates than their normal-weight counterparts [7].

It is generally thought that survival paradoxes arise due to biases in study design.

These may be mechanistic in origin, for instance, a statistical analysis plan may be poorly designed due to incomplete understanding of the underlying disease process. In this example of the obesity paradox, potential explanations to the observed protective effect of obesity include selection bias due to the study being conducted only among diseased (diabetic) individuals, and reverse causation due to factors like smoking related comorbidities [2]. Overall, an explicit mechanistic understanding of how these (and other) study design biases might interact to cause the obesity paradox could help minimize bias in the design of future studies on the topic (Aim 1).

In the infectious disease literature, multiple modes of transmission for a given disease can interact [8] and lead to findings that are difficult to interpret. For instance, despite the fact that the screening of household contacts (household contact tracing) for tuberculosis (TB) has been widely adapted and is considered effective in the low-burden settings [9, 10], results from high-burden settings have been mixed [11–14]. In high burden settings, the majority of transmission occurs outside the household and between not known contacts [15, 16]. However, sharing a household with an individual who has active TB is still a risk factor for infection [17, 18]. The spatially heterogeneous distribution of TB [19, 20] further complicates the design of effective interventions, because different settings might benefit from different types of screening programs (e.g. household contact tracing compared with neighborhood tracing). In light of these complicated multilevel transmission patterns, design of screening interventions should be carefully considered in a mechanistic manner and in the context of a spatially explicit setting (Aim 2).

Finally, yet another unexpected finding from surveillance data is the low norovirus attack rates in school and daycare settings. Norovirus outbreaks are considered explosive in nature [21] due to a rapid onset of cases and dramatic symptomology among diseased individuals [22]. Additionally, children have the highest incidence rates [23–25] and are thought to be drivers of transmission in the community [5]. These rapid outbreak growth rates and the role that children play in community transmission, would lead one to expect that outbreaks in schools and daycares would have high attack rates and exhaustion of susceptibles. However, attack rates are relatively low in these settings ($\sim$15% and $\sim$20% in daycares and schools, respectively – data from the National Outbreak Reporting System (NORS) [26]). In general, a clear mechanistic understanding of why explosive norovirus outbreaks in daycares and schools do not lead to the majority of children becoming infected can provide insight into the biological and epidemiological drivers of these transmission patterns and ultimately, inform the design of effective interventions (Aim 3).

Overall, mathematical models are a useful tool which can enable a greater understanding of counterintuitive findings in epidemiology. Specifically, formulating a model requires that causal mechanisms are explicitly defined. Next, simulating the model enables the researcher to gain insight into what factors (or combination of factors) can explain the patterns seen in the data e.g., two modes of transmission might counteract each other. Furthermore, models can integrate risk factors on multiple levels contributing to the mechanism in question e.g., endogenous factors like immunity along with population level factors like contact rates. Models may therefore be used as *in silico* laboratories to explicitly encode alternative causal mechanisms and to ask counterfactual questions. Ultimately, epidemiological findings only appear counterintuitive when there is a lack of understanding or clarity about the underlying mechanisms which lead to what is observed in the data. This dissertation uses models to resolve this lack of understanding and to gain insight into the transmission patterns and natural history of diseases.

1.1 Aim 1: Mathematical modeling of directed acyclic graphs to explore competing causal mechanisms underlying epidemiological study data

1.1.1 Directed Acyclic Graphs

Directed Acyclic Graphs (DAGs) are networks which in epidemiology are used to graphically depict causal relationships among variables of interest in a study (framed as Bayes nets in this setting [27]). DAGs encode conditional dependencies and are used to identify which variables should be measured and adjusted for (i.e., conditioned upon) to obtain an unbiased effect estimate of an exposure on a disease [28].

In general, apparent statistical associations may arise through any of the following [29]:

- Chance due to random variation
- Causal effect e.g. ‘a’ causes ‘b’ or ‘b’ causes ‘a’
- ‘a’ and ‘b’ share a common cause that was not conditioned on (i.e., confounding bias)
- ‘a’ and ‘b’ share a common effect that was conditioned on (i.e., collider bias or selection bias)

DAGs can be used to identify all of the above associations except for random variation. Variables on a DAG include the exposure, outcome, and covariates, connected by *directed* edges which indicate causal effects. The absence of an edge between variables implies a lack of a causal relationship between those variables [29]. A variable ‘a’ is a direct cause of ‘b’ if there is an arrow from ‘a’ to ‘b’. On the other hand, ‘a’ is an indirect cause of ‘b’ if there is only one connection between ‘a’ and ‘b’, and that connection is mediated by another variable [29]. Any sequence of arrows connecting two variables on a DAG (regardless of direction) is a ‘path’. A sequence following the direction of arrows is a ‘directed path’ while a path which starts with an arrow pointing into the exposure and ends with an arrow pointing to the outcome is a ‘backdoor path’ [30]. A DAG is *acyclic* in that no directed path forms a closed (feedback) loop. To be considered ‘causal’, all common causes of any pair of variables must also be included on the DAG [31].

In general, the direction of arrows and relationships between variables on DAGs can be used to identify study design biases in a straightforward manner. First in general, confounding bias occurs when a pair of variables share a *common cause* that is not adjusted for in the analysis (the common cause may be direct or indirect). Second in general, collider bias (i.e., selection bias) occurs when two variables share a *common effect* upon which itself or its descendants are adjusted for in the analysis. After a DAG is drawn, d-separation criteria are used to define statistical dependencies between variables. These criteria depend on three key assumptions [29]:

- The Causal Markov Assumption: Any variable ‘a’ is independent of ‘b’ conditional on the direct causes of ‘a’ unless ‘b’ is an effect of ‘a’.

- Faithfulness: If ‘a’ effects ‘b’ through two pathways (one negative and one positive), the pathways must not completely cancel each other out. Faithlessness leads to statistical independence between variables even when the causal structure of a DAG implies dependence.

- No Random Error: Large enough sample size such that statistical associations are not due to random chance

On a DAG, two variables that are not causally related (i.e., no directed path between them) are statistically independent (i.e., no association between them) if every backdoor path between them is blocked. Typically to obtain an unbiased effect estimate, the goal is to adjust for the minimally sufficient set of covariates that blocks all backdoor paths and controls for confounding [30]. The workflow for determining the set of covariates that must be conditioned upon to achieve statistical independence and adjust for confounding is described elsewhere [30]. In general, a path is blocked when there is either (1) a collider (and its descendants) in the path between the two variables that *is not* conditioned upon, or (2) a non-collider in the path between the two variables that *is* conditioned upon. Again, once all backdoor paths between exposure and outcome are blocked, an unbiased effect estimate can be calculated.

Because DAGs require that assumptions of the causal relationships between key variables are made explicit in addition to identifying sources of bias, they may also reveal variables that need to be measured or ambiguous relationships between variables that should be investigated further. Ideally, DAGs are constructed before a study to aid in identifying which variables must be measured and adjusted for in the analysis and to help design an appropriate analysis plan.

Overall, although DAGs summarize the complete set of known relationships between variables relevant to a given study question [28, 30, 32], they are not well suited for comparing competing causal mechanisms because they are nonparameterized, require that ambiguous variable relationships are simplified and are reductionist in that they typically only examine a single causal link (i.e., between exposure and outcome) within a chain of events. Various sensitivity analyses can be conducted (e.g., to see the effects of unmeasured confounders), but exploring multiple effects at once can be difficult to interpret. Thus, DAGs are useful tools for minimizing bias and measuring associations, but do not synthesize an entire mechanistic chain of events.

1.1.2 Compartmental Models

Compartmental models (CMs) are parametric causal frameworks which are typically formulated as ordinary differential equations (ODEs) and used to simulate flows between states on the population level over time [32, 33]. There is a long history of using compartmental models to simulate disease transmission and progression in infectious diseases [34] and disease progression in chronic diseases [35]. For a detailed example of a CM representing infectious disease transmission with ODEs written out, see Aim 3 Section 1.3.1. Once a model is defined, it can be fit to epidemiological data (e.g. surveillance data, weekly incidence counts, or serology data) using a likelihood to match the model results with the data. Identifiability analysis may be used to determine which parameter values can be estimated based on the data and structure of the model. Finally, the fitted model can be used to obtain values or distributions for unmeasured parameters (assuming they are identifiable), and to examine counterfactual scenarios like the effects of interventions on the incidence or progression of the disease. Furthermore, sensitivity analyses can be conducted to see how results change when exploring uncertainty in the model (i.e., by varying parameter values), and to examine competing causal mechanisms (i.e., by changing the model structure).

Contrary to DAGs, mathematical models are not meant to simulate all causes underlying a given mechanism. Rather, only the key mechanism(s) directly related to the study question should be incorporated. In this way, CMs seek to balance parsimony [36] with realism. Furthermore, by operationalizing a sequence of events, mathematical models can synthesize *a priori* knowledge to directly simulate an entire chain of events (as opposed to DAGs which tend to focus on a single cause and effect). For instance, infectious models typically simulate individuals the natural history of disease e.g., progressing from susceptible to infectious to immune (see Section 1.3.1) whereas a DAG is only meant to provide insight into a single association like maternal smoking and infant mortality (see Figure 1.1). The parameter estimation process and resulting model fit can be informed by many different types of data (e.g., meta-analyses, laboratory studies). Additionally, other scenarios can be explicitly encoded into the model when data collection is untenable due to ethical constraints, limited resources, or is logistically impossible (e.g. counterfactual scenarios). Limitations of compartmental models should be considered during the model development stage and when interpreting model results. For instance, compartmental models typically assume homogeneous (mass action) mixing which may not be realistic and can affect the disease dynamics [37, 38]. This can be addressed by adding different demographic groups mixing at different rates (i.e., by incorporating different compartments), but can quickly lead to a combinatorial explosion of equations. Alternatively, a network model can be used to explicitly simulate contact heterogeneity [38] (see Section 1.2.1 for more details). Another limitation is the fact that parameterizing models requires assumptions about the values of transition rates that are often times not known empirically e.g., the duration of the period from exposure to the disease to actually presenting symptoms [39]. However to account for this uncertainty, large scale sensitivity analyses can be conducted to understand how the model results change in response to changing parameter values (e.g., by conducting Latin hypercube sampling (LHS)) [40]. Another limitation is that deterministic models assume large population sizes which is not realistic in certain settings e.g. daycare center [37]. To address this, a stochastic version of the CM can be created to simulate small populations (see Section 1.3.2 for more details). Another limitation is exponentially distributed transition rates which can result in some proportion of the population transitioning to a subsequent state immediately upon entering a new state. This is not realistic, but on the population level, the model simulates the accurate average transition time. Additionally, substates can be added to approximate a more realistic distributions of transition times e.g., a gamma distribution.

Ultimately, CMs are a flexible tool which can be used synthesize various types of data and simulate an entire causal mechanism. Detailed counterfactual questions can be explicitly coded into the model and examined. Importantly, assumptions of the model including what is and is not being simulated should be carefully considered while designing the model and interpreting its results.

1.1.3 Survival Paradoxes in Epidemiology

Study design biases that are not properly adjusted for and lack of understanding about predominant causal mechanisms can lead to counterintuitive results in epidemiological studies. For instance, the low birth-weight (LBW) paradox has shown that LBW children have a lower infant mortality rate if they are born to smoking mothers compared to those born to non-smoking mothers [41]. One would expect that those born to smoking mothers would have worse health outcomes given the risks associated with maternal smoking [42]. Although survival paradoxes are referred to as ‘paradoxical’, it does not mean that we lack understanding about what their causes are. Rather, the explanations underlying them might be complex and multifaceted. Furthermore, different explanations might be equally valid in different study settings.

Although it has been proposed that this and similar paradoxes (see below) are physiologically true [43, 44], the general consensus in the literature tends to be that smoking always leads to worse health outcomes and therefore, studies typically explain the paradox by revealing biases in their design (potentially resulting from incomplete understanding of the mechanisms/causal processes at play). DAGs can provide a useful framework for explaining how underlying bias in study design can cause the paradoxical findings of LBW paradox. For instance, conditioning on LBW (a mediator in the pathway between maternal smoking and infant mortality) can cause collider bias if there is an unmeasured confounder (e.g. birth defects) between LBW and infant mortality. This would change the direction of the association between maternal smoking and infant mortality because LBW is a collider due to the unmeasured confounder [1, 41] and conditioning on a collider creates a spurious negative association between its causes (i.e., between maternal smoking and infant mortality). See Figure 1.1 for an illustration using a DAG.

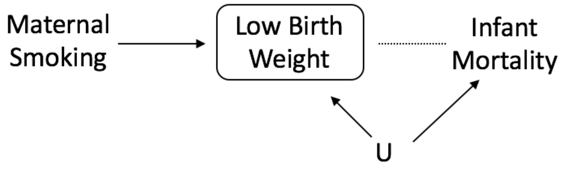


Figure 1.1: From [1], the exposure, 'Maternal smoking' does not have a causal association with the outcome, 'Infant mortality'. However, the existence of an unmeasured risk factor 'U', turns 'low birth weight' into a collider therefore creating a spurious negative association between 'low birth weight' and 'mortality' and therefore between 'Maternal smoking' and 'Infant mortality'

In addition to the LBW paradox, numerous other epidemiological paradoxes exist, such as the obesity paradox, i.e. the finding that obese ever-smokers with certain diseases such as diabetes have lower mortality rates than their normal weight counterparts [2]. One would expect that obese individuals would have worse outcomes than normal weight individuals [7]. Ultimately, a wide range of physiological and statistical explanations exist for these survival paradoxes and statistical methods are limited in their ability to flexibly assess and compare different theories. DAGs are qualitative and even if they show the presence of a certain bias, they do not provide any information on its quantitative effects. Although the magnitude and direction of potential biases can be explored in sensitivity analyses, examining multiple biases at the same time in a rigorous manner can become complex and intractable.

1.1.4 Aim 1: Overview

In Aim 1 we developed a workflow to use a previously published mapping between CMs and DAGs [32] to simulate epidemiological studies. We applied used this workflow to evaluate how different unmeasured biological mechanisms can potentially generate bias leading to the obesity paradox.

1.2 Aim 2: Exploring the impact of variation in spatial patterns of community contact on the effectiveness of household- vs. community-based screening interventions for Tuberculosis

1.2.1 Network Models in Epidemiology

In spite of the flexibility and relative simplicity of using CMs to incorporate different causal mechanisms, simplifying assumptions can limit their accuracy. For example, in many settings, homogeneous mixing is not a sufficient representation of reality and can alter the course of the dynamics of disease transmission substantially [38]. An individual's contact network tends to be determined by factors like age, sex, race, social structure, spatial structure, and behavior [38, 45]. Homophily or the tendency for people to be attracted to individuals similar to themselves is thought to be the primary driving force behind social networks [45]. Therefore, for diseases that are transmitted more homogeneously across the population (e.g. a local measles outbreak), the homogeneous mixing assumption might be reasonable. On the other hand for diseases where social contact heterogeneity or community structures are more critical (e.g. HIV), incorporating more complex mixing patterns might be important to accurately represent dynamics [38]. One potential solution for relaxing the assumption of homogeneous mixing is to add additional structure into a CM to account for different contact rates between groups (e.g., a set of compartments for high mixing groups and a set of compartments for low mixing groups) however, this can quickly lead to a combinatorial explosion of compartments if there are multiple groups. An alternative solution that is more scalable is to use a network model to explicitly define person-to-person contact patterns. Homogeneous mixing in a CM is equivalent to a fully connected network (i.e., where each individual is connected to all other individuals) [46]. Modeling disease spread over networks dates back to the 1980s with the first paper simulating acquired immune deficiency syndrome transmission [47].

Network models are made up of nodes (e.g. individuals) and edges which are the connections between nodes [48]. The average number of contacts per individual (i.e., average (expected) degree), the average connection length, and the average path length between 2 randomly selected nodes are important parameters which are typically used to characterize networks. There are numerous different types of networks on which infectious disease transmission has been modeled. For example, random networks(also called Erdös-Rényi networks) are distinct from many other kinds in that spatial position does not affect the probability of forming connections [49]. Additionally, lattices are networks in which adjacent individuals placed on a grid are connected. Lattices have high clustering (local connections) and long path lengths [50]. Small world networks are characterized by high clustering and short path lengths [51]. Scale-free or power law scaling networks incorporate heterogeneity in average degrees (e.g. to represent super-spreaders) [48, 52, 53]. Additionally, exponential random graphs models are networks in which the probability of a connection between nodes is independent of any other connection on the network [54]. Finally, spatial networks [55] are a broad class of networks in which nodes are distributed in a defined area and connections are drawn based on spatial structure, e.g. using an explicitly drawn by a connectivity kernel. The parameters that inform the connectivity kernel typically include the distance between nodes, the average degree, and the the spatial distance or standard deviation which defines the average connection length between contacts [3, 55]. These parameters can be tuned to create different network characteristics e.g., primarily long range connections vs. primarily short range connections.

Overall, network models are more computationally expensive than standard CMs, but if the spatial distribution or contact patterns of a population is of interest, they are a useful tool for simulating disease transmission.

1.2.2 Transmission of TB

On the population-level incident cases of TB tends to cluster in high burden hotspots indicating that transmission occurs in a spatially heterogeneous manner e.g., [20].

TB is transmitted through close and causal contact in a variety of contexts. Specifically, close contacts are those that are frequent and prolonged, e.g within households, while casual contacts are shorter and likely to occur with more people, e.g. gatherings among community members. Both contribute to the spread of TB [17, 18, 56]. Close contacts can concentrate infection within a tightly clustered unit, and casual contacts seed transmission in other clusters. Unexpectedly, numerous studies have found that in high incidence regions, most TB transmission occurs outside the household and between not-known social contacts e.g., [15, 16, 57]. However, sharing a house is still a risk factor for TB transmission [17, 18].

1.2.3 Latent TB

Individuals who become infected with TB initially enter a period of latency (latent TB infection (LTBI)) in which they are asymptomatic. It has traditionally been estimated that up to one-third of the global population has LTBI [58]. However, a recent analysis modeled the trends of the annual risk of TB infection on the country level and estimated the human LTBI reservoir to be closer to one-quarter of the global population with wide age and regional variation [59]. Individuals who have LTBI may progress to TB months, years or even decades [60] after the initial infection, but the vast majority $\sim 90\%$ of immuno-competent individuals never progress to TB [61]. A recent review examining studies before and after the introduction of chemotherapy found that the majority of LTBI cases that progress to TB do so within 2 years of infection [62].

Among those with LTBI who do not progress to TB, some individuals may clear the infection without any treatment (though studies are sparse) [59]. Alternatively, if detected e.g. through contact tracing (see below) individuals with LTBI may be treated with isoniazid preventative therapy (IPT) which inhibits the synthesis of mycolic acids essential for the cell wall of *mycobacterium tuberculosis* [63]. Importantly, impacts of trials that have administered IPT to individuals at high risk for TB have shown limited long term population-level impacts [11, 13].

Overall, the large human reservoir of LTBI is asymptomatic and therefore not easily detected without active case finding programs (see below). To accelerate reductions in TB transmission by preventing new cases of TB, it is essential to find cost-effective ways to target and treat these individuals.

1.2.4 TB Screening Methods

The standard approach to case ascertainment and treatment for individuals with active TB since the 1990s has been direct observed treatment – short course (DOTS) which involves passive surveillance, followed by chemotherapy administered under direct observation [64]. Although these current measures have been successful in reducing global TB incidence by $\sim 2\%$ annually, it has been suggested that more aggressive case ascertainment measures are needed to find recently infected individuals and including those who have LTBI and are asymptomatic.

When an individual is found to have active TB through passive surveillance (i.e., they go to clinic with symptoms and get diagnosed), their household and community contacts may be subsequently screened for TB (contact tracing). This systematic investigation of contacts can help rapidly find new cases of TB or LTBI. In addition, contact tracing can identify targets for IPT which can in turn prevent a first episode of TB or LTBI reactivation [65–67]. Although contact tracing has been conducted in high-income, low-incidence countries for decades [9, 10], resource poor settings have traditionally relied on DOTS. In recent years, there has been increasing support for a contact tracing based approach in low- and middle-income settings. In fact, the World Health Organization recommends that contact tracing of household and close contacts be conducted in low- and middle-income countries for individuals with smear positive pulmonary TB, multi-drug or extensively drug resistant TB, HIV and for children < 5 [65]. However, lack of resources prevents countries from widespread adoption of contact tracing [68].

One meta-analysis found that in low- and middle-income settings, the majority of TB transmission occurs between contacts in the first year after the index case becomes exposed [69]. A randomized control trial conducted in Zambia and South Africa used household contact tracing (HHCT) i.e. screened the household contacts of index cases, in an effort to reduce the prevalence of TB in HIV-endemic communities among 1.2 million individuals. Although a reduction in both TB incidence and prevalence did occur, the results were not statistically significant [11]. Another randomized control trial conducted in Vietnam conducted HHCT among 10,964 individuals with pulmonary TB and found that household contact tracing plus the standard passive case finding was significantly more effective than passive case finding alone [12]. Other trials have yielded no significant population-level effects of contact tracing interventions [13, 14]. Finally, a mathematical modeling analysis showed that HHCT with the provision of IPT had a modest effect on population-level TB risk [70].

Overall, if sufficient resources are available, contact tracing shows promise as a method for rapidly finding new cases of TB or LTBI, but the exact interventions and protocols need to be clearly considered and defined to be cost-effective.

1.2.5 Aim 2: Overview

In Aim 2, we developed an individually-based network model with household and community contacts to examine how variation in spatial and contact heterogeneity impacts the effectiveness of different household and community targeted screening programs.

1.3 Aim 3: Examining the discrepancy between explosive norovirus outbreaks and relatively low attack rates in day-care and school settings

1.3.1 Infectious Disease Transmission Model

CMs are commonly used to simulate infectious disease transmission and progression. A particular infectious disease transmission model that has gained popularity in recent years is the susceptible-infectious-water-recovered (SIWR) model (an extension of the classical susceptible-infectious-recovered (SIR) model [34], and a variation on the Environmental Infection Transmission System modeling framework [71]), which is used to simulate disease transmission on the population level in the context of its surrounding environment. Numerous other adaptions of SIR models exist such as the susceptible-infectious-susceptible (SIS), susceptible-exposed-infectious-recovered (SEIR), and maternal derived immunity-susceptible-infectious-recovered (MSIR) models [72].

The SIWR model is used to simulate disease transmission both directly from person-to-person and indirectly through environmentally-mediated pathways [73]. A

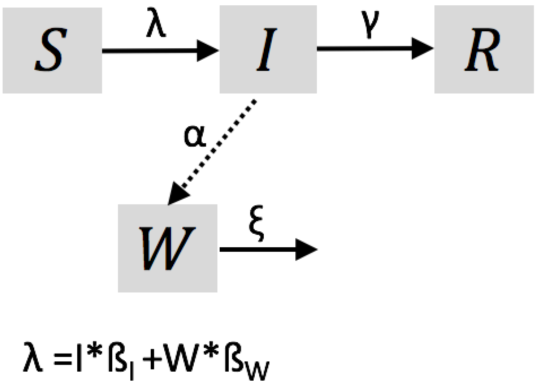


Figure 1.2: The SIWR model of environmentally transmitted disease. The proportion of the population that is susceptible, infectious, and recovered are in the S, I, and R compartments, respectively. The pathogen concentration in the water is tracked in the W compartment. The dotted line indicate shedding of pathogen. Finally λ, the force of infection is calculated based on the person-to-person (β_I) and the water-to-person (β_W) transmission rates and the number of infectious individuals and the amount of contamination in the water.

schematic of the model is shown in Figure 1.2. ODEs corresponding to the SIWR model are as follows:

SIWR Model Equations

$$\begin{aligned} \dot{S} &= -\beta_W W S - \beta_I S I, \\ \dot{I} &= \beta_W W S + \beta_I S I - \gamma I, \\ \dot{W} &= \alpha I - \xi W, \\ \dot{R} &= \gamma I, \end{aligned} \tag{1.1}$$

where the proportion of the population that is susceptible, infectious, and recovered are in the S, I, and R compartments, respectively. The pathogen concentration in the water is tracked in the W compartment. Next, β_W and β_I are the water-to-person and person-to-person transmission rates, γ is the recovery rate for infectious individuals, α is the shedding rate of pathogen from infectious individuals into the water, and finally ξ is the pathogen decay rate in the water. The water compartment can be used to track other sources of environmental contamination such as numbers of pathogens on fomites.

Overall, the SIWR framework provides an example of how mathematical models can be used to represent causal mechanisms across multiple levels that drive underlying disease transmission.

1.3.2 Stochastic Formulation

In general, small populations are more likely to become extinct due to lack of genetic diversity (which leads to more vulnerability) with the mean time to extinction increasing exponentially with population size [74, 75]. Analogously, infectious disease outbreaks occurring in small populations have a lower mean time to extinction than outbreaks occurring in larger populations [74]. As mentioned previously, CMs rely on the assumption that the model population size is large (see section 1.1.2), therefore to model outbreaks in small populations a stochastic formulation of the model can be implemented to account for the effect of small population size on dynamics. There are numerous ways to implement stochasticity into a disease transmission model including stochastic differential equations which contain a random noise variable [76], exact continuous-time Markov chain model with the Gillespie algorithm [77], and a more computationally efficient approximation of the Gillespie algorithm using τ-leaping [78]. The τ-leaping formation is an Euler approximation which performs all reactions over a fixed (user-defined) time interval (τ) before updating the population distribution across disease states.

Overall, stochastic extinction and variation are important to account for when simulating disease outbreaks in settings with small populations, e.g. daycare centers.

1.3.3 Epidemiology of Outbreaks

Norovirus is the leading cause of diarrhea requiring medical care in the United States (US) [79]. Among children < 5, the incidence of norovirus is approximately 3 times higher compared with other age groups in the community with adults > 65 having the second highest incidence rates [23, 80]. Overall, the highest risk populations for severe norovirus are children, elderly and the immuno-compromised with children often being identified as the primary drivers of transmission [22, 81].

A systematic review examined published norovirus outbreaks between 1993 and 2011, and found that outbreaks occurring in food service settings were the most common (35%) while those occurring in school and daycare settings were unexpectedly the least common (10%) [82].

Norovirus outbreaks originate from a variety of sources including food [83], water [84], and fomites [85], with person-to-person transmission propagating the disease across settings and age-groups [39, 86]. Outbreaks are typically described as being explosive in nature with rapid onset of cases due to high infectivity of norovirus virions [87], dramatic symptomology [22],long periods of post-symptomatic shedding [88], limited immunity conferred by natural infection (see below) [89, 90], and extended environmental persistence [91].

Despite the explosive nature of norovirus, and the fact that children drive transmission, attack rates in school and daycare settings are low ($\sim$15% and $\sim$20% in daycares and schools, respectively – data from NORS [26]). Overall, understanding why attack rates are low in these settings, requires an understand of the mechanisms driving transmission.

1.3.4 Molecular Biology and Immunity

Noroviruses are a genetically diverse group of single stranded RNA viruses with 29 different genotypes (strains) able to infect humans [92, 93]. This extensive diversity makes it difficult to develop broadly protective immunity either through natural infection or vaccination [94]. The most dominant strain is GII.4 which rapidly evolves to create novel viruses every 2-4 years. This genetic drift is analogous to influenza viruses [95] and has been successful in evading immunity within humans [96]. After natural infection there is limited cross-strain protection and a short-lived duration of immunity [79, 97]. Empirical research from human challenge studies has generally shown that immunity can last up to 2 years [98]. Although sero-prevalence studies have found high percentages of the children with norovirus antibody titers [99, 100], there is no established correlate of protection to determine norovirus immunity.

Individuals can have innate resistance to norovirus depending on their genetics. A functional FUT2 gene confers greater susceptibility to certain viruses (e.g. HIV, Rhinovirus, and Influenza) including noroviruses [95, 101, 101, 102]. The FUT2 gene enables the synthesis of certain types of histo-blood group antigens on the gut epithelium which are receptors for noroviruses and contribute to resistance to certain strains [95, 101, 101]. Additionally, individuals with A or O blood groups are also thought to be more susceptible to norovirus (compared with AB or B) due to better binding of norovirus to the saliva of those individuals, but this varies by strain as well [103–105].

1.3.5 Aim 3: Overview

In Aim 3, We developed an transmission model and calibrated it to CDC surveillance data from NORS. We incorporated different model features including population immunity, stochasticity, and exclusion of diseased individuals to understand what mechanisms can effectively recreate explosive outbreaks with low attack rates in daycare and school settings.

CHAPTER II

Mathematical modeling of directed acyclic graphs to explore competing causal mechanisms underlying epidemiological study data

2.1 Abstract

Directed acyclic graphs (DAGs) are used in epidemiological studies to understand causal processes and determine appropriate analytical approaches for a given exposure and outcome. Compartmental models (CMs) depict flows between disease states on the population level and can also be used to represent different causal mechanisms. In this chapter, we use and extend a mapping between DAGs and CMs to show how DAG–derived CMs can be used to compare competing causal mechanisms by simulating epidemiological studies. Specifically, by restructuring our DAG and corresponding CM to represent competing hypotheses, we can see how robust our simulated epidemiological study results are to different biases in study design and underlying causal mechanisms. As a case study, we simulated the obesity paradox: the apparent protective effect of obesity on mortality among diabetic ever-smokers, but not among diabetic never-smokers. Given that we would expect obesity to confer poor health outcomes among all groups, especially those at higher

risk of mortality (e.g., ever-smokers), this paradox has been studied extensively in the chronic disease literature. Here, we used a CM derived from a published DAG to simulate a longitudinal cohort study and examined how changing the underlying causal mechanisms (i.e. CM structure) can lead to the obesity paradox. We found that incorporating study design bias (i.e., not adjusting for age-varying mortality rates or reverse causation), can lead to the obesity paradox. Ultimately, we show how mathematical modeling of DAGs can be used to simulate epidemiological studies, inform analyses, and explore underlying biases that may be important for understanding epidemiological study data.

2.2 Introduction

Designing analyses to accurately estimate the effect of an exposure on outcome requires understanding how variables relevant to a study question are causally related to each other. Directed acyclic graphs (DAGs) are diagrams used to graphically map causes and effects to separate associations due to causality versus those due to bias. Compartmental model (CMs) depict parameterized flows between disease states over time [32, 33] and can be used to explicitly represent mechanisms underlying disease progression or transmission [34, 35]. Given the causal nature of both DAGs and CMs, a question arises of whether these two approaches may be linked. Indeed, Ackley et al. provided a formal mapping from the basic building blocks of DAGs (e.g. causality, confounding and selection bias) to CMs [32]. See Figure 2.1 for an example illustration and Appendix Section A.1 for a more in-depth comparison between DAGs and CMs. A DAG and CM are defined as ‘corresponding’ if they represent the same conditional independencies. Despite this published mapping, an in-depth exploration of its utility and examples deriving CMs from more realistic

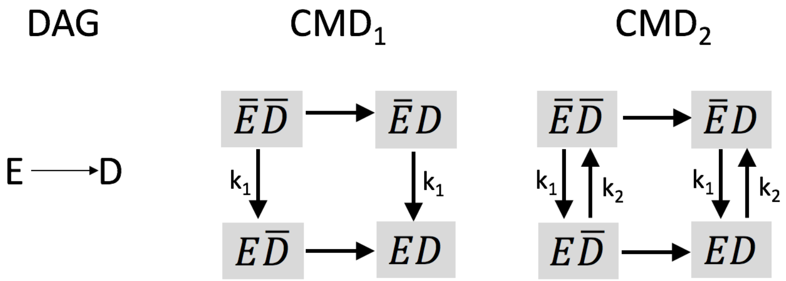


Figure 2.1: From left to right: A simple DAG showing causality wherein exposure E causes outcome D. Next, in CMD_1, We will assume that E and D are both dichotomous and this corresponding CM will have 2^n states where $n = 2$ since there are 2 random variables on the DAG. Additionally, D status does not affect E status. The $\bar{X}$ notation denotes *not* X, so $\bar{E}$ is unexposed. Thus the rates at which individuals become exposed (i.e. go from $\bar{E}$ to E) are the same whether or not they have D—equal rates are denoted by the same parameter value and if the parameter symbol is not indicated, distinct rates are assumed. This CM is further asserting that once an individual becomes diseased or exposed, they cannot return to the non-diseased or non-exposed state. In CMD_2, we see that individuals can move from E to $\bar{E}$, but their D status does not affect the rate at which they transition as indicated by the equal rates between $\bar{E}D$ to $\bar{E}\bar{D}$ and ED to $E\bar{D}$.

DAGs such as those in published literature has not previously been done.

In this chapter, we extended the work by Ackley et al. by simplifying the mapping to reduce the combinatorial explosion of CM compartments that results from realistic DAGs (taking advantage of simplifications to the CM that can be included when conditioning on a variable and tracking mortality). We then developed an operationalized workflow which uses this mapping to simulate epidemiological studies. We illustrated our findings by deriving a CM from a published DAG representing the obesity paradox, the phenomenon wherein obese ever-smoking diabetics have lower mortality rates than their normal weight counterparts. We examined competing hypotheses underlying the obesity paradox by incorporating different potential biases into our CM and then simulating study data. Our method can be applied to nearly any DAG or study question to gain insight into what underlying causal mechanisms can explain patterns observed in epidemiological data. This insight can be used to reduce bias in study designs and ultimately obtain more accurate effect measures of an exposure on outcome.

2.3 Methods

2.3.1 Overview of the Obesity Paradox

The obesity paradox is the apparent protective effect of obesity on mortality among individuals with chronic diseases such as heart failure, stroke, or diabetes [2, 106–108]. In this analysis, we examined results from an observational study conducted by Preston et al. in which obese, ever-smoking (but not never-smoking) diabetics had lower mortality rates than their normal weight counterparts [2].

Figure II.2(a) shows the published DAG from the observational study [2] representing the obesity paradox. The exposure is body mass index (BMI) and is coded as either overweight/obese (BMI $\geq$ 25 kg/m^2) or normal weight (BMI = 18.5-24.9 kg/m^2) and the outcome is mortality. Individuals are considered to have diabetes or prediabetes if their hemoglobin A1c is less than 5.7%, or if they have been previously diagnosed. Smoking is a common risk factor for diabetes, mortality, and BMI, and is coded as ever-smoking ($\geq$ 100 cigarettes over the course of an individual's lifetime) or never-smoking ($<$ 100 cigarettes). The mortality rates were age-standardized according to the 2000 census using age groups 40-59 and 60-74. For simplicity of notation, we will refer to pre-diabetics and diabetics as 'diabetics' and overweight and obese as 'obese'.

To determine how to obtain an unbiased effect measure of BMI on mortality, we can refer to the structure of the DAG from Preston et al. (Figure II.2(a)). Overall, if we assume that there are no other sources of bias in the study, and no other common causes of the variables on the DAG, an unbiased effect estimate would require that we adjust for smoking status. Diabetes is a common cause of smoking and BMI (or collider). Although conditioning on a diabetes will create a spurious association between its causes i.e., selection bias [109], adjusting for smoking removes this. Additionally, diabetes is a mediator on the path between BMI and mortality. To account for the fact that we are conditioning on a mediator, we can assume that there are no additional unmeasured confounders and only consider the controlled direct effect of BMI on mortality i.e., when diabetes is held constant [110]. See Appendix Section A.2 for a more details.

There are numerous potential explanations for the obesity paradox. Example explanations due to bias in study designs include reverse causation [2], confounding, selection bias [2, 111], or inaccuracy of BMI in representing body composition [112]. Study design bias may be due to underlying causal mechanisms that have not properly been adjusted for in the analysis (e.g. reverse causation). Causal explanations include the fact that obese individuals may receive better medical treatment [113], or might be specific to the chronic disease e.g. obese individuals may be protected from plaque formation on their arteries through a greater mobilization of endothelial progenitor cells [114].

For the purposes of this study, we will define the obesity paradox based on the qualitative results of the Preston et al. study i.e., the obesity paradox occurs when obese *never*-smoking diabetics have *higher* rates of mortality than normal weight never-smoking diabetics and obese *ever*-smoking diabetics have *lower* rates of mortality than normal weight ever-smoking diabetics. We also assumed that comparable individuals who are obese or ever-smokers always have higher mortality rates than their normal weight or never-smoking counterparts, respectively. In other words, we only considered biases in study designs (specifically reverse causation or selection bias) as potential explanations, rather than examining situations where we model obesity as being biologically protective.

2.3.2 Workflow Summary

We propose the following workflow to simulate epidemiological studies and conduct statistical analyses on CMs derived from DAG:

1. **DAG and Study Design**. Design or use an existing DAG representing the causal processes related to a given exposure and outcome, and then plan an epidemiological study to simulate using the model. Using the DAG, determine which variables will be controlled for in the statistical analysis (see Step 5 below). In our analysis, we started with a published DAG [2].

2. **DAG→CM Mapping**. Derive a CM from the DAG using the mapping described by Ackley et al. [32].

   (a) Because multiple CMs may match the given DAG, decide the appropriate CM based on the chosen study design and realistic mechanisms for the process of interest. In our CM, individuals can transition from never-smoking to ever-smoking, but not back to never-smoking. In general, the research question and hypotheses will guide how to correctly derive a CM from a given DAG since the correspondence between DAGs and CMs is not one to one. [32].

   (b) Potentially reduce the state-space for the chosen CM based on the study design. In our analysis, the study design conditions on diabetes, so we can simplify the model state-space to only include the diabetic states.

3. **Simulation and Sampling**. Simulate the chosen study population using the CM based on predefined ranges of parameter values and initial conditions. In our analyses, we simulated a yearlong longitudinal cohort study among diabetics aged 40-74 (this matched the population in the observational study) for each sampled parameter set.

   (a) Parameter and initial condition values and ranges can be determined based on the mechanism of interest, existing data, the literature, or simply broad ranges that encompass the plausible space of values (as were used in our analysis). Values may be (for example) uniformly sampled from these ranges using Latin Hypercube Sampling (LHS) [115].

   (b) Simulation of the study using the chosen CM can be implemented in a variety of ways, e.g. as ordinary differential equations or as a stochastic model.

4. **Generate Simulated Data**. Generate a simulated dataset based on the outcome of interest and measurement details of the study (e.g. number of follow-up time points, variables measured, potential sampling or measurement error). In our case, individuals were followed up once at the end of the study, we made a single simulated dataset for the entire study because individuals were followed up once at the end of the study. For simplicity and because we simulated a very large study (1,000,000 individuals), we did not examine issues of sample size or measurement error. We subsequently calculated person-time (to estimate time at risk for the study population over the course of the study) and incident mortality by disease state.

5. **Analysis and Evaluation**. Run statistical analyses or calculate outcomes using the simulated data in Step 4. Analyses may include calculation of single effect estimates and/or a wide range of statistical regression methods (depending on what analyses are of interest/planned for the study). Next, evaluate the results to examine how the causal relationships and parameters included in the model affect potential biases and patterns of interest in the data. In our analysis, we calculated mortality rate ratios or (MRRs) to compare normal weight to obese individuals within different smoking strata and then assessed whether each given model and study design could recreate the obesity paradox.

6. **Revision and Exploration**. Based on the results of Step 5, potentially alter the study design and/or DAG to explore alternative biases and causal mechanisms, then re-run the workflow. We did this by simulating epidemiological studies assuming different unadjusted study design biases (i.e., reverse causation and selection bias).

In the remainder of this paper, we simulate studies assessing how different underlying causal mechanisms might lead to the obesity paradox to illustrate the utility of this workflow.

2.3.3 Simulating a Longitudinal Cohort Study

We simulated a yearlong cohort study to examine the relationship between obesity and mortality among diabetics ages 40-74. We followed up participants once at the end of the study to calculate person-time and incident mortality by disease state. We started with a population of 1,000,000 people and (for the age-structured models mentioned below) weighted according to their age group distribution in the

2010 United States (US) census [116]. See Appendix Section A.7 for details on age-weighting for our study population and A.8 for a more detailed age-weighting scheme not implemented in our study.

2.3.4 Alternative CMs

We used 4 different models to explore how our simulated datasets change with different proposed underlying causal mechanisms. See Figure 2.2 for all DAGs and corresponding CMs. We begin with Model 1, a direct conversion of the published DAG from Preston et al. [2] to a CM. See Appendix Section A.3 for details on how we converted this DAG and reduced the number of compartments on the CM. After following the workflow for Model 1, we explored other possible mechanisms that might lead to the obesity paradox. Model 2 incorporated age-varying rates and was age-weighted according to the US census [116]. We split our population into a younger age-group (ages 40-59) and an older age-group (ages 60-74) and simulated the same model within strata of age. See Appendix Section A.6 for details on how we incorporated age into the DAG and CM. Model 3 represents reverse causation due to chronic obstructive pulmonary disease (COPD), a co-morbidity associated with diabetes for which smoking is a risk factor that can induce cachexia (loss of weight and muscle mass) and cause higher mortality rates [117–120] (thereby increasing mortality among a subset of normal weight ever-smokers). Individuals with comorbid diabetes and COPD can transition into an ‘unhealthy’ compartment, U. Individuals in U have lost weight due to cachexia and also have higher mortality rates than their normal weight ‘healthy’ counterparts (i.e. normal weight ever-smoking individuals with COPD who have not undergone cachexia). See Appendix Section A.9 for details on the underlying biological mechanism and how we incorporated reverse causation into the DAG and CM. Finally, Model 4 is a combination of Models 2 and 3. See Appendix Section A.11 for details on how we incorporated age and reverse causation into the DAG and CM.

In all CMs, once individuals die, they cannot move between disease states and we no longer track them, therefore to reduce the dimensionality of our model, mortality is an outgoing flow from each compartment and was not included in the set of disease states. Overall, we made minimal assumptions about parameter values to derive generalizable insight into the mechanisms driving the obesity paradox.

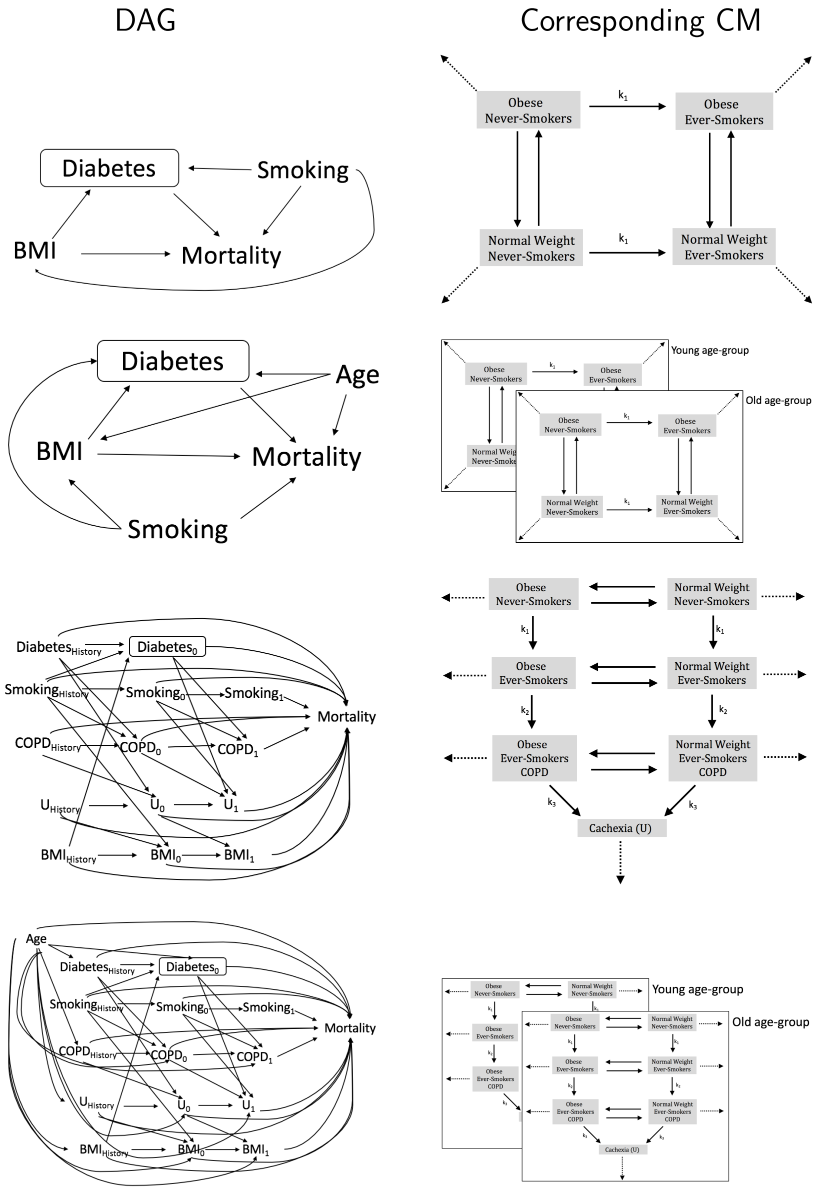


Figure 2.2: All DAGs and corresponding CMs used in our obesity paradox simulation study. DAGs (left column) and corresponding CMs (right column) for each model. By row: 1. Preston et al. [2] DAG; 2. adding in age-varying mortality rates; 3. reverse causation and 4. combined model. Mortality rates are denoted by dotted lines. Rates with no labels (including mortality rates) may all be distinct. Where 'BMI' is the exposure and 'Mortality' is the outcome. 'Age' and 'Smoking' confound the relation between 'BMI' and 'Mortality' The box around 'Diabetes' indicates that the study population is conditioned on individuals with diabetes. Cachexia is represent by 'U', and COPD is chronic obstructive pulmonary disease. With respect to the longitudinal DAGs, 'History' denotes status before the study, '0' denotes baseline, and '1' represents the end of the study i.e., one year follow up.

2.3.5 Parameterization of the CM

We conducted a sweep of parameters (transition and mortality rates) and initial states (denoted ‘parameter sets’) using LHS [115] to uniformly sample values from pre-defined ranges [40, 115]. Specifically, we allowed all compartment transition rates to vary from 1% to 20% per year. For example, this results in between 1% to 20% of obese ever-smokers becoming normal weight over the course of the 1 year study. Although 20% is unrealistically high (especially in the general population), we intentionally set a large range of parameter values to ensure that we capture realistic ranges and to see if any extreme scenarios might lead to the obesity paradox. Furthermore, we placed no restrictions on the number of individuals starting in each state and only ensured that the total number of individuals across all disease states equaled the study population at the start of the simulation. See Appendix A.12 for more details on the calculation of initial conditions. We imposed biologically realistic restrictions on the mortality rates such that ever-smokers have a higher mortality rate than their never-smoking counterparts (i.e., within weight strata), and obese individuals have a higher mortality rate than their normal weight counterparts (i.e., within smoking strata). In the age-structured models, older age group mortality rates for a given disease state were determined by multiplying the younger age group mortality rate of the same state by a scaling factor between 1 and 2. Finally, in the reverse causation models, we derived the mortality rate in the U compartment by multiplying the mortality rate of normal weight healthy ever-smokers with COPD by a cachexia scaling factor between 1 and 2 (similar to the age scaling factor in Model 2). Overall, each model represents different underlying causal mechanisms and running a model on a given parameter set represents a single simulated study. See Appendix Section A.13 for more details on sampling transition and mortality rates for each model and Table 2.1 for all LHS ranges.

Table 2.1: LHS Ranges

Parameters	Range	Models
Normal weight never-smoking to obese never-smoking	1% to 20%	All models
Obese never-smoking to normal weight never-smoking	1% to 20%	All models
Smoking initiation rate	1% to 20%	All models
Normal weight ever-smoking to obese ever-smoking	1% to 20%	All models
Obese never-smoking to normal weight never-smoking	1% to 20%	All models
COPD incidence rate	1% to 20%	Model 3 and combined model
Normal weight ever-smoking with COPD to obese ever-smoking with COPD	1% to 20%	Model 3 and combined model
Obese ever-smoking with COPD to normal weight ever-smoking with COPD	1% to 20%	Model 3 and combined model
Cachexia initiation rate	1% to 20%	Model 3 and combined model
Baseline mortality rate	1% to 10%	All models
Add on for smoking	0% to 10%	All models
Add on for obesity	0% to 10%	All models
Age-varying mortality scaling factor	1 to 2	Model 2 and combined model
Add on for COPD	0% to 10%	Model 3 and combined model
Cachexia (U) scaling factor	1 to 2	Model 3 and combined model

2.3.5.1 Data Generation and Statistical Analysis

After running each model with 10,000 randomly sampled parameter sets [121], we calculated person-time and incident deaths per compartment for each study (i.e. for each model and sampled parameter set). See Appendix Sections A.14 and A.16 for more information on these calculations. Next, we calculated MRRs comparing normal weight to obese individuals within smoking strata to measure the effect of BMI on mortality. As mentioned, to recreate the obesity paradox (as per [2]), the MRRs from the simulated data must simultaneously show normal weight never-smokers with *lower* mortality rates than their obese counterparts, and normal weight ever-smokers with *higher* mortality rate than their obese counterparts.

In Model 1, we measured all compartments and calculated the MRRs directly from the simulated data. In Model 2 (age), we initially did not adjust for age as a confounder. Rather, MRRs were calculated by taking the sum of incident deaths divided by the sum of person-time for a given disease state across age-groups. As a sensitivity analysis, we did adjust for age by externally standardizing the MRRs to the unexposed (obese) group [122]. Finally, in Model 3 (reverse causation), our study design did not initially adjust COPD or related complications (i.e., cachexia). Therefore individuals with COPD were measured together with ever-smokers (e.g., in our study population, all normal weight individuals with COPD including those with cachexia were measured together with normal weight ever-smokers). The MRRs were calculated in the same way as we did for Models 1 and 2. Therefore, we initially did not adjust for COPD or cachexia. As a sensitivity analyses, we adjusted for reverse causation by excluding all individuals with COPD (including those with cachexia) at baseline. Finally, in the combined model, we ignored age, COPD, and cachexia in our initial analysis, and then adjusted for age only, COPD only and finally, age and COPD.

All simulations and analyses were conducted in R version 3.3.3 [123]. Compartmental models were run using the 'deSolve' package [124].

2.4 Results

Overall, we found that not adjusting for study design bias in our CMs resulted in the obesity paradox. See Table 2.2 and Figure 2.3 for all results.

Table 2.2: Results from All Analyses

Models	Unadjusted Analysis	Adjusted Analysis
Model 1: Published DAG	**No obesity paradox**	NA
Model 2: Adding in age-varying mortality rates	**Obesity paradox occurs** – when there are more younger obese or more older normal weight (selective survival bias)	Adjusting for age stops the obesity paradox from occurring
Model 3: Reverse causation	**Obesity paradox occurs** – more than in Model 2 because we created a mechanism that directly affects normal weight individuals (reverse causation bias)	Excluding those with COPD at baseline stops the obesity paradox from occurring (for 1 year, but not 5 year study)
Model 4: combined	**Obesity paradox occurs** – Predominant mechanism is reverse causation	Interactive effects between biases

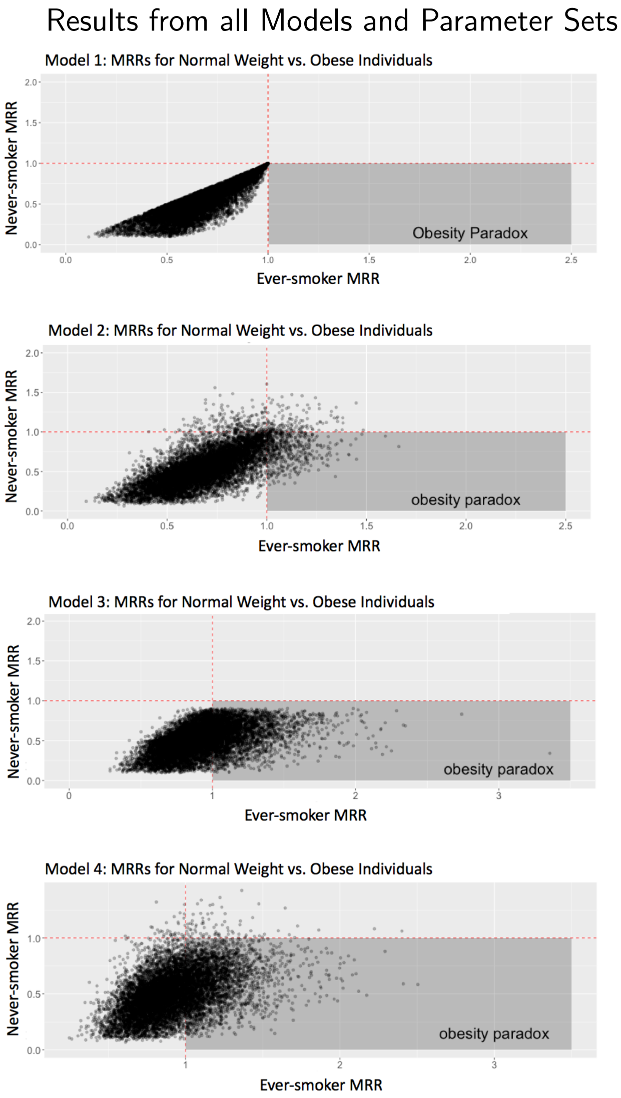
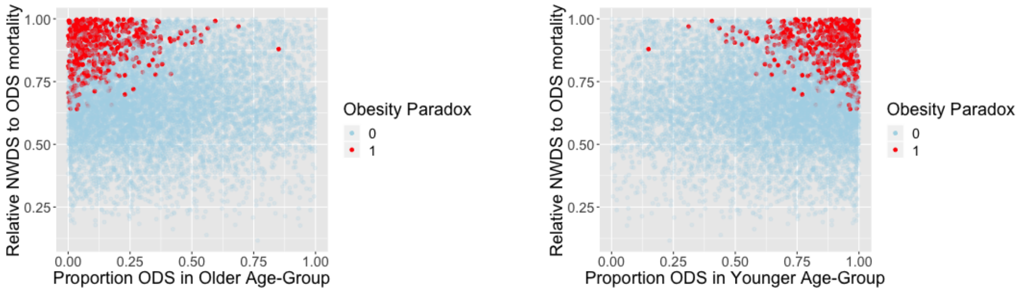


Figure 2.3: Results from all model runs. Each plot represents the MRR comparing normal weight to obese for never-smokers against the MRR for ever-smokers for each of the 10,000 LH-sampled parameter sets with each point representing a single simulated study. The obesity paradox occurs when obese *never*-smoking diabetics have *higher* rates of mortality than normal weight never-smoking diabetics and obese *ever*-smoking diabetics have *lower* rates of mortality than normal weight ever-smoking diabetics. By row: 1. Preston et al. [2] 2. adding in age-varying mortality rates; 3. reverse causation and 4. combined model.

In Model 1, we did not see the obesity paradox because the MRRs from the simulated data were simply the ratio of the CM mortality rate parameters (see Figure II.3(a)). For instance, the ever-smoker MRR is just the mortality rate of normal weight diabetic ever-smokers ($NWDS$) divided by the mortality rate of obese diabetic ever-smokers (ODS). See Appendix Section A.17 for more details. Due to the structure of Model 1 and the restrictions we placed on the parameter values, mortality rates for normal weight individuals were always lower than (or at the very least equal to) their obese counterparts therefore, all ever-smoking MRRs were ≤ 1. Overall, Model 1 cannot simulate a protective effect of obesity on mortality among diabetic ever-smokers.

Next, in Model 2, the obesity paradox did occur in a subset of studies. (See Figure II.3(b)). Overall, among model runs that resulted in the obesity paradox, there were generally either more younger individuals in the obese ever-smoking compartment and/or more older individuals in the normal weight ever-smoking compartment. This caused the mortality effects of age to counterbalance those of obesity, resulting in the obesity paradox. In other words, for the obesity paradox to occur, age-varying mortality must be sufficiently high and work together with the relative age distribution of individuals across disease states. This is analogous to selective survival bias in which obese ever-smoking individuals are more likely to die before they reach older ages, thus there would tend to be more older normal weight ever-smokers than older obese ever-smokers. An illustration of this is the trade-off between the proportion of old vs. young individuals in the obese diabetic ever-smoking (ODS) compartment and the relative mortality rate of normal-weight diabetic ever-smokers ($NWDS$) vs. ODS (shown in Figure 2.4). The majority of parameter sets that resulted in the obesity paradox show the proportion of individuals in the older age-group among all ODS is $< 50\%$. Additionally, the effect of obesity on mortality is relatively low (i.e., the $NWDS$ mortality rate is consistently similar to the ODS mortality rate in the parameter sets that resulted in the obesity paradox). Finally, as the proportion of individuals in the older age group increases, the effect of obesity on mortality decreases even more. This is analogous to obesity becoming less risky as individuals age [125]. In the age-standardized sensitivity analysis, no runs resulted in the obesity paradox (results not shown).

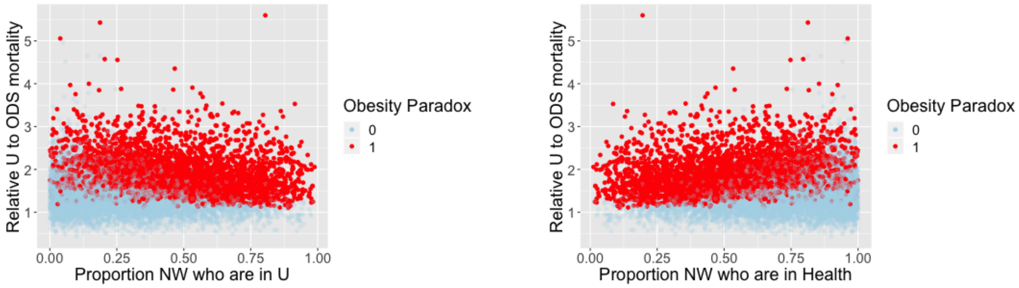


(a) ODS in Older Age-Group

(b) ODS in Younger Age-Group

Figure 2.4: Age-weighting and relative mortality among Model 2 runs. The proportion of obese diabetic ever-smokers (ODS) who are old (left) and young (right) at the beginning of the simulation is displayed on the x-axis i.e., if equal to 0.5, half of the individuals in ODS are in the older age group and half are in the younger age group. The relative mortality of normal diabetic ever-smokers ($NWDS$) to ODS is displayed on the y-axis i.e., if equal to 0.5, the $NWDS$ mortality rate would be half of the ODS mortality rate. Parameter sets that resulted in the obesity paradox are in red (coded as '1') and sets that did not result in the obesity are in blue (coded as '0').

In Model 3 (compared with Model 2), more runs resulted in the obesity paradox (Figure II.3(c)). This is due to the fact that the reverse causation mechanism differentially affects normal weight ever-smoking individuals (compared with obese ever-smoking individuals). Therefore, the obesity paradox depends on (1) the relative obese and normal weight (healthy and unhealthy) mortality rates and (2) the distribution of individuals in healthy and unhealthy compartments. On the other hand, in Model 2, age-related mortality affects normal weight and obese individuals in the same manner and thus relies on the population distribution across more compartments i.e., both age groups in ever-smoking obese and normal weight compartments. Because healthy and unhealthy normal weight ever-smokers are measured together in our observational study, the unhealthy mortality rate increases the combined (healthy and unhealthy) normal weight ever-smoking mortality rate such that the overall normal weight ever-smoking mortality rate is higher than the obese ever-smoking mortality rate and the obesity paradox occurs. This mechanism of weighting the overall normal weight ever-smoking mortality rate is revealed in the relative proportion of individuals starting in different disease states (Figure 2.5). For instance, for runs in which the obesity paradox occurs, the relative mortality rate of individuals who are unhealthy compared to those who are obese ever-smokers increases when fewer normal weight individuals start in the unhealthy compartment.



(a) Proportion of Normal Weight Individuals who are Unhealthy

(b) [Proportion of Normal Weight Individuals who are Healthy

Figure 2.5: Age-weighting and relative mortality among Model 3 runs. The proportion of normal weight individuals (ODS) who are unhealthy (left) and healthy (right) at the beginning of the simulation is displayed on the x-axis. The relative mortality of unhealthy individuals to ODS is displayed on the y-axis i.e., if equal to 0.5, the U mortality rate would be half of the ODS mortality rate. Parameter sets that resulted in the obesity paradox are in red (coded as '1') and sets that did not result in the obesity are in blue (coded as '0').

The results from our sensitivity analyses reveal that excluding individuals with COPD at baseline reduced the number of model runs that result in the obesity paradox to 1 (compared with 3,114 in the unadjusted version). If we run the study for 5 years, only 82 model runs resulted in the obesity paradox (results not shown). This highlights the importance of inclusion and exclusion criteria in an initial study population in recreating the obesity paradox.

Finally, in the combined model, we found that the reverse causation mechanism leads to the obesity paradox substantially more than the age-weighting (selective survival) mechanism. This is evidenced by the fact that in 91.7% of all runs in which the obesity paradox occurred, the normal weight ever-smoking unhealthy (U) mortality rate is higher than both the obese ever-smoking and obese COPD mortality rates within each age strata. The results from our sensitivity analyses reveal that when we control for both age and COPD, the obesity paradox is avoided almost completely. Interestingly, when we standardize age or exclude individuals with COPD only, certain parameter sets that didn't previously result in the obesity paradox, now demonstrate the obesity paradox. This indicates a 'two wrongs make a right' interactive effect between these two biases: for instance, if there are more younger normal weight individuals this might counteract the effects of a high proportion of individuals starting in U in the unadjusted model, but if we adjust for age only, the proportion starting in U may result in the obesity paradox.

2.5 Conclusion

We have developed a workflow that can be used to explicitly examine the underlying conditional independencies of DAGs. This method provides a systematic way to quantitatively evaluate bias and provide insight into the causal relationships between variables in a study. Our workflow can be applied to nearly any study question assuming standard assumptions (e.g., assuming no faithfulness violations on DAGs [126]). Modeling DAGs and conducting simulated studies can provide insight into how to design sound observational studies and analysis plans. For instance, if results from a simulated study don't match expected results, this may provide insight into unmeasured and/or unadjusted covariates, or interacting biases as we found in our obesity paradox simulation study. Although traditional analyses using DAGs would have likely found the same main sources of bias (i.e., unadjusted covariates), our method also identified some additional biases (that would not have been easily identified using traditional methods (e.g., interaction between age and reverse causation, excluding individuals with COPD at baseline and then running the study for 5 years). Overall, we simulated epidemiological study data in a structured manner based on the conditional independencies of DAGs to test different hypotheses.

We successfully recreated the obesity paradox by deriving a compartmental model from a published DAG [2] and then incorporating two different unadjusted biases. In Model 1, we found that direct conversion of the published DAG was not able to recreate the obesity paradox. In Model 2, we incorporated age-varying mortality and found that the relative proportion of individuals in different age groups across disease states can create a selective survival bias causing the obesity paradox. In Model 3, we found that reverse causation caused by an unmeasured disease state can more effectively cause the obesity paradox compared with the age-varying mortality model. The reverse causation mechanism was more effective because it differentially affected normal weight ever-smoking individuals (compared with obese ever-smoking individuals). Finally in the combined model, we observed how different biases can interact to cause or prevent the obesity paradox from occurring. Overall, adjusting for biases in these models (sensitivity analyses) made the obesity paradox nearly non-existent, indicating that incorporating bias and not adjusting for it correctly is required to recreate the obesity paradox (assuming the protective effect of obesity is not truly present, and that we have sufficient sample size). Ultimately even with very general parameter assumptions for our model, we were able to relax specific assumptions about parameter values and initial conditions to derive general insight into what causal mechanisms may drive the obesity paradox.

Limitations of this study include the fact that the DAGs we use are overly simplified (despite our use of a published DAG) and do not represent the complete state of knowledge about the relationships between variables relevant to the study question. We decided to use relatively simple DAGs to more effectively illustrate our workflow. It is simple to make more realistic DAGs by adding additional demographic characteristics e.g. race, socioeconomic status, access to medical treatment and including these would simply require vectorizing our equations further (as we did for the extension from Model 1 to Model 2). However, since we are not fitting these models to study data, we would have added more parameters to our models without truly adding any information. Because each new DAG variable doubles the number of equations in the CM, this would add complexity without insight. We aimed to strike a balance between realism and parsimony in our models to isolate and examine the qualitative effects of individual causal mechanisms of interest. For instance, the effects of race may counteract the effects of age leading to overly complicated results (i.e. identifiability issues may obscure the larger point). A potential future direction is to construct larger DAGs from the literature and make simplifying assumptions to reduce the corresponding CM's dimensionality (such as including only one variable among a colinear set). For instance, suppose both BMI at baseline and BMI history are included on a given DAG, one could assume that history is a proxy for baseline BMI among e.g. adults [127] and collapse these two variables into a single BMI variable. The robustness of results to this simplifying assumption can also be explored using our workflow. Relatedly, our workflow could also be used to identify which variable(s) on a DAG are sufficient or necessary to replicate a particular pattern in the data (e.g., by systematically removing variables and simulating the results). Finally, individually-based models may be used for study questions requiring more detailed demography. Another weakness is that our crude estimate of person-time (see Appendix Section A.14) will not work if the dynamics of the model are very fast. It is possible to calculate person-time precisely by tracking the flows in and out of compartments separately.

Strengths of this study include the methodological contributions to using CMs in conjunction with DAGs to understand patterns seen in the data. We extended the mapping that Ackley et al. provided [32] and proposed a method for comparing simulated data with epidemiological study data. This method can be expanded for different types of epidemiological analyses and can also be used for different purposes e.g. relaxing statistical assumptions, multifaceted sensitivity analyses or exploring counterfactual scenarios. We were able to show that the DAG presented in the Preston et al. paper did not on its own adequately describe the obesity paradox and then proposed alternative mechanisms and DAGs that could recreate the obesity paradox. Furthermore, we gained insight into what hypothetical causal mechanisms could result in the obesity paradox with limited data informing our model. Additionally, conducting the random sweep (i.e. LHS) of the parameters and initial conditions allowed us to account for uncertainty and draw general qualitative conclusions about the structure of the model and its effects on our statistical results. Ultimately, our workflow can help explicate causal mechanisms to explore whether or not DAGs are valid representations of hypotheses in question even when data is limited. Additionally, CMs derived from DAGs can be used as a testing ground for competing causal mechanisms to determine which ones can most closely explain patterns seen in observational study data. This represents a departure from the standard paradigm of fitting CMs to epidemiological data where instead, we operationalize causal relationships depicted on the DAG to simulate epidemiological study data.

Additional future research can include other statistical analyses. For instance, a Poisson regression model (for count data) can calculate MRRs and can be useful if conditioning on multiple variables (see Appendix A.18). Alternatively, simulated data can be individuated and other types of regression models can be run. Model parameters can be tuned to quantitatively recreate specific datasets which might be useful for gaining insight into specific study results or a specific target population. Additionally, model parameters can be informed directly from data. For instance, see Appendix Section A.17 for notes on how to parameterize the mortality rates from data. Similarly, the data collection process itself can be simulated in the compartmental model, allowing one to assess how issues such as measurement error or insufficient power might affect the relationships reflected in the DAG.

Overall, we presented here a new utility for CMs derived from DAGs: testing hypotheses to understand patterns seen in study data. We also proposed a method to compare simulated data with epidemiological study data that can be used to test competing hypotheses. We used our method to determine that a DAG from the literature was not complete and could not explain study results (i.e., it could not recreate the obesity paradox) by itself. We therefore simulated two alternative causal mechanisms and derived corresponding DAGs that could recreate the qualitative results of the study. Ultimately, simulating study data by operationalizing the causal relationships on DAGs can provide insight into how to design sound observational studies and analysis plans.

CHAPTER III

Exploring the impact of variation in spatial patterns of community contact on the effectiveness of household- vs. community-based screening interventions for Tuberculosis

3.1 Introduction

Tuberculosis (TB) is the leading cause of death from infectious disease worldwide killing 1.6 million individuals in 2017 [128]. Additionally, it has recently been estimated that $\sim 25\%$ of the global population has a latent TB infection (LTBI) [59]. Individuals with LTBI cannot transmit TB, but are at risk for progressing to active disease months to years after the initial infection. The majority of individuals with LTBI who progress to active TB do so within the first 2 years following infection [62]. This immense burden of latent disease poses particular challenges for TB control and elimination: Without guidance on how to efficiently find those recently infected individuals who are most likely to manifest infectious TB, the dramatic reductions in global TB incidence laid out in the World Health Organization's (WHO) END TB goals [128] are unlikely to be achievable.

Evidence from high-incidence settings has suggested that screening of the household contacts of individuals with active TB (or household contact tracing (HHCT)) is an under-utilized tool for finding and treating (i.e., with preventive therapy (IPT) [65, 67]) recently-acquired latent infections [17, 18]. HHCT has been successfully used in high-income, low-incidence countries for decades [9, 10]. Although a recent contact tracing trial in a high-incidence settings, showed promise for finding new cases of TB [12], other trials have yielded no significant population-level effects [11, 13, 14]. These mixed results raise questions about whether contact tracing in LMICs can be optimized to quickly find new cases of active and latent TB to reduce population-level risk in a cost effective manner.

On the individual level, TB is transmitted through close or casual contact with an infected individual. Close contacts are repeated and occur over extended periods e.g., within households or workplaces [129], while casual contacts are more ephemeral and may occur between individuals who do not know each other such as in a bus or a store [130]. Both types of contact contribute differentially to the spread of TB [17, 18, 56], with close contacts concentrating infection, and community contacts serving as bridges between these more tightly clustered units. On the population-level within cities, the distribution of TB is spatially heterogeneous. For example, incident cases can cluster forming high burden hotspots [20, 131–136]. A mathematical modeling analysis found that targeting hotspot transmission can provide an efficient way to reduce city-wide TB risk in Rio De Janeiro [19]. Given the complexity of different contact networks, and the spatial heterogeneity of TB, screening interventions targeting one transmission cluster e.g., a household or neighborhood, may have far-reaching effects on other distal parts of the population.

The potential of HHCT in high-burden settings may not stem primarily from the role of household transmission: In settings with a high burden of TB, a majority of co-prevalent household infections have been shown to be genetically discordant [15, 137], indicating that outside-household, community-based exposures may be driving household-based clustering of infection. However, this should not be taken as evidence that households are unlikely to be an effective site of intervention. Instead, clustering of genetically distinct infections within households may be suggestive of common *community* exposures that are shared between household members (e.g. through overlapping community contacts). Indeed, sharing a household with a TB case in a high incidence setting is still a meaningful risk for infection [17, 18] and a key risk factor for developing active disease [138, 139]. Overall, we hypothesize that targeting households for TB interventions is more effective than community-based active case finding [18].

Overall, due to the contact and spatial heterogeneity underlying TB transmission, and the fact that shared community risk factors (e.g. overlapping contacts) may be driving incident cases within households, localized community risk may be detectable from the household level. In other words, clustering of risk within households may help efficiently identify localized hotspots of community transmission. Overall, it is necessary to examine whether this spatial and contact heterogeneity can be leveraged to improve the effectiveness of screening interventions for TB.

We present here a spatially explicit network model with individuals and households representing TB transmission within an urban area. Different networks were used to represent different spatial and community contact patterns. Key parameters in our model were informed from a cohort study in Lima, Peru that implemented HHCT and administered IPT [17]. We used our model to examine the performance of three screening interventions: HHCT, community contact tracing (community CT), and community-based active screening (control intervention). Specifically, we explored how each intervention performs across all networks, and then within a variety of different settings to gain generalizable insight into how spatial and contact heterogeneity can be leveraged to increase their effectiveness at reducing population-level TB incidence.

3.2 Methods

3.2.1 Data

The model parameters were derived from several different published data sources including reviews of historical studies before the onset of chemotherapy e.g. for the recovery rate of individuals with active TB [140], other contemporary systematic reviews e.g. for the late latency progression rate [61]. Additionally, the household transmission rate was derived from the large prospective cohort study that conducted HHCT in Lima, Peru [17, 67]. See Table 3.1 for a full list of model parameters and their sources.

3.2.2 Spatial Contact Network

To capture the heterogeneous spatial distribution of TB as well as differential household and community contact and transmission, we developed a spatially explicit network model consisting of close (household) and casual (community) contacts. Our model has 100,000 individuals divided evenly into 20,000 households, with 5 individuals in each household. Households are represented by fully-connected sub-networks. Each household was placed randomly in a 2 dimensional space of normalized area equal to the total number of households. Community contacts are determined by a Gaussian (or normally distributed) connectivity kernel adapted

Table 3.1: TB Parameter Values and Uncertainty Ranges

Parameter	Description (units)	Value	Source	Uncertainty Ranges
θ	Life expectancy (years)	74.78	[141]	Fixed
ϵ	Early latency progression rate (/yr)	Years 1-5: {0.0866, 0.0355, 0.0112, 0.0074, 0.0024}	[142]	Each substate rate will be multiplied by 0.0817 to 0.0905
τ	Late latency progression (/yr)	0.0005	[61]	Fixed
γ	Recovery rate (/yr)	0.12	[140]	0.09 to 0.15 [70]
κ	Active TB mortality rate (/yr)	0.12	[140]	0.05 to 0.4 [70]
β_{HH}	Household transmission rate	0.21	[18]	0 to 0.315
β_{uC}	Sampled community transmission rate, this is divided by average degree to derive β_C		Range determined by expected number of cases (see 3.2.4 for details)	0 to 7
M	Imported case rate (/yr)		Assuming 2 to 20 individuals migrate with active TB per year	2 to 20
cdr	case detection rate (/yr)	1	Assuming individuals are detected an average of 1 year after progressing to active TB	Fixed
txd	Treatment duration (months)	6	[143]	Fixed
iptd	IPT duration (months)	6	[144]	Fixed
ω	Amount of immunity conferred by current disease state (%)	For ω_S, $\omega_E L$, $\omega_F L$, ω_R, ω_I, ω_T, ω_D: {0, 80%, 80%, 80%, 100%, 100%, 100%}	[145, 146]	Fixed
Network Parameters				
n	Average degree i.e., the sum of the number of community contacts and number of household contacts	{50, 100, 150, 200, 250, 300, 350, 400}	NA	
σ Standard deviation	Average Degree Distance	{0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 3.25, 3.5, 3.75, 4, 4.25, 4.5, 4.75, 5}	NA	
ρ Network density		Fixed at 1	NA	

from Lang et al. [3]. This kernel controls the overall neighborhood size in which contacts are formed and numbers of contacts, with the probability of a connection forming between community contacts given by:

$$f(d) = (n-4)\frac{e^{(-d^2/2\sigma^2)}/2\pi\sigma^2}{\rho} \tag{3.1}$$

where $(n-4)$ is the average number of total contacts (average degree) offset by 4 which is the number of household contacts. Next, d is the distance between nodes, σ is the average connection radius, and ρ is the density of the network. TB can only be transmitted between connected individuals. See Figure 3.1 for a schematic of the household and community network structure.

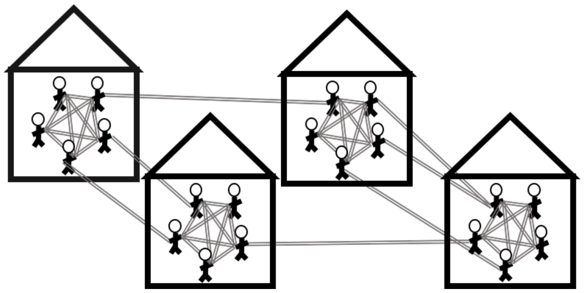


Figure 3.1: Schematic of Network Structure: All individuals are fully connected within their households. Individuals form community contacts based on a Gaussian (normally distributed) Connectivity Kernel [3]. Networks consist of 100,000 individuals divided evenly into 20,000 households.

We generated a wide array of networks with different input parameter values. For each set of network parameters, we generated 10 network realizations to account for random variation. Specifically, we varied the average degree or number of community contacts (n) and average connection radius (network standard deviation) or the distance within which individuals form community connections (σ). We examined a wide range of networks to determine how clustering and variation in community contact affected incidence and the performance of screening interventions. Overall, we generated networks covering a large range of parameter space from very local and close-range community connections to approximately spatially random and long range community connections. We assumed constant household and total population sizes, and static network connections. See Appendix Figures B.1 and B.2 for features of generated networks (for one set of realizations) by different n and σ values.

3.2.3 TB Transmission Model

We adapted a model of coupled household and community TB transmission previously published by Kasaie et al. [70] to simulate TB transmission on the generated networks. See Figure 3.2 for the model schematic. In this section, we will first outline the **natural history** model, which includes the set of possible state transitions for infected individuals in the absence of intervention. We will then outline the **intervention model** which adds intervention-relevant states and state transitions to the underlying natural history model.

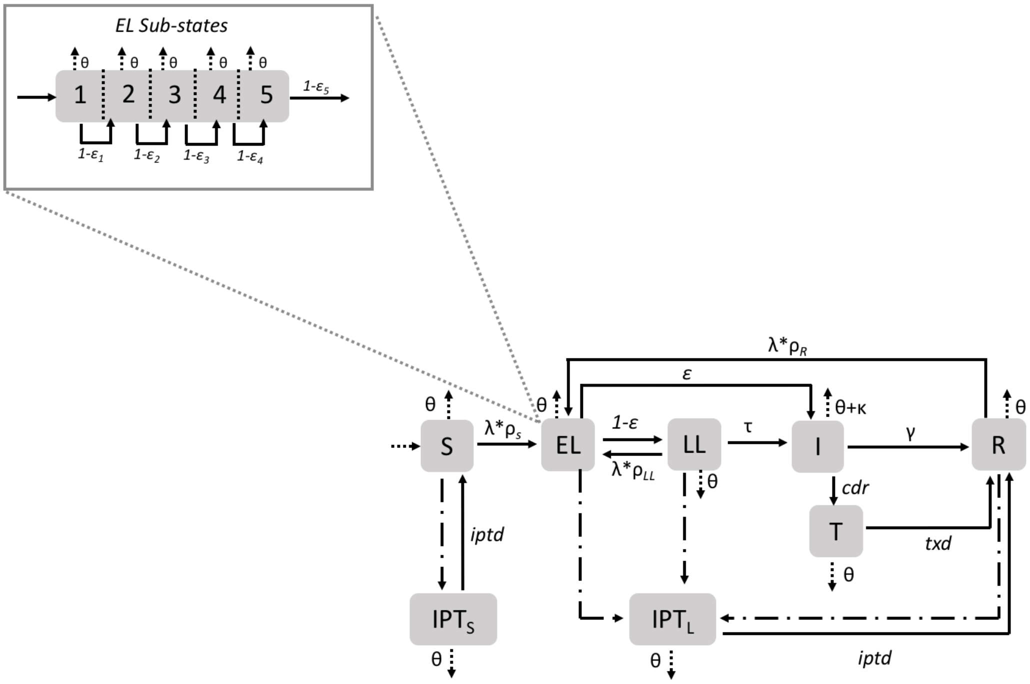


Figure 3.2: Schematic of TB Transmission Model. Where S is susceptible; EL is early latent; LL is late latent; I is active TB; R is recovered; T is treatment; IPT_S is susceptible individuals who are given preventive therapy and IPT_L are all other individuals given preventive therapy (i.e., those who originate in EL, LL or R. Individuals transitioning to treatment states are represented by dot-dashed lines. The rate at which individuals transition to the IPT states depends on the screening scenario and number of contacts. Vital dynamics are represented by dotted lines.

3.2.3.1 Natural History Model

The natural history model includes 5 main states: susceptible and not infected (S), early latent or those infected within the last 5 years, but have not progressed to active TB (EL), late latent or those infected more than 5 years ago, but have not progressed to active TB (LL), infectious or those with active TB (I), and recovered or those who have previously been infected and/or diseased, but were cured either spontaneously or through treatment or preventive therapy (R).

In our model, susceptible (S), late latent (LL), and recovered (R) individuals can become infected (or reinfected in the case of individuals in the LL and R states) according to the force of infection which can be divided into household, community, and external transmission (from imported TB cases), and is defined as:

$$FOI = [\beta_{HH} I_{HH} + \beta_C I_C + \frac{M}{N}](1 - \omega) \tag{3.2}$$

Where β_{HH} and β_C are the **per-contact** transmission rates of household and community contacts, respectively. The number of active TB household and community contacts are I_{HH} and I_C, respectively. Additionally, to simulate the TB risk due to migration, we average the impact of imported cases over the entire population to derive the external force of infection. Specifically, M is the active TB case migration rate and N is the total population. Because prior infections confer limited protective immunity [145, 146], the force of infection experienced by individuals with prior infection (relative to a fully susceptible individual) is modulated by ω. For instance, fully susceptible individuals have no added protection, but those who have been infected and subsequently recovered are 80% less likely to contract TB.

Once an individual becomes infected, they move into the non-infectious early latent state (denoted EL). Between 10 and 20% of individuals in this state will progress to the infectious (diseased) active TB state (I) within 5 years with annual rates of progression decreasing with time since infection (see EL sub-states in Figure 3.2) [142]. After 5 years, the remaining 80% to 90% of individuals in the EL state who do not progress to I transition to the non-infectious late latent state (LL). Individuals in LL may also progress to I, but the rate of progression is substantially lower than EL such that only 5-10% of individuals in late latency progress over the course of their lifetimes [61]. I individuals may die from TB, or spontaneously recover (and move to R) [140]. Individuals in LL or R may become reinfected and move back to the EL_1 sub-state.

Any individual may die of natural causes and upon death is immediately replaced with a susceptible individual. To avoid placing a susceptible individual into an endemic hotspot, we implemented a reshuffling scheme based on [147], in which individuals who die from TB are replaced with one of their non-infectious contacts (i.e., they could be in any state except for I) who in turn, are replaced with one of their non-infectious contacts and so on until the 3^{rd} replacement is made. Replacements take on the same contact network (i.e. household and community contacts) of those that they replace. Then, a new susceptible individual is born into that 3^{rd} individual's previous location. This shuffling scheme maintains the network structure and constant number of individuals per household over the course of the simulation. This scheme also decreases the number of susceptible individuals put directly into a location with high TB transmission (which would ultimately overestimate the amount of TB transmission occurring) by forcing births of susceptible individuals to occur in separate part of the network.

3.2.3.2 Intervention Model

In our model, we implemented two types of screening interventions: passive and active case finding. First passive case finding included treatment for passively detected primary household cases (T), representing the current standard of care in high-burden settings [148]. Second active case finding explored the impact of delivering screening and treatment for LTBI and active TB disease to all of an individual's household contacts, a selection of their community contacts, or both. A recent model-based analysis examined the direct effects conferred by HHCT compared with community based active screening and found that HHCT resulted in a higher reduction in cases [18]. Here, we explored the impacts of these screening interventions on the population-level in a spatially explicit environment.

Passive Case Finding Individuals with active TB who were found through passive surveillance were given treatment (moved to the T state). Once in the treatment state, individuals were assumed to be non-infectious and eventually recovered (moved to R). Cases detected through passive surveillance triggered the active case finding interventions (below).

Active Case Finding For the active case finding interventions, once an infectious individual was detected through passive surveillance, either a random selection of community members or the household or community contacts of the index case (depending on the screening scenario) were screened for TB.

If any of the additionally screened individuals had active TB, they were given treatment and moved to T. All other screened individuals not found to have active TB (including those who were susceptible, recovered, or latent) were given IPT. Individuals in the IPT compartments cannot become infected or progress to active TB. If individuals transition to IPT from the S compartment, they returned to S at the end of treatment. On the other hand, if individuals originated in the EL, LL, or R compartments, they moved to R at the end of treatment.

3.2.3.3 Screening Interventions

We ran the model with following screening intervention scenarios:

Passive Screening: Passive screening [143, 149] in which a fixed proportion of active TB cases were diagnosed based on the case detection rate. We assumed that on average, it takes one year for individuals with active TB to be detected. Additionally, each newly detected case triggered the active screening scenarios below and represents real-world circumstances in which passive surveillance is ongoing even when other screening interventions are introduced.

Active Case Finding: In all active case finding scenarios below, screened individuals (e.g., contacts of the index case) who were found to have active TB were given treatment and screened individuals who did not have active TB were given IPT.

- **Community non-targeted (control scenario):** Community non-targeted active screening and IPT – adapted from [18, 70]. For each case discovered through passive surveillance, 4 individuals were selected at random from the entire population during the same time step and screened for TB.

- **Community-targeted (community CT):** For each case discovered through passive surveillance, up to 4 community contacts of the index case were screened for TB during the same time step. If an individual had less than 4 community contacts, all of their community contacts were screened (e.g., if they only have 3 community contacts, all 3 were screened)

- **Household targeted (HHCT):** This scenario is adapted from [18, 70]. For each case discovered through passive surveillance, all 4 household contacts of the index case were screened for TB during the same time step.

- **Combined Scenario:** In the combined scenario, we implemented both community CT and HHCT simultaneously and compared this to community nontargeted active screening of 8 individuals.

3.2.4 Simulation Strategy

We ran the model with 3,040 uniformly sampled parameters drawn from predefined parameter ranges to account for uncertainty in their values using Latin Hypercube Sampling (LHS) [115]. The majority of the parameter ranges were set to explore whether our model outcomes were robust to changes in these parameters (i.e., to conduct sensitivity analyses). The household transmission rate range was derived from the cohort study in Lima, Peru [18], and the community transmission rate range was selected such that the annual incidence of the model matched a range of target incidence levels. See Appendix Figure B.3 for distribution of incidence levels by model runs. We selected this range to explore which screening interventions are effective across different settings. For instance, higher incidence settings might saturate transmission to the point where contact tracing does not have a noticeable effect on population level risk of TB. We randomly selected a network for each parameter set such that each network type was run 20 times (there are 152 network types (10 realizations each) with different combinations of average degree and network connection radius values). For each parameter set, we ran the transmission model 2 times using a different random number seed to account for stochastic variation.

Because passive surveillance of infectious individuals followed by treatment is the current standard of care, it was continuously implemented throughout the entire simulation. Next, the active case finding interventions were only run after the model had initially reached steady state. See Figure 3.3 for a timeline of interventions. In each run, we seeded the model with a single randomly placed infectious individual, and ran it (with passive surveillance only) until it reached steady-state (burn-in 1) i.e., for 2,000 one-month time steps or ~ 167 years. At that point, we implemented active screening interventions and ran the model until it reached steady state again (burn-in 2) i.e., for another 2,000 time steps. Overall, we implemented all 4 different screening interventions and a passive surveillance only scenario (discussed above in Section 3.2.3.3) for each parameter set and random number seed combination to directly compare how these interventions performed on the exact same transmission model run accounting for stochastic variation.

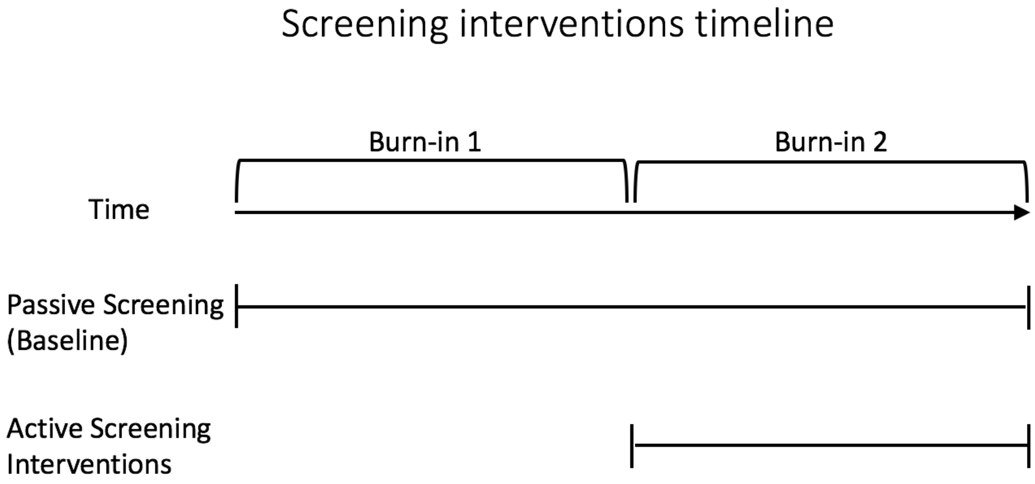

Screening interventions timeline
Burn-in 1
Burn-in 2
Time
Passive Screening (Baseline)
Active Screening Interventions

Figure 3.3: Timeline of interventions for each simulation run. Passive surveillance was run until the model reached steady state (burn-in 1), and then active screening interventions were implemented and the model was run until it reached steady state again (burn-in 2).

Examining the Impact of Single Screening Interventions To make comparisons between screening scenarios, we first calculated the 5-year average annual incidence rate at the end of burn-in 2 (i.e., post-active screening intervention incidence) for each parameter set, random number seed, and screening scenario combination. We defined the incidence as the number of new I cases over the previous 12 one-month time steps. We then calculated rate ratios (RRs) by comparing the burn-in 2 incidence level of each screening scenario for a given parameter set, and random number seed combination to the burn-in 2 incidence level of the reference group for the same parameter set, and random number seed combination.

Overall, we examined the impact of screening interventions using 2 reference groups. First, to determine if additional allocation of resources would contribute to a reduction in TB incidence (i.e., the extent of protection conferred by each screening intervention), we assessed how active case finding interventions performed compared with passive surveillance only. Second to determine how best to allocate resources among different contact tracing interventions, we examined how community and household targeted contact tracing interventions performed compared with the control intervention scenario.

RRs for each screening scenario were compared across all parameter sets and also, within strata of network parameters (average degree and average connection radius) and the 5-year average annual incidence level at burn-in 1 (i.e., pre-active screening intervention).

Examining the Joint Effects of Screening Interventions After assessing the effects of individual screening interventions, we compared the joint effects of the combined screening scenario to the component effects of HHCT and community CT to determine if there was interaction between screening programs that could confer an additional benefit without added cost. We assessed interaction on the additive scale because it is recommended that biologic interaction be quantified as such [150]. Additionally, because our risk factors (i.e., screening interventions) were preventative, we first re-coded the RRs such that the interventions conferring more protection became the reference group [151]. Therefore in our analysis, positive interaction corresponded to synergy between screening interventions and negative interaction corresponded to effects from the combined intervention being less than the sum of each individual interventions (i.e., parallelism or less than additive).

We first calculated the relative excess risk due to interaction (RERI) to determine if there is additive interaction on the rate scale:

$$RERI = RR_{HC} - RR_{0C} - RR_{H0} + 1 \tag{3.3}$$

where RR_{HC} is the RR of the combined scenario, RR_{0C} is the RR of the community contacts only scenario and finally, RR_{H0} is the RR of the HHCT only scenario. An RERI $\neq 0$ indicates interaction on the additive scale with > 0 indicating positive interaction and < 0 indicating negative interaction.

For the control reference group, we calculated RRs comparing each intervention to its corresponding control scenario. Specifically, because we screened up to 8 individuals per index case in the combined scenario, we used community non-targeted active screening of 8 individuals as the reference group. For the HHCT community CT interventions, we used community non-targeted active screening of 4 individuals as the reference group. For the passive surveillance only reference group, the post-burn-in 2 incidence levels of all screening scenarios was compared to the post-burn-in 2 incidence levels of passive surveillance only.

3.3 Results

To determine how screening interventions performed across different scenarios, we first examined whether different key model inputs (i.e., community compared with household transmission rates, and network parameters) affect TB incidence levels. Next, we examined how single screening interventions performed across different strata of these factors. Finally, we looked at potential interactions between screening interventions.

3.3.0.1 Factors Affecting Population-level TB Incidence

To determine what factors increase population-level TB incidence in our model, we first examined the relationship between the ratio of the annual number of TB infections (not active TB cases, though the number of new infections corresponds with the number of active TB cases) attributed to community vs. household transmission and annual incidence levels (Figure 3.4). Specifically for each of these metrics, we calculated yearly averages over the last 5 years of the simulation before burn-in 1. We also only included incidence levels between the 2.5^{th} and 97.5^{th} percentiles to remove outliers. Overall, for higher incidence levels, we observe that more infections are caused by community transmission. Even though the maximum household transmission rates are higher than the maximum community transmission rates (see Table 3.1), community transmission is more important in driving these higher incidence levels.

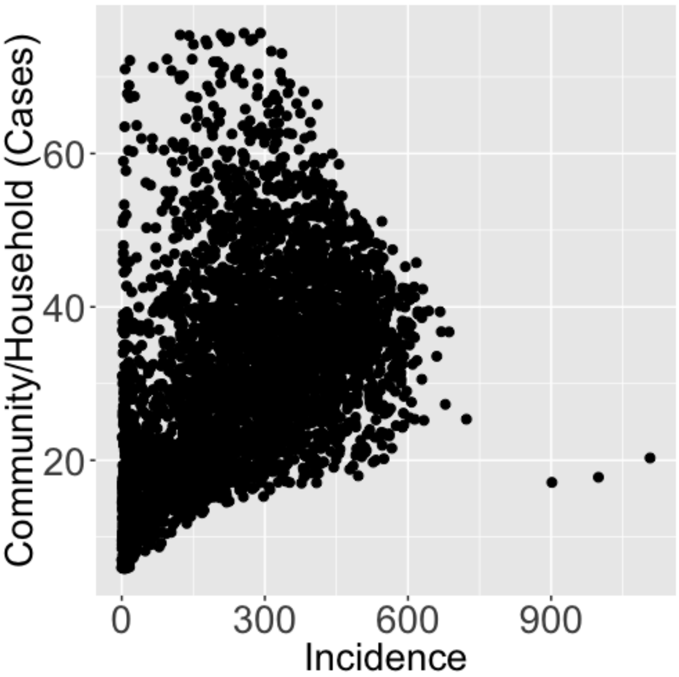


Figure 3.4: Annual number of TB infections attributed community vs. household transmission (y-axis) transmission and incidence levels at the end of burn-in 1 (x-axis).

We next examined how network parameters and transmission rates might together affect population-level incidence (Figure 3.5). Overall, we can see that for the community transmission rate to effectively cause higher incidence levels, the average degree and average connection radius must be sufficiently high (Figures III.5(a) and III.5(b)). On the other hand, there does not appear to be any increasing trend in incidence levels corresponding to household transmission rates (Figures III.5(c) and III.5(d)).

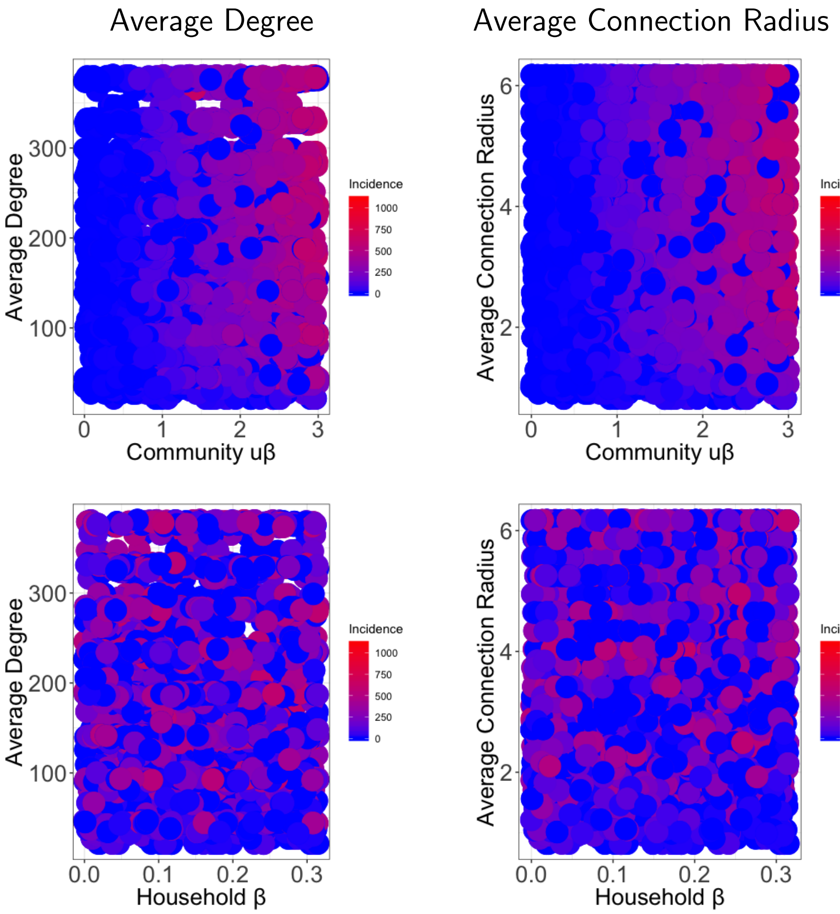


Figure 3.5: Network parameters (i.e., average degree, average connection radius) and transmission rates colored by incidence levels. Average degree (left column) and average connection radius (right column) for each model. By row: 1. network parameters vs. community transmission rates 2. network parameters vs. household transmission rates.

3.3.0.2 Protection Conferred by Single Interventions by Community Transmission Rate

We next fit splines to the relationship between β_{uC} and all RRs within the 2.5^{th} to 97.5^{th} percentiles, and plotted the results among different screening scenarios. See Figure 3.6 for results using passive surveillance only as the reference group.

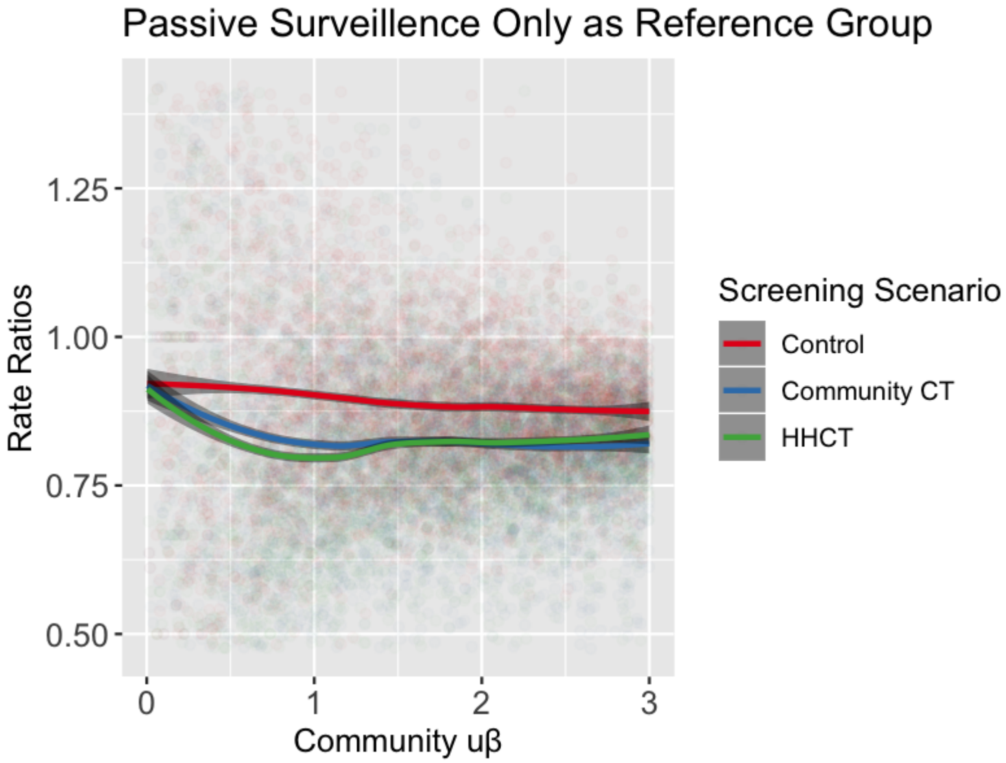


Figure 3.6: Fitted splines representing relationship between the unscaled community transmission rate (β_{uC}) and RR (among RRs within the 2.5^{th} to 97.5^{th} percentiles). Passive surveillance only is the reference group. Dots represent individual model runs colored by screening scenario and lines are the splines (with 95% confidence intervals in shaded regions) which were calculated using the loess method in R [4].

The average RRs in all 3 scenarios (control, HHCT, and community CT) were < 1 regardless of the community transmission rate indicating an overall protective effect. Contact tracing interventions performed substantially better than the control intervention. Finally, HHCT performs better than community CT for lower community transmission rate values, but worse than community CT for higher community transmission rate values. However, there is substantial overlap in the 95% confidence intervals. Trends were the same when comparing the performance of screening interventions with the control intervention as the reference group. The only notable difference was the passive surveillance only intervention conferred RRs > 1

(Appendix Figure B.4).

3.3.0.3 Protection Conferred by Single Interventions by Network and Incidence Strata

We next calculated the mean RRs (among RRs within the 2.5^{th} to 97.5^{th} percentiles) and standard deviation (SD) within strata of network parameters and burn-in 1 incidence. For each network parameter or incidence level strata, we defined ‘high’ as $>$median across all model runs and ‘low’ as $<$median across all model runs. See Table 3.2 for performance of screening interventions with passive surveillance only as a reference group in order of most effective scenario to least effective scenario.

Table 3.2: RRs of Screening Interventions–Passive Surveillance as Reference Group

Screening Intervention	Connection Radius Strata	Average Degree Strata	Incidence Strata	Mean RR (SD)
Community CT	low	low	low	0.75 (0.18)
HHCT	low	low	low	0.76 (0.17)
Community CT	low	low	high	0.78 (0.1)
HHCT	low	low	high	0.8 (0.09)
Community CT	high	low	high	0.82 (0.09)
Community CT	high	low	low	0.82 (0.16)
HHCT	low	high	low	0.82 (0.18)
HHCT	high	low	low	0.83 (0.17)
HHCT	low	high	high	0.83 (0.09)
HHCT	high	high	low	0.83 (0.17)
Community CT	low	high	high	0.84 (0.09)
HHCT	high	low	high	0.85 (0.1)
HHCT	high	high	high	0.85 (0.09)
Community CT	low	high	low	0.86 (0.18)
Community CT	high	high	low	0.87 (0.17)
Control	low	low	high	0.87 (0.1)
Community CT	high	high	high	0.87 (0.09)
Control	low	low	low	0.87 (0.18)
Control	low	high	high	0.88 (0.09)
Control	high	low	high	0.89 (0.1)
Control	high	high	high	0.89 (0.09)
Control	low	high	low	0.89 (0.17)
Control	high	low	low	0.9 (0.17)
Control	high	high	low	0.93 (0.17)

In general, contact tracing interventions performed better than the control intervention. Furthermore, screening interventions performed better in lower incidence, lower average degree, and lower connection radius strata however, there was substantial variability.

The mean incidence was 203.4 (SD: 127.4) per 100,000 person-years among the top third (of 24 total) most effective screening scenarios and 253.3 (SD: 149.9) per 100,000 person-years among the bottom third. The mean community average degree was 102.1 (SD: 48) among the top third and 192.3 (SD: 93.7) among the bottom third. Finally among connection radii, the mean radius was 3.2 (SD: 1.5) and 3.7 (SD: 1.3) for the top and bottom thirds, respectively. When examining results only among HHCT and community CT, the mean of all RRs in the high incidence strata is 0.84 (SD: 0.1), while it is only 0.81 (SD: 0.18) in the low incidence strata. Next, the mean of all RRs in the high average degree strata is 0.85 (SD: 0.13), while it is only 0.79 (SD: 0.15) in the low average degree strata. Finally, the mean of all RRs in the high average connection radius strata is 0.85 (SD: 0.13), while it is only 0.8 (SD: 0.15) in the low connection radius strata. Therefore, within strata of incidence and network parameters, a lower network connection radius and lower average degree results in slightly better performance of screening scenarios. Again, there is substantial variability within strata.

Trends were the same when comparing the performance of screening interventions with the control intervention as the reference group. The RRs were higher and the passive surveillance only interventions conferred RRs > 1 (see Appendix Table B.1).

3.3.1 Impact of Combined Screening Interventions

Finally, we implemented a combined community CT and HHCT intervention to determine if there were any circumstances (i.e., incidence level or network type) in which screening interventions modify the effects of each other. We first fit splines to the relationship between β_{uC} and all RRs within the 2.5^{th} to 97.5^{th} percentiles, and plotted the results among different screening scenarios. See Figure 3.7 for results with the passive surveillance only reference group. The combined intervention performs substantially better across the majority of community transmission rate values. See Appendix Figure B.5 for results with the control intervention is the reference group. In this analysis, the screening interventions perform worse as the community transmission rate increases.

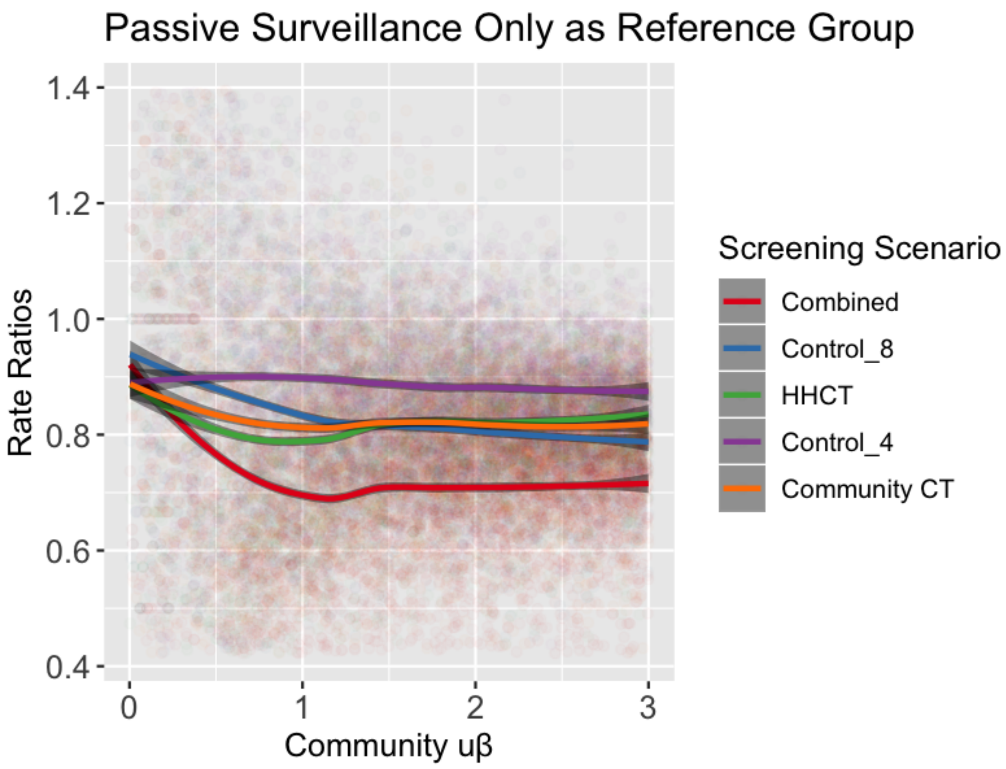


Figure 3.7: Fitted splines representing relationship between the unscaled community transmission rate (β_{uC}) and RR (among RRs within the 2.5^{th} to 97.5^{th} percentiles). Passive surveillance only is the reference group. Dots represent individual model runs colored by screening scenario and lines are the splines (with 95% confidence intervals in shaded regions) which were calculated using the loess method in R [4]. '$Control_4$' and '$Control_8$' are community non-targeted screening of 4 and 8 individuals, respectively.

See Table 3.3 for RERIs with passive surveillance only as a reference group.

We see negative interaction (RERI<0) for the majority network parameter and incidence strata settings. However, there is substantial variability. There appears to be greater negative interactive effects in strata of low incidence. But, we do not see any trends in average degree or average connection radius. See Appendix Table B.2 for RERIs with the control intervention as a reference group. Trends were similar when examining the effect of screening interventions with the control intervention

Table 3.3: RERIs of Combined Screening Interventions–Passive Surveillance as Reference Group

Connection Radius Strata	Average Degree Strata	Incidence Strata	Mean RERI (SD)
low	high	low	-0.06 (0.45)
high	high	low	-0.06 (0.38)
low	low	low	-0.06 (0.43)
high	high	high	-0.03 (0.19)
low	high	high	-0.03 (0.21)
high	low	high	-0.02 (0.21)
high	low	low	-0.01 (0.37)
low	low	high	0 (0.24)

as a reference group.

3.4 Conclusions

Overall, we found that HHCT and community CT provided substantially more protection against TB compared with passive surveillance only and the control intervention. There did appear to be slight differences between the contact tracing interventions as the community transmission rate increased, with community CT performing slightly better for higher community transmission rate values (see Figure 3.6 and Appendix Figure B.4). Although there was substantial variability in screening intervention results across strata of network parameters and incidence (as seen in the SDs of the RR estimates in Table 3.2), we did find general trends wherein screening interventions tended to confer more protection in the low incidence, low average degree, and low network connection radius strata (see Table 3.2 and Appendix Table B.1). Importantly, network parameters and incidence are correlated with each other, i.e., a low network connection radius corresponds with a lower average degree (see Appendix Figure B.2) which in turn may lead to lower incidence (see Figures 3.5), so it is not entirely clear which of these network features is altering the effectiveness of screening interventions. Presumably, transmission is too strong in higher incidence settings and on networks with higher average degrees or higher standard deviations for screening interventions to be as effective.

Consistent with previous literature, results from our model (Figure 3.4) revealed that transmission in higher incidence settings is driven by the community [15, 137]. We also found that high incidence levels only occur when there is both high community transmission rates and high community average degree and/or high community average connection radius 3.5. Thus we can conclude that despite the fact that household transmission rates are on average higher than community transmission rates in our model, community transmission leads to higher incidence levels (compared with household transmission) because incidence is affected by both the intensity of transmission and the number of contacts. Furthermore, in our model, individuals have on average many more community contacts than household contacts. These results are consistent with a recent mathematical modeling analysis that found that contact saturation (as what might occur within a household) and super-spreading (among individuals with many contacts) might lead to higher incidence levels [152].

Due to the fact that high incidence levels are driven by community transmission and that community CT does marginally better than HHCT when the community transmission rates are higher (3.6), it may be preferable to focus on community CT (instead of HHCT) in high incidence settings. On the other hand in low incidence settings, disease is very low in the community and screening community contacts would likely find fewer cases of active or latent TB. Therefore, HHCT tracing may be preferable in low incidence settings. Also consistent with standard practice in low- and high-incidence settings [9, 65], contact tracing interventions in our model tended to perform better in lower incidence settings (see Section 3.3.0.3 for results).

The combined intervention (i.e., both HHCT and community CT) resulted in negative interaction in nearly all settings, with greater negative interaction at lower incidence levels. In other words, there was no synergy between contact tracing interventions and actually, the effects of the HHCT and community CT interventions were lower when the they were combined. This is sensible because both interventions are likely to prevent many of the same cases e.g., via overlapping indirect protection [153, 154]. Overall, the combined intervention resulted in a greater reduction in TB risk on the population-level compared with any single intervention (Figure 3.7), but it conferred less protection than the sum of contact tracing interventions.

Limitations of our analytical methods include the wide range of assumptions required to generate a parsimonious and computationally efficient model. First, we assumed static network connections over the course of the entire simulation. Allowing dynamic network connections would cause infection to escape from local clusters more easily [147], therefore in low average degree and high incidence network types, we might over-estimate the effects of HHCT compared with the control scenario because it will be more efficient when infection is concentrated in clusters. On the other hand, our model covers a range of incidence levels and average degrees with the majority of runs having relatively low prevalence, rare infection events and a high average degree which should in general approximate a random community sample. We also assumed that number of individuals per household is fixed with 5 people per household. Although this is unrealistic, keeping the household size fixed allowed for us to examine how variation in community contact alone might impact the effectiveness of different screening interventions. Future work can examine how variations in household size might interact with different community contact structures to change transmission in our network. Modeling dynamics between multiple strains of tuberculosis are beyond the scope of our model and study questions therefore, we did not incorporate re-infection events among individuals with early latent TB since that would just prolong the time an individual spends in the early latent state [155]. Finally, our births reshuffling scheme based on [147] is not realistic, however, it allows for the population size to be constant and prevents our model from consistently placing susceptible individuals into transmission hotspots.

Future work may consider how to optimize screening interventions in different network connectivity scenarios. One recent simulation model employed adaptive approaches to find undiagnosed HIV cases and found that accounting for spatial correlation may modestly improve case finding [156]. Additionally, other networks may be formulated to consider how more heterogeneity in contact may also change the effectiveness of screening interventions, e.g. in networks with skewed community contact distributions.

With the exception of the combined intervention, in our analysis, the maximum number of individuals that are screened for each case is 4. This is consistent with the fact that screening interventions are more effective in lower incidence strata. We chose a limit of 4 such that community and household screening interventions would screen the same number of individuals to facilitate a direct comparison. Another potential future direction could include identifying the number of contacts needed to screen to obtain a target reduction in community incidence. Alternatively, it would be useful to explore how many individuals must be screened before the addition of screened individuals leads to diminishing returns. We also found that results did not differ substantially by reference group for HHCT and community CT, thus more work should be done to more precisely classify when HHCT or community CT should be used.

Overall, it is important to consider how spatial and contact heterogeneity can be leveraged to optimize screening interventions. Despite a large amount of variability within settings, screening interventions conferred reductions in TB incidence. We found that screening interventions performed better in settings with low incidence levels, low average degree and a low connection radius. Overall, our results suggest that prior knowledge about the community contact structure and endemic incidence level can help determine whether or not a specific contact tracing screening intervention will be effective.

CHAPTER IV

Examining the discrepancy between explosive norovirus outbreaks and relatively low attack rates in daycare and school settings

4.1 Abstract

Background: Norovirus outbreaks are notoriously explosive, with dramatic symptomology and rapid disease spread. Children are particularly vulnerable and drive norovirus transmission due to their high contact rates with each other and the environment. Despite the explosive nature of norovirus outbreaks, attack rates in school outbreaks remain low with the majority of students not reporting symptoms.
Methods: We next explore the biological and epidemiological mechanisms that may underlie epidemic norovirus transmission dynamics using a disease transmission model. We compared different model scenarios, including a partially immune population, stochastic extinction, and an individual exclusion intervention, and we calibrated our model to daycare and school outbreaks from national surveillance data.
Results: Including innate resistance and acquired immunity recreated the low attack rates observed in daycare and school outbreaks. Partial immunity alone

resulted in outbreaks that were substantially faster than what was observed. The addition of individual exclusion (to a partially immune population) extended outbreak durations by reducing the amount of time that symptomatic people contribute to transmission resulting in attack rates and outbreak durations more consistent with the surveillance data.

Conclusions: Incorporating both a partially immune population and individual exclusion is sufficient to recreate explosive norovirus dynamics, with more realistic outbreak durations (compared with immunity alone), and relatively low attack rates in school and daycare venues.

4.2 Introduction

Norovirus is the leading cause of acute gastroenteritis across all ages in the United States (US), with 19 to 21 million cases occurring per year [157]. The role of children in transmission has recently been highlighted in a mathematical modeling analysis [5], which found that pediatric vaccination would result in substantially higher protective population-level effects when compared with vaccination of the elderly. After the onset of a school outbreak, person-to-person transmission can propagate the disease to cause secondary cases in households and the broader community [158–165]. Understanding how norovirus spreads within venues can inform design of effective interventions to reduce overall population-level risk.

Norovirus transmission can occur directly through person-to-person contact [86] and indirectly through water [84], food [83], or fomite-mediated pathways [22, 85, 166, 167]. Symptomatic individuals efficiently spread virus through vomiting and defecation [22]. After symptoms resolve, individuals continue to shed for an average of $\sim$2 weeks [88]. Norovirus transmission is sustained through the combination of efficient and prolonged human shedding [22], and extended environmental persistence [91, 168–170]. Additionally, norovirus is highly infectious, with an infectious dose of 18-2800 virions (peak viral concentration per gram of stool reach levels of 10^9) being sufficient to cause infection [87, 171]. These features of transmission as well as a lack of long-lasting immunity in human hosts [89, 90], contribute to venue-level norovirus outbreaks potentially exhibiting rapid, explosive growth rates [21, 172]. These explosive epidemic growth rates would be expected to correspond to high attack rates with exhaustion of susceptibles, similar to other highly transmissible diseases like measles [173, 174]. However, despite this explosive tendency and the important role that children play in transmission, attack rates (ARs) in school and daycare outbreaks are relatively low ($\sim$15% to $\sim$20% in daycares and schools, respectively, based on data from the National Outbreak Reporting System (NORS) [26]).

There are multiple explanations for the combination of explosive outbreaks and low ARs observed in outbreak data. First, $\sim$20% of the US population lack a functional FUT2 gene, conferring innate resistance to some norovirus genotypes [175]. Furthermore, depending on age, up to $\sim$90% of children < 5 years of age have norovirus antibodies titers potentially indicating acquired immunity [99, 100], although the level of protection conferred by these of antibodies is not known [176]. Second, the Centers for Disease Control and Prevention (CDC) recommends various interventions to prevent and control norovirus outbreaks, including isolation of individuals during the symptomatic period [177] which may also reduce transmission [178]. Finally, stochastic extinction may lead to outbreaks ending without exhaustion of susceptibles, especially for smaller populations [179] e.g. daycares. Any combination of these factors may contribute to low ARs and rapid cessation of outbreaks within venues.

In this paper, we employ mathematical models to explore underlying mechanisms leading to disease transmission dynamics that can explain the observed norovirus epidemiology within daycare and school venues. Given the epidemiological features discussed above, explaining norovirus dynamics requires a detailed representation of the mechanisms driving transmission. Here, we examine which mechanisms are sufficient to explain the epidemiological patterns seen in outbreaks using a transmission model calibrated to CDC NORS surveillance data.

4.3 Methods

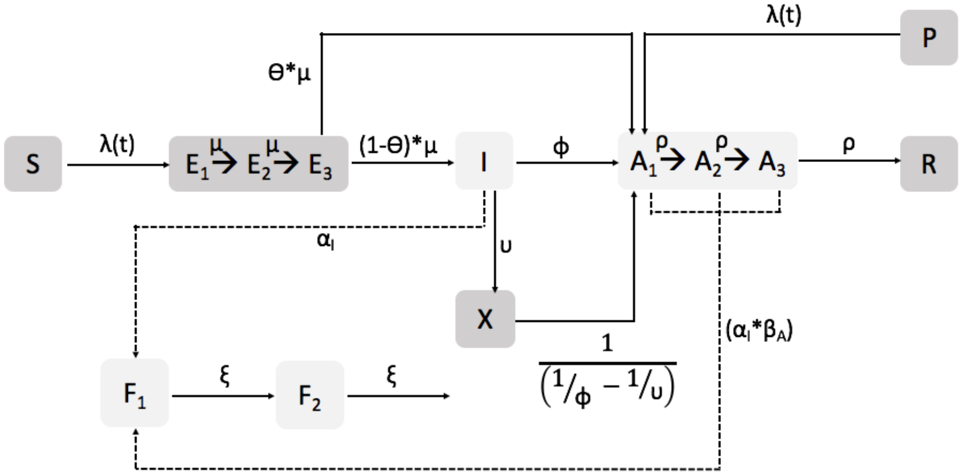
Our model (adapted from [5]) represents a school or daycare center that we calibrated to NORS outbreak surveillance data. We randomly sampled parameter values from realistic predefined ranges to account for uncertainty. For each model scenario (described below), we derived a distribution of parameter sets which best recreated the distribution of ARs, durations, and populations observed in NORS using sample-importance-resampling [180]. Sample-importance-resampling involves sampling parameter sets from a prior distribution (i.e., the predefined parameter ranges), calculating likelihoods to compare model outcomes from each parameter set to the NORS data, and resampling the parameter sets with replacement using the likelihoods as weights to create a posterior distribution. We ran each model scenario using the resampled parameter set distribution, and both graphically and quantitatively evaluated which models and underlying mechanisms best matched the observed NORS data, using Kullback-Leibler divergence (KL) [181].

4.4 NORS Dataset

We calibrated our model to outbreak duration (in days), student AR, and student population size data from NORS, a CDC-operated internet-based surveillance system through which state, territorial, and local health departments within the US can enter outbreak information [26, 182] (See Appendix Table C.2 for summary statistics of dataset).NORS collects various data including AR, duration, primary mode of transmission (e.g. foodborne or person-to-person), other etiological information, lab confirmation, secondary transmission data, health outcomes such as hospitalizations or deaths, sex and age distribution, and location. Our dataset includes all school and daycare outbreaks in NORS that occurred from 2009–2016 which indicated norovirus as the only suspected or confirmed etiology. We classified a given venue as daycare or school based on self-reported classification by the reporting agency. According to the NORS categorization, 'daycare' is the aggregate of both daycare and preschool and 'school' includes all other school-aged venues (i.e., elementary school, middle school, and high school) [183]. In total, there were 989 norovirus outbreaks in schools and 329 in daycares, which comprised 4.6% of all outbreaks reported through NORS during 2009–2016 (i.e., 1,318 of 28,580 total outbreaks across all modes and etiologies). We only included outbreaks with complete data (i.e., not missing for total students exposed, AR, and outbreak start and end dates), and with ARs and durations within the 5^{th} and 95^{th} percentiles of the dataset to remove outliers and to calibrate our model to data generally representative of common norovirus outbreaks.

4.4.1 Model Structure

We model transmission in a school or daycare venue, to capture the features of the NORS calibration data. All analyses were conducted in R version 3.2.4 [184].



$$\lambda(t) = [I + \beta_A(A_1 + A_2 + A_3)]\beta_{HH} + (F_1 + F_2)\beta_{FH}$$

Figure 4.1: Model schematic for a single venue. Our model is an extension of [5]. All compartments involved in the force of infection (Equation 4.1) are in light gray. The full force of infection equation is also shown in the figure above. Susceptible individuals (S) may be infected and pass through a latent period (E_1 to E_3) before becoming symptomatically (I) or asymptomatically infectious (A_1 to A_3). Social distancing or individual exclusion is represented by (X). During infection, individuals may shed pathogens onto environmental fomites (F_1). As pathogens on the fomites decay, they move to (F_2), which represents biphasic decay. Additionally, individuals become immune (R) following infection, and may have innate resistance (R) or acquired immunity and be partially immune (P) at the start of the outbreak. All parameter values are shown in Table 4.1.

4.4.1.1 Transmission Model for Daycare centers and Schools

Transmission occurs directly through person-to-person contact or indirectly through fomite-mediated pathways i.e., shedding and pickup of virions in the environment.

Individuals start as susceptible S, partially immune P, or fully recovered R depending on acquired immunity and innate resistance status (Figure 4.1). Susceptible and partially immune individuals become infected according to the force of infection $\lambda(t)$, which is based on: symptomatic individuals (I), asymptomatic individuals (A_1, A_2, and A_3), environmental fomite pathogen concentration (F_1 and F_2), and human-to-human (β_{HH}), and fomite-to-human (β_{FH}) transmission rates. Excluded individuals do not contribute to the force of infection:

$$\lambda(t) = [I + \beta_A(A_1 + A_2 + A_3)]\beta_{HH} + (F_1 + F_2)\beta_{FH}, \tag{4.1}$$

where β_A represents the reduction in efficiency of asymptomatic transmission compared with symptomatic transmission.

Infected individuals move through a gamma-distributed latent period (E_1, E_2, and E_3) and then either become symptomatic or asymptomatic [185]. The latent period is gamma distributed to represent the empirical distribution of the data [39]. The gamma-distributed asymptomatic period (A_1, A_2, and A_3) represents post-symptomatic shedding and exhibits reduction in shedding by stage (e.g., individuals in A_2 shed less than individuals in A_1). Partially-immune individuals can become infected, but not diseased (symptomatic). Symptomatic individuals may become excluded (X) for the remaining duration of their disease (e.g. if sent home from school/daycare) and do not contribute to transmission while they are excluded. All non-excluded symptomatic and asymptomatic individuals of a given age-group shed pathogen into the environment. Norovirus pathogen decay on fomites occurs in a biphasic pattern with an initial rapid die-off followed by a period of slower decay [91, 186]. Finally, all individuals who become infected eventually progress to the fully recovered. See Appendix Table C.1 for initial condition ranges, and Section C.1 for the model description and equations. Also, see Table 4.1 for parameter ranges.

Table 4.1: Transmission Model Parameter Values and Uncertainty Ranges

Parameter	Description	Estimate/Uncertainty Ranges (units)	Sources
μ	Rate of transition through each latent state	$\frac{1}{(\frac{1.15}{3})}$ $(days^{-1})$	Systematic review examining incubation periods for different types of viral gastroenteritis [187]
θ	Proportion of latent individuals that don't become symptomatic	0.3 (unitless)	Volunteer study [185]
ϕ	Transition rate from symptomatic compartment (I) to asymptomatic compartment (A_1)	$\frac{1}{1.25}$ $(days^{-1})$	Cohort study examining the natural history of calcivirus infection in the community [188]
ρ	Recovery rate	$\frac{1}{(\frac{15}{3})}$ $(days^{-1})$	Review examining norovirus shedding duration data [88]
α_I	Shedding rate for diseased $(I$, A, $X)$ individuals	0 to 10,000,000 ($\frac{pathogens}{day}$)	Study quantifying symptomatic and asymptomatic shedding in nursing home and hospital outbreaks [171]
σ	Rate of reduction in shedding	0.2 (unitless)	See above [171]
ξ	Biphasic decay rate of norovirus in the environment	$\frac{1}{14}$ to $\frac{1}{0.333}$ $(days^{-1})$	Lab-based study quantifying the persistence of murine norovirus on a variety of surfaces [91]
β_A	Reduction factor for asymptomatic shedding and transmission (compared with symptomatic individuals)	-4 to -0.09691; sampled in log space (unitless)	See above [171]
Transmission Rates			
β_{HH}	Human-to-human transmission rate	$\frac{1}{Population}$ to $\frac{70}{Population}$ (infection/time)	Approximation of R_0. Range from a review of norovirus mathematical models [189]
β_{FH}	Fomite-to-human transmission rate derived by multiplying a scaling factor [0,2] by β_{HH}	0 to $2\beta_{HH}$ (unitless)	Limited empirical data on fomite to human transmission therefore, we allowed for wide range of values which can increase or decrease rates relative to β_{HH}
Exclusion Parameters			
υ	Time spent in symptomatic compartment, (I), before becoming excluded (X)	$\frac{1}{(\frac{1}{24})}$ to 1 $(days^{-1})$	Individuals are symptomatic and mixing normally for between 1 to 24 hours before being excluded

4.4.2 Model Scenarios

We considered the following model scenarios to examine mechanisms that can recreate the features of norovirus transmission:

- **Baseline Model**: A fully susceptible population with no individual exclusion.

- **Immunity Model**: A partially immune population with no individual exclusion. Twenty percent of individuals have innate resistance, are assumed to be fully immune, and start in R [175]. Individuals with acquired immunity are assumed to be partially immune.

- **Individual Exclusion Model**: A fully susceptible population with individual exclusion. Excluded individuals (X) do not contribute to transmission.

- **Combined Model**: A partially immune population with individual exclusion.

See Appendix Section C.2 for more details about the model scenarios. All models were simulated in a daycare setting and separately, in a school setting. We randomly sampled starting population sizes from the distribution of exposed student populations in the NORS data. Each model was simulated deterministically and stochastically for all population sizes. Low numbers of susceptible individuals are more accurately represented by a stochastic model compared with a deterministic model, because factors like stochastic extinction might affect dynamics. Therefore, we calibrated using the stochastic model only and then compared the results to the deterministic model as a sensitivity analyses to explore whether stochasticity improved the calibration. To address other assumptions about our modeling framework, we conducted additional sensitivity analyses varying seeding of the initial outbreak using our best-calibrated model, and separately re-ran all models with staff added into the model (see Appendix Sections C.9 and C.11 for details).

4.4.3 Calibration

We ran the model with 10,000 randomly sampled parameter and initial condition sets (collectively denoted ‘parameter sets’) using Latin Hypercube Sampling (LHS) [115]. Each parameter set was run in a school setting and separately, in a daycare setting. The only distinction between the model setup for schools and daycares is the starting population size (which in each setting is taken from the NORS outbreak data in the corresponding setting). We then calibrated each venue–specific model separately to its corresponding NORS data using sample-importance-resampling [180]. The NORS data includes attack rates and populations for students and staff. Here, we considered two versions of the model and calibration—one model which included only students (using the student attack rate and population data), and a sensitivity analysis which included staff members and students separately in the model (in which we sampled a matrix of contact rates within and between the two groups; see Appendix Section C.11 for details). Only one overall outbreak duration was provided in the NORS data, which was used for calibration in both model versions. The full calibration results for the staff and student model were similar to the student-only model (see below and Appendix Section C.11), and so for simplicity we present only the student model.

For a given venue–specific model, parameter sets were excluded from calibration if, for the corresponding model run, the outbreak was ongoing when the simulation ended (60 days set according to the distribution of the NORS data). We note that all the outbreaks in the NORS data used for calibration had ended by this point (the maximum duration in NORS was 40 days for daycare and 32 days for school). We derived a kernel density estimate (KDE) of the 3-dimensional probability distribution of NORS student ARs, student population size, and outbreak durations. KDE estimates were computed using the KS package in R which was designed for kernal smoothing of multidimensional data [190, 191]. The likelihood estimate of a given parameter set was calculated by taking the NORS KDE value which corresponded to a given AR, population size, and outbreak duration from the model results. For each simulation, the AR was defined as the total number of symptomatic individuals divided by the total population. The outbreak duration was defined as the number of days from the first symptomatic incident case to the last symptomatic incident case (we rounded up if the number of incident cases was 0.5 in the deterministic model). Finally, we resampled the parameter sets 5,000 times with replacement using the likelihoods as weights to obtain a final array of parameter sets that, if used as inputs for the model, could most closely recreate the NORS data distribution. See Appendix Table C.2 for calibration ranges derived from NORS.

To determine the best-fitting model, we derived a KDE of the calibrated model results, and calculated the KL divergence to measure the difference between the each calibrated model and NORS KDEs [181]. We also examined pairwise scatter plots of ARs, outbreak durations, and population sizes for calibrated model runs and compared them to the NORS data to assess the model calibration graphically.

4.5 Results

4.5.1 NORS Data

Our final dataset (after removing incomplete data and outliers) consisted of 163 daycare outbreaks and 393 school outbreaks. The median student population, student AR, and outbreak duration for daycare outbreaks were 75 people (range: 7, 410), 21.6% (range: 4.6%, 69.2%), and 13 days (range: 2, 40 days), respectively. The median population, AR, and duration for school outbreaks were 420 people (range: 6, 6486), 15.3% (range: 4.6%, 68.4%) and 8 days (range: 1, 32 days), respectively. See Appendix Table C.2 for details.

4.5.2 Model Comparisons

Figure 4.2 shows the median attack rates and durations for the NORS data and each of the models. Among the mechanisms examined here, partial immunity (included in the immunity and combined models) was best able to recreate the relatively low attack rates consistent with NORS data. This is due to the fact that individuals who are partially immune may become infected, but not symptomatic. Additionally, even though infected partially immune individuals contribute to transmission, they are not detected as diseased and do not count in the overall attack rate. The individual exclusion model also generates some simulations with low attack rates (see Figure 4.2, Figure 4.3 and the scatter plots in the Appendix Figures C.8, however because the great majority of simulations have higher attack rates, the overall calibrated distribution of ARs tends to be higher than the partial immunity models.

In terms of model fit, all models yielded a wide range of attack rates and generally shorter durations than the NORS data. The individual exclusion model yielded slightly longer durations than the other models (i.e., 5 days (95% credible interval (CI): 2 to 21 days) and 5 days (95% CI: 0 to 26 days) for daycares and schools, respectively). In comparison to the individual exclusion model, the baseline model durations were shorter (i.e., 4 days (95% CI: 2 to 12 days) and 4 days (95% CI: 0 to 10 days) for daycares and schools, respectively). As discussed above, the immunity model best captured the low attack rates in NORS. Furthermore, examining the pairwise comparison between attack rates and populations, we can see that the immunity and combined models best match the NORS data. See Appendix Table C.4 for all durations and Figures C.2 and C.5 for populations plotted against attack rates. Likely because of the balance in fit of these different features, the combined and individual exclusion models performed best according to the KL divergence (see Table 4.2 for all KL divergences). However, because of the improved performance of the combined model in terms of attack rate and visual fit in the pairwise scatter plots comparing calibrated models to NORS data (Figure 4.3 and Appendix Figures C.8), we selected the combined model as the overall best-fit model.

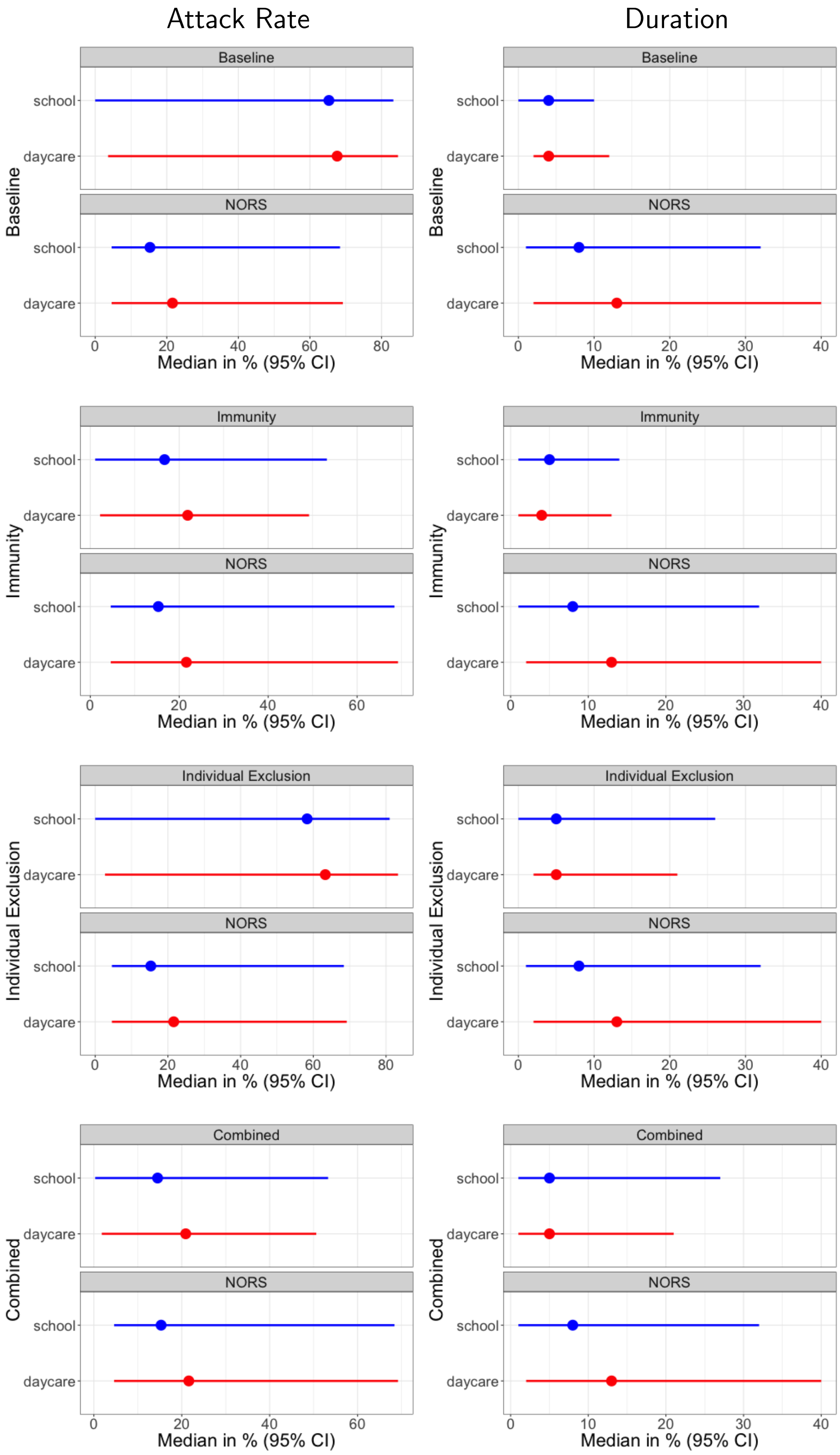


Figure 4.2: ARs (left column) and durations (right column) for each model compared with NORS data. Each plot corresponds to a different model structure (indicated on the y-axis label).

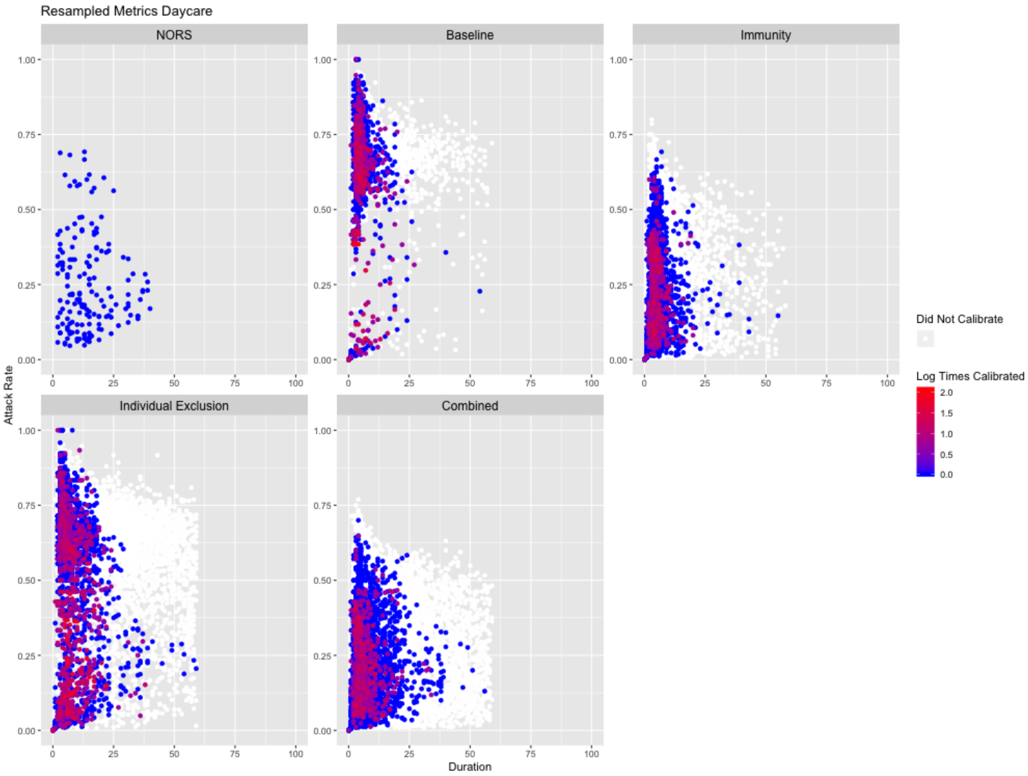


Figure 4.3: Attack rates vs. duration results from resampled parameter and initial conditions for the NORS daycare data. NORS data is shown in the top left. Points correspond to parameter sets and are colored by the amount of times they were resampled.

In the combined model, the daycare and school-aged ARs had medians of 20.9% (CI: 1.8% to 50.6%) and 14.5% (95% CI: 0.3% to 53.3%), respectively (shown in Figure 4.2 and Appendix Table C.3). These are slightly lower, but relatively consistent with the NORS ARs which were 21.6% (range: 4.6% to 69.2%) and 15.3% (range: 4.6% to 68.4%) for daycares and schools, respectively. The daycare and school-aged durations were 5 days (95% CI: 1 to 21) and 5 days (95% CI: 1.7 to 27), respectively. These are somewhat lower than the NORS outbreak duration medians which were 13 days (range: 2 to 40) and 8 days (range: 1 to 32) for daycares and schools, respectively.

Individual exclusion led to longer outbreak durations (see Appendix Table C.4) indicating that it may slow the spread of norovirus. This is due to the fact that excluded individuals transmit for a shorter period of time while they are symptomatic (i.e., between 1 hour and 1 day before being excluded) compared with non-excluded individuals (i.e., 1.25 days). This reduction in transmission time likely prevents the outbreak from spreading as fast as the baseline or immunity models. In the model, individual exclusion alone does provide slight protection from becoming infected or symptomatic (as indicated by the reduction in individual exclusion attack rates compared with the baseline model). Even though the reduction in transmission time can slow the spread of norovirus and prevent a limited number of individuals from becoming infected, attack rates are still high at $\sim$60% which corresponds to almost complete exhaustion of susceptibles (i.e., an attack rate of 70% because 30% of susceptible individuals become asymptomatic when they are infected).

For all models the median outbreak durations were less than the average time it took the first incident cases to fully recover (i.e., less than the ~ 16-day infectious period). Thus, outbreaks tended to end within a single infectious period indicating a rapid spread of disease within venues. Therefore, the combined and immunity models provided a mechanism for explosive outbreaks without the entire population becoming diseased. Importantly, all median outbreaks were substantially faster than what was observed in the NORS data. However, adding individual exclusion to the immunity model (i.e., the combined model) resulted in a more realistic distribution of outbreak durations closer to what was observed in the NORS data. The remaining analyses use the combined model, because this calibrated to the NORS data the best and replicated the low attack rates in the NORS data. See Figure 4.2 for forest plots, and Appendix Tables C.3 and C.4 attack rates and durations.

Model	Daycare	School
Stochastic Simulations		
Baseline	10.49	10.52
Immunity	7.82	9.86
Individual Exclusion	6.9	7.6
Combined	**6.45**	**8.22**
Sensitivity Analyses on Combined Model		
Seeding: Varying Pathogens in Environment	8.07	8.89
Seeding: Diseased Individual	5.98	8.32

Table 4.2: Kullback Leibler (KL) divergence for each model compared with the NORS data kernel density estimated distribution. Smaller KL divergence indicates a more similar distribution to NORS (i.e. less information difference between the NORS and the model distribution). The combined model is shown in bold. The sensitivity analyses include varying the environmental contamination at the start of the outbreak ('Seeding: Varying Pathogens in Environment'), and seeding with a single diseased individual ('Seeding: Diseased Individual'). Results for the deterministic models calibration are shown in the Appendix Table C.10.

4.5.3 Sensitivity Analyses

We conducted sensitivity analyses on the combined model to ensure that our results were robust to certain simplifying assumptions. We examined deterministic versions of the models, a model including staff transmission, and different seeding scenarios. Overall, the deterministic models fit worse than all corresponding stochastic models. The student and staff model had qualitative results consistent with the main analysis. Finally, when seeding with an infected individual (instead of in the environment), the original combined model had a lower KL divergence for schools, but had a higher KL divergence in the daycare model. See Appendix Sections C.9, and C.11 for more details, and Tables C.5 and C.6 for age-specific ARs and durations.

4.6 Conclusions

Our analyses suggest that a partially immune population is sufficient to recreate explosive norovirus outbreaks and relatively low ARs within venues. Furthermore, although individual exclusion resulted in only a modest reduction in the final number of symptomatic cases, it appears to slow transmission. Overall, the combined model was able to generate outbreaks with both low attack rates and a wider distribution of outbreak durations.

The combined model successfully recreated the general trends of the NORS data (i.e., relatively fast durations with low ARs) on the venue-level and more consistently generated runs with slightly longer durations and lower attack rates compared with the other models (see Figure 4.3 and Appendix Figure C.4). Although the initial LH sampled simulations of the combined model were able to recreate almost the entire joint distribution of ARs and durations observed in the NORS data, a large fraction of the initially LH sampled simulations had very low outbreak durations (and these low duration simulations were weighted highly in the sample-importance-resampling as they are also common in NORS). This skewed the calibrated distributions of durations in the combined model toward the lower end of the NORS data, so that the median model-generated durations were lower than those in NORS. Therefore, for a model to be able to calibrate well, it is required to both (1) recreate the joint distribution of attack rates and durations in the NORS data, and (2) evenly distribute model runs across the distribution. In general, stochasticity provided more variation in model results and therefore led to a distribution of model runs closer to the NORS data distribution when compared with deterministic model (see Appendix Section C.10 for more details). Although the combined model did not evenly distribute model runs across the NORS distribution, it did tend towards outbreaks with slightly longer durations (due to individual exclusion) and therefore calibrated better than the immunity model. We also conducted two sensitivity analyses varying seeding to explore how this might affect outbreak durations and found that the seeding with a single infected individual scenario performed approximately as well as the combined model in the main analysis. However, varying the number of pathogens starting in the environment fit worse than the combined model according the KL divergence (see Appendix Section C.9).

Additionally, although daycares generally had lower population sizes in the NORS data, they had longer durations. However, in all of our models, lower population sizes resulted in shorter outbreak durations. Overall in the NORS daycare data, there appears to be a trade-off between population size and outbreak duration. Specifically, lower population sizes have shorter durations and higher population sizes tend towards longer durations. See Appendix Figure C.3 for the joint distribution of daycare population sizes and outbreak durations in NORS. Importantly, the outbreak durations may have been misreported in NORS e.g., due to missing the detection of the first case. Therefore, we focused on recreating the general trends of the data.

Weaknesses of the calibration dataset include limitations of NORS reporting. Because NORS relies on passive surveillance, there is under-reporting of outbreaks which would likely result in the distribution of ARs and durations from reported outbreaks not being representative of reality [86]. Some outbreaks conducted interventions (e.g. decontamination), but we did not include this in our model, because the majority ($> 90\%$) did not report on this. The exposed population reported (i.e., the denominator of the ARs) may correspond to a single classroom, grade-level or entire school, and is not consistent across outbreaks (and the method of determining the exposed population is not specified in the data set). This may result in lower

ARs (if the exposed population is overestimated) and could have substantial effects on how well our models calibrate. In our initial exploration of the NORS data, we plotted ARs vs. duration while stratifying on exposed population size (for populations < 200) and found that the overall distribution appeared similar across strata (see Appendix Figure C.1). Furthermore, classifications of school or daycare venues relied on self-reporting. Many venues have a mixture of different age groups and therefore, classifications were likely not consistent.

Weaknesses of our model and analytical methods include the wide range of assumptions required to generate a parsimonious model. First, we did not explicitly include the processes of waning immunity and the existence of different norovirus strains because we are simulating a single outbreak. These processes are represented by varying the distribution of individuals who start as partially immune compared with fully susceptible. We are further assuming that partially immune individuals may become infected, but not diseased. In fact, reality is more complicated in that individuals with acquired immunity may become symptomatic or asymptomatic. However, altering the relative proportion of asymptomatic vs. symptomatic individuals i.e., by varying the initial conditions of those who start as partially immune should account for these effects. Another limitation of this model is its compartmental rather than individual-based structure, which does not include explicit contact network effects such as clustering and degree heterogeneity. One form of contact heterogeneity is the different transmission rates for staff and students—while our main model did not include staff transmission, the staff and student sensitivity analysis included these features and had results consistent with the main analysis i.e., immunity lowers attack rates and individual exclusion may make outbreak durations longer. For more details see Appendix Section C.11.

Results from our modeling analyses suggest that both immunity and individual exclusion are important for understanding outbreak dynamics in school and daycare venues. Future analyses should consider how interventions can leverage these factors e.g., using decontamination and vaccination interventions together to prevent outbreaks. For instance, if individual exclusion leads to slower transmission, other interventions may be able to work in conjunction with individual exclusion to stop transmission altogether. Furthermore, the importance of immunity in reducing outbreak attack rates reveals how useful a vaccine may be in preventing outbreaks.

4.7 Acknowledgements

We would like to thank Molly Steele Environmental Health Emory University, Zachary Marsh Centers for Disease Control and Prevention, and Dr. Lisa Prosser Department of Pediatrics and Communicable Diseases at the University of Michigan Medical School for their helpful comments, advice, and support. This Study was funded by the Joint Initiative for Vaccine Economics, Phase 5, a cooperative agreement between the University of Michigan and the Centers for Disease Control and Prevention (U01IP000965). The views expressed herein are solely those of the individual authors and do not necessarily reflect the views of the Centers for Disease Control and Prevention (CDC). This information or content and conclusions are those of the authors and should not be construed as the official position or policy of, nor should any endorsements be inferred by the CDC or the US Government.

CHAPTER V

Conclusion

Overall, this dissertation sought to understand what leads to seemingly counterintuitive findings in epidemiology. We used a variety of modeling tools to explicate causal mechanisms that could potentially represent the key drivers underlying disease transmission and progression.

In Aim 1, we simplified and extended a previously published mapping to convert between DAGs and CMs [32], and developed a workflow that can be used to simulate epidemiological studies and understand patterns seen in data. Ultimately, our method can be applied to nearly any study question or DAG (e.g., assuming there are no faithfulness violations on the DAGs [126]). We used our method to assess bias in study designs, and specifically explored what types of bias can lead to normal weight ever-smokers (not never-smokers) having higher mortality rates than their obese counterparts (i.e., the obesity paradox). Because we did not simulate a physiologically protective effect of obesity on mortality, we found that in general the obesity paradox occurs when covariates (i.e., age or COPD leading to reverse causation) are unadjusted in the statistical analysis. Interestingly our modeling framework enabled us to find that study design biases can interact and that adjusting for only one type of bias (e.g. confounding by age and not reverse causation) in a study that previously did not result in the obesity paradox, might cause the obesity paradox to occur. Other potential applications of our workflow include obtaining estimates for unmeasured variables on a DAG and relaxing key epidemiological assumptions e.g. no interference between units assumption [192].

In Aim 2, we leveraged spatial and contact heterogeneity to optimize TB screening interventions across a range of settings and levels of endemic incidence. Consistent with previous literature we found that transmission in higher incidence settings is driven by the community [15, 137]. This is due to the fact that incidence is affected by both the intensity of transmission and the number of contacts and individuals in our generated networks had substantially more community contacts than household contacts. Both HHCT and community CT interventions led to overall reductions in TB incidence. This is consistent with previously published randomized control trials [11] and analyses [70]. Screening interventions tended to perform better in settings with lower incidence and on networks with lower average degree and a lower average connection radius. Future analyses can examine whether or not other spatial features such as heterogeneity in community contact (e.g. very clustered networks with longer range connections) can lead to more effective contact tracing interventions. Other interventions may use adaptive approaches e.g., accounting for spatial correlation [156]. Overall, our results suggest that prior knowledge about the community contact structure and endemic incidence level can help determine whether or not a specific contact tracing screening intervention will be effective.

Finally in Aim 3, we proposed a combination of (realistic) mechanisms that might explain explosive daycare- and school-level norovirus outbreaks with relatively low attack rates. We found that partial immunity alone resulted in outbreaks that were faster than what was observed in surveillance data. The addition of individual exclusion reduced the amount of time that symptomatic people contributed to transmission and extended outbreak durations (making them more consistent with the surveillance data). Therefore, our combined model which incorporated both of these features, calibrated the best to the surveillance data. These results could be used to inform the design of effective interventions. For instance, future analyses could consider how interventions can leverage the fact that individual exclusion slows outbreak dynamics by simultaneously implementing other interventions (e.g. decontamination) to stop transmission.

Overall, Aims 1 and 3 used models as an explanatory tool which can be used to improve study design or intervention programs. Aim 2 used models to attempt understand drivers of transmission to improve disease control. Epidemiological findings are seemingly counterintuitive when there is a lack of understanding about what leads to observed patterns in data. This dissertation highlighted the importance of resolving this lack of understanding. Explicitly simulating the causal mechanisms underlying biological and epidemiological processes is a useful way to understand these patterns. This can, ultimately, provide insight into the design of sound studies and the optimization of interventions.

APPENDICES

APPENDIX A

Appendix for Chapter 2

A.1 Compartmental Model and Directed Acyclic Graph Comparison

See Table A.1 for a general comparison between Directed acyclic graphs (DAGs) and Compartmental models (CMs).

DAGs are non-parameterized causal diagrams used to graphically map causes and effects to aid in designing epidemiological studies. DAGs summarize the complete set of known relationships between variables relevant to a given study question [28, 30, 32]. A necessary precursor for a DAG to be considered causal is that all known common causes of any pair of variables on the graph must also appear [32]. Once relationships between variables are synthesized, a researcher can identify what must be measured and/or controlled for to eliminate confounding and selection bias [28, 109]. On DAGs, statistical associations between variables may be produced by (1) cause and effect (unbiased), (2) common causes (confounding bias), (3) common effects (collider bias) [109]. DAGs are therefore used to separate associations due to causality versus those due to bias. Figure A.1 shows illustrations of how associations are formed on a DAG. The work flow to structure and then determine which variables to adjust for on DAGs is described elsewhere e.g. [29, 30, 193, 194]. Once assumptions about the causal relationships between variables are made explicit and potential confounders and/or colliders are revealed, a study design and statistical analysis plan can be created such that an unbiased effect estimate of a given exposure on outcome can in principle be calculated.

Table A.1: DAGs vs. CMs

DAGs	CMs
Non-parameterized	Parameterized
Temporality represented	Temporality represented
Used to identify bias, gaps in knowledge and plan analyses in studies	Used to conduct in silico studies e.g., understand mechanisms, ask policy questions
Synthesize all *a priori* knowledge	Synthesize key knowledge, but strive for parsimony
Causal if all common causes are included	Depict flows over time i.e., simulate causal processes

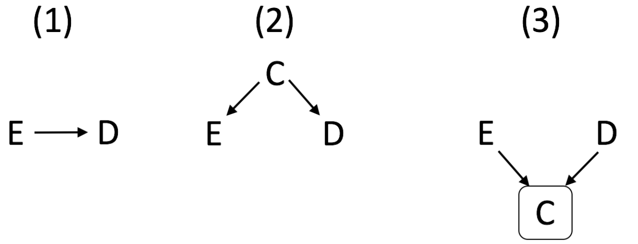


Figure A.1: The sources of association between variables become evident on a DAG. E is the exposure, D is the outcome or disease, and C is the covariate (1) cause and effect, (2) common causes or confounding, and (3) common effect or collider bias e.g. selection bias.

CMs simulate parameterized flows between disease states over time and are themselves a form of causal diagram [32, 33]. Specifically, CMs can be used to explicitly simulate mechanisms underlying disease transmission or disease progression and are often fit to population level data [34, 35]. Unlike causal DAGs which must in principle include all common causes, CMs often must balance realism with parsimony, and often only include the causal processes most relevant to the hypothesis [36].[1]

Once a model schematic is created, it may be converted into ordinary differential equations (ODEs) or simulated stochastically for smaller sample sizes. Data can then be integrated from a variety of sources as model inputs. For instance, data can inform the number of people which start in each disease state (initial conditions) or the transition rate parameter values. Furthermore, model outputs can be fit to a variety of data types [195], such as an epidemic curve or cancer incidence time-series[35, 196]. A fitted model can then be used to (1) estimate transition model parameters and initial conditions, (2) determine which parameters should be measured in future field studies, and (3) examine counterfactual scenarios when data collection is untenable due to ethical constraints or limited resources.

A.2 Model 1: Determining What to Adjust for on the Preston et al. DAG

To determine what to adjust for in a statistical analysis, we can refer to the structure of the DAG from the observational study (Figure II.2(a)).

Diabetes is a collider or a common effect of smoking status and BMI. The study is conditioned on diabetics (denoted by the box around diabetes on the DAG) since it will only be conducted among individuals with diabetes. Conditioning on a collider

[1] However, we note that DAGs technically also require parsimony in that (for example), all mediators between a given cause and effect are not included. Additionally, in practice DAGs often do not include all common causes, e.g. if it is not fully clear whether certain features are causally related.

creates a spurious association between its causes (in this case: smoking status and BMI) also called selection bias [109]. Additionally, diabetes is a mediator on the pathway from BMI to mortality. Conditioning on a mediator typically causes bias when there are unmeasured confounders i.e., between mediator and outcome or exposure and mediator. However, for simplicity, we will assume that there are no additional unmeasured confounders. Even though smoking confounds the association between mediator and outcome (and mediator and exposure), it is measured and we will adjust for it. Other issues related to conditioning on a mediator may arise due to exposure-mediator interaction i.e., if the effect of BMI on mortality is affected by diabetes status. This is addressed by the fact that we will only consider the controlled direct effect of BMI on mortality i.e., when the mediator value is held constant [110]. Next, smoking status is a common cause of BMI and mortality and therefore confounds their association. If we assume that there are no other sources of bias in the study, and no other common causes of the variables on the DAG, an unbiased effect estimate of BMI on mortality would require that we adjust for smoking status. For instance, we will estimate the effect of BMI on mortality in separate smoking strata to remove the spurious associations. We would therefore expect that examining the association between BMI and mortality in a population of diabetics among ever-smokers and then separately among never-smokers would remove the bias and the protective effect of obesity on mortality. However, this was not found to be the case in the Preston et al. study [2]. Therefore, either other biases exist and are not evident due to an inaccurate DAG or incorrectly categorized variables, or obesity truly is protective against mortality among ever-smoking diabetics.

The work flow to structure and then determine which variables to adjust for on DAGs is described elsewhere e.g. [29, 30, 193, 194].

A.3 Deriving a Corresponding CM from the Preston et al. DAG

The DAG of interest in this analysis is taken from Preston et al. [2], shown in Figure II.2(a). We initially operationalized the causal relationships between variables in the DAG from Preston et al. (Figure A.3) by creating a corresponding CM A.3, using the method of Ackley et al. [32]. See Appendix Equations A.1 for the ODEs of the full model.

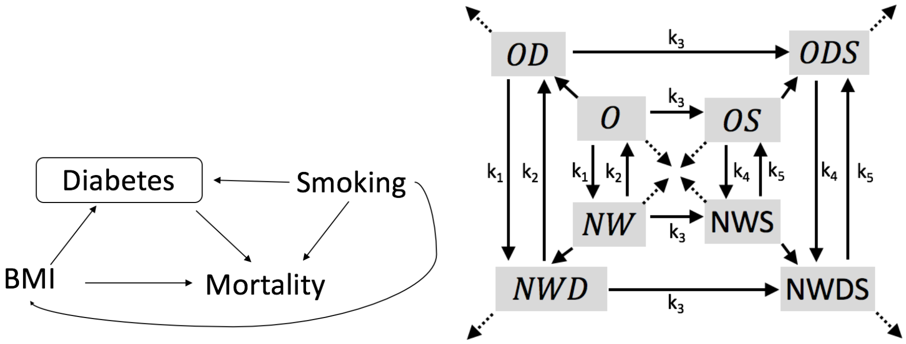


Figure A.2: (left) DAG representing the obesity paradox from Preston et al. [2] (right) CM: Schematic of the single age group compartmental model diagram corresponding to the DAG. NW represents normal weight individuals; O represents obesity; D represents diabetes, and S represents smoking. Individuals in any given compartment can die. Each arrow represents flows between states and rates that are equal to each other have the parameter. For instance, diabetes status does not affect the rate at which an individual transitions from obese to normal weight, therefore OD to NWD and O to NW have the same rate. We specify where transition rates are the same between compartments by labeling the model schematic accordingly and using the same parameter to represent equal rates in the equations. Mortality rates are denoted by dotted lines. Rates with no labels (including mortality rates) may all be distinct.

We enumerated disease states based on all possible combinations of random variables appearing in the DAG, i.e. there are 2^4 possible states, since we have 4 binary variables: diabetes, obesity, smoking, and mortality. Once individuals die, they cannot move between disease states and we no longer track them, therefore to reduce the dimensionality of our model, mortality is an outgoing flow from each compartment and was not included in the set of disease states. This reduced our model to 2^3 possible states. Next, we included all biologically plausible transitions between states. For instance, an individual can become an ever-smoker, but cannot return to being a never-smoker, also a diabetic individual cannot become non-diabetic.

Since the study population is conditioned on diabetics, we further simplified our model to only include diabetic compartments. This step reduced our model to 2^2 possible states. See Figure II.2(b) for the simplified model schematic and Appendix Equations A.2 for the ODEs.

A.4 Obesity Paradox Model 1 Full Equations

Below are the ODEs used to simulate the flows between disease states for the full model derived from the Preston et al. DAG. The model schematic is shown in Figure II.2(b). The model equations are given by:

$$\begin{aligned}
\dot{NW} &= -k_2NW + k_1O - k_{nw}NW - k_7NW - k_3NW \\
\dot{O} &= k_2NW - k_1O - k_oO - k_3O - k_8O \\
\dot{NWS} &= k_3NW - k_{nws}NWS - k_9NWS - k_5NWS + k_4OS \\
\dot{OS} &= k_3O + k_5NWS - k_4OS - k_6OS - k_{os}OS \\
\dot{NWD} &= k_7NW - k_2NWD + k_1OD - k_{nwd}NWD - k_3NWD \\
\dot{OD} &= k_8O - k_{od}OD - k_3OD + k_2NWD - k_1OD \\
\dot{NWDS} &= k_9NWS + k_3NWD - k_{nwds}NWDS - k_5NWDS + k_4ODS \\
\dot{ODS} &= k_6OS + k_3OD - k_{ods}ODS - k_4ODS + k_5NWDS
\end{aligned} \tag{A.1}$$

where NW, O, NWS, and OS are normal weight and obese non-diabetic never-smokers and normal weight and obese non-diabetic ever-smokers, respectively. The corresponding compartments for diabetics are NWD, OD, $NWDS$, and ODS. The mortality rates begin with a k and are labeled according to their corresponding compartment. For instance, k_{nwd} is the mortality rate for normal weight diabetic never-smokers. All other parameters are transition rates between disease states.

A.5 Simplified Model 1 and 2 Equations

Below are the ODEs used to simulate the flows between disease states for the simplified model that includes state transitions for diabetic individuals only. There is only one set of equations for Model 1 while each age group has its own set of equations, transition and mortality rates in Model 2. The model schematic is shown in Figure II.2(b).The model equations are given by:

$$\begin{aligned} N\dot{W}D &= -k_2NWD + k_3OD - k_{nwd}NWD - k_1NWD \\ \dot{OD} &= -k_{od}OD + k_2NWD - k_3OD - k_1OD \\ N\dot{W}DS &= k_1NWD - k_{nwds}NWDS - k_4NWDS + k_5ODS \\ O\dot{D}S &= -k_{ods}ODS - k_5ODS + k_4NWDS + k_1OD \end{aligned} \tag{A.2}$$

where NWD and OD are normal weight and obese diabetic never-smokers, respectively, and $NWDS$ and ODS are the corresponding normal weight and obese ever-smokers. The mortality rates begin with a k and are labeled according to their corresponding compartment. For instance, k_{nwd} is the mortality rate for normal weight diabetic never-smokers. All other parameters are transition rates between disease states. The initial state variables (initial conditions) are in units of people.

A.6 Model 2: Adding Age to the Original CM

Although age was not explicitly depicted on the original DAG (Figure II.2(a)), the analysis conducted by Preston et al. standardized mortality rates according to US census ages. Because age is a confounder in the relationship between the exposure, BMI and outcome, mortality, we should adjust for it in the statistical analysis to obtain an unbiased effect estimate. However, we initially ran the same statistical analysis as done for model 1 to see if not adjusting for age correctly could result in the obesity paradox. The purpose of this exercise is analogous to a sensitivity analysis in that, we investigate how unmeasured bias may have altered our study data and results.

To incorporate this into our study, we split our population into a younger age-group (ages 40-59) and an older age-group (ages 60-74) to explore how age-varying rates might lead to the obesity paradox. Among older adults (i.e., our study population), age affects obesity status [197], diabetes status [198], and mortality. Smoking initiation rates are quite low after age 40 i.e., $\sim$1% so we will assume that this rate is the same regardless of age-group [199]. See Figure II.2(c) for the Model 2 DAG. Alternatively, for an example of how to conduct more detailed age-weighting by individual ages, see Appendix A.8.

Apart from considering age, The model schematic (Figure II.2(b)) and model equations for Model 2 are the same as Model 1 (Equations A.2). To add in age-varying rates, we included one set of equations for the younger age group and one identical set of equations for the older age group. The equations are otherwise the same, but parameters and initial conditions vary between age-groups, which accounts for the effects of age. We assumed that everyone remains in their given age group over the course of the study (one year).

A.7 Age-Weighting for Age-Structured Models

We age weighted our model using weights from the 2010 census [116]. Specifically, the proportion of individuals in the young age group (ages 40-59) in the US population is 0.2771295 while the proportion of individuals in old age group (ages 60-74) is 0.1247997. Thus for a total study population of 1,000,000 individuals, there are

- $1,000,000\frac{0.2771295}{0.2771295+0.1247997} =$ **689,498.3** young individuals
- $1,000,000\frac{0.1247997}{0.2771295+0.1247997} =$ **310,501.7** old individuals

A.8 Alternative Age Weighting Example

Here, we show how to age-weight a model such that individual ages may be represented. Although this was not used in the analysis, certain datasets or study questions may necessitate a more detailed age weighting scheme.

To demonstrate how to age standardize a model using weights from the 2010 census [116], we use the following example. Let's assume that we run a simulation over time range 40-74 and that time represents age. Thus, time step 51 to 52 represents the distribution of all individuals who are age 51 by disease state over the course of the year. We will also take the model state variables to be fractions of the initial starting population.

First, divide each 5-year age group (from the census) by the total number of individuals in the census population to calculate age weights for the following groups ('40 to 44 years', '45 to 49 years', '50 to 54 years', '55 to 59 years', '60 to 64 years', ‘65 to 69 years’, ‘70 to 74 years’). We can assume that individuals are evenly distributed over each five year interval and so we can further divide each weight by 5 to get the age specific weight.

Next, we can calculate the proportion of the population in each state at a given age from the simulated dataset by dividing each value by the total fraction of the population alive at that age (i.e. summing over all disease states). For example, let’s say at time 51 the proportion of individuals in the normal weight diabetic never-smoking compartment is $\frac{0.02}{0.95}$ where 0.95 is the total proportion of the initial population still alive at age 51. Thus, 2.1% of people at age 51 are normal weight diabetic never-smokers.

Finally, we can multiply the age weights calculated from the census by the corresponding age proportions in the adjusted simulation data output and then calculate the person-time of the weighted dataset A.4 using the trapezoidal rule and the incident mortality A.5 by multiplying the mortality rates of the simulated dataset by the age weights.

A.9 Model 3: Adding Reverse Causation in the Original Model

We next tested the hypothesis that reverse causation may cause the obesity paradox. We first assumed that our observational study design was the same as in previous models and that we did not account for reverse causation in our data collection or statistical analyses (i.e., we ran the statistical analysis based on the original DAG in Figure II.2(a)). We made changes directly to the model to test this alternative underlying causal mechanism and then made a corresponding DAG. Again, since we assumed that our study design is the same, we did not change the statistical analysis. See Figures II.2(f) and II.2(e) for the new model and corresponding DAG, and see Appendix Equations A.3 for the model equations.

As mentioned previously, complications from comorbid diabetes and other diseases such as COPD may induce weight loss [117, 118] and also, increase the risk of mortality [119, 120]. We therefore extended Model 1 to simulate how undiagnosed COPD and associated complications may be a risk factor for mortality and also affect the exposure, BMI. Because these complications affect the outcome and exposure, this model incorporates reverse causation in that the higher risk of mortality may precede changes in the exposure. In our extended model, normal weight and obese ever-smokers can transition into COPD disease states, marked with a 'C'. We assumed that comorbidity of diabetes and COPD only occurs among ever-smokers since smoking is a key risk factor for COPD [200]. Additionally in our extended model, individuals with comorbid diabetes and COPD can then transition into the 'unhealthy' compartment, U. Individuals in U have lost weight due to cachexia and also have higher mortality rates than their normal weight 'healthy' counterparts (i.e. normal weight ever-smoking individuals with COPD who have not undergone cachexia). We assumed that BMI does not affect the rate at which individuals get COPD or transition into U. In some cases (depending on parameter values), individuals in U may also have higher mortality rates than obese ever-smoking individuals with COPD. Importantly, for the statistical analysis, 'unhealthy' individuals are measured as normal weight ever-smoking diabetics since our original study design did not measure COPD or the occurrence of cachexia. Individuals with diabetes are at an increased risk for developing COPD [201], so it is also possible that even if individuals with COPD at baseline in our cohort study were excluded, participants may have developed COPD and moved into the unhealthy disease state over the course of the study (this would be more likely for longer prospective studies more than 1 year). We examined this in two sensitivity analyses in which we exclude individuals with COPD at baseline and run the study for 1 year and also, 5 years.

Because this reverse causation mechanism relies on exposure status changing due to a risk factor for the outcome, mortality, the corresponding DAG is longitudinal to represent time-varying exposure and covariates. Specifically, this DAG incorporates changes to BMI over time. The exposure is BMI_0 the BMI measurement at baseline, and the outcome is cumulative mortality at the end of the study. More details for converting between longitudinal DAGs and CMs can be found in [32].

A.10 Reverse causation Model 3 Equations

Below are the ODEs used to simulate the flows between disease states for the reverse causation model. The model schematic is shown in Figure II.2(f). The model equations are given by:

$$\begin{aligned}
\dot{NWD} &= -k_4 NWD + k_5 OD - k_1 NWD - k_{nwd} NWD \\
\dot{OD} &= k_4 NWD - k_5 OD - k_1 OD - k_{od} OD \\
\dot{NWDS} &= k_1 NWD - k_6 NWDS + k_7 ODS - k_2 NWDS - k_{nwds} NWDS \\
\dot{ODS} &= k_1 OD + k_6 NWDS - k_7 ODS - k_2 ODS - k_{ods} ODS \\
\dot{ODSC} &= k_8 NWDSC - k_9 ODSC - k_3 ODSC + k_2 ODS - k_{odsc} ODSC \\
\dot{NWDSC} &= k_9 ODSC - k_8 NWDSC - k_3 NWDSC + k_2 NWDS - k_{nwdsc} NWDSC \\
\dot{U} &= k_3 NWDSC + k_3 ODSC - k_u U
\end{aligned} \tag{A.3}$$

where NWD represents normal weight diabetic individuals; OD represents obese diabetic individuals; ODS are obese diabetic ever-smokers; $NWDS$ are normal weight diabetic ever-smokers; $ODSC$ and $NWDSC$ are obese and normal weight ever-smoking diabetic individuals with COPD; and U are unhealthy individuals with comorbid COPD and diabetes who have undergone cachexia. These individuals have higher mortality rates than their unhealthy counterparts (i.e., $NWDSC$) and in some cases than $ODSC$ but are measured together with healthy normal weight individuals due to the design of our observational study. Finally, mortality rates are labeled according to their corresponding compartment. For instance, k_{nwd} is the mortality rate for normal weight diabetic never-smokers. All other parameters are transition rates between disease states. The initial state variables (initial conditions) are in units of people.

A.11 Combined Model

In the combined model, we incorporated both reverse causation and age-dependant mortality. Apart from considering age, the schematic (Figure II.2(f)) and ODE equations (A.3) are the same as Model 3, but are vectorized such that each age group has its own set of equations. See Appendix Figure II.2(g) for corresponding DAG which incorporates both reverse causation and age varying mortality. Because COPD prevalence increases with age [202], we allowed rates of transition to COPD to vary between age groups. Furthermore, because cachexia increases with age [203], we allowed rates of transition to U to vary between age groups. Finally, as reflected in Model 2, age affects BMI, diabetes and mortality. The MRR calculations are the same as conducted for the other models. We also ran various sensitivity analyses to see if adjusting for bias can keep the obesity paradox from occurring. Specifically, we (1) adjusted for age by standardizing to the unexposed population, (2) excluded individuals with COPD at baseline, and (3) combined 1 and 2.

A.12 Initial Condition Calculations

To determine how to distribute the population across disease states, the proportion of individuals in each state was randomly sampled using LHS [115]. For each model run, the sum of all sampled population fractions for the initial states must equal 1 to ensure uniformly sampled proportions. We randomly sampled proportions of the population in each disease state then multiplied the sampled proportions by the number of individuals in the given age group to get numbers of people starting in each disease state.

For instance in model 1 there are 4 states, we sampled 3 (total states - 1) values between $[0,1]$. Then we appended 0 and 1 onto the vector and sorted e.g., {0,0.1,0.4, 0.5, 1}, generating cut-points for the interval $[0,1]$, allowing the interval to be divided uniformly at random among the four states. Next, we took the lagged differences between the elements in the vector i.e., in this example {0.1,0.3,0.1,0.5} to get the start proportions for each state. This process avoided sampling in a specific order which would more frequently result in the last disease state have a lower proportion. We finally multiplied the proportions by census weights to determine how many individuals start in each state in this example: {40,192.92, 120,578.76, 40,192.92, 200,964.6}, totalling 401,929.2.

A.13 Mortality Rate Add-Ons

We imposed biologically realistic restrictions on the mortality rates such that ever-smokers have a higher mortality rate than their never-smoking counterparts (i.e., within weight strata), and obese individuals have a higher mortality rate than their normal weight counterparts i.e., within smoking strata).

Specifically, we set ODS mortality $\geq NWDS$ mortality $\geq NWD$ mortality and ODS mortality $\geq OD$ mortality $\geq NWD$ mortality. We did this by sampling a baseline mortality rate between 1% to 10% per year and then sampled 'add-on' mortality rates between 0% and 10% for obesity and separately smoking, thus the minimum mortality rate for obese, diabetic ever-smokers is 1% and the maximum is 30% per year.

A.13.1 Model 2: Age-varying mortality

All transition rates between disease states were allowed to vary by age group with the exception of smoking initiation. We set smoking initiation to be the same, because as mentioned, we assumed that the initiation rates were quite low anyway in these ages i.e., after 40 years of age [199]. Older age group mortality rates for a given disease state were determined by multiplying the younger age group mortality rate of the same state by a scaling factor between 1 and 2. We chose a maximum of 2 because it is a rough approximation of the relative mortality rates for the younger compared older age groups in the US according to the Centers for Disease Control and Prevention [204], although the age-groups are slightly different than in our model. Overall, within a given age-group the same restrictions on the relative mortality rates across disease states were used (as was used in Model 1). See Table 2.1 for all parameter ranges.

A.13.2 Model 3: Reverse Causation

We placed the same biologically plausible restrictions on the relative mortality rates as we did for Model 1 (i.e., a baseline mortality rate and add-ons for obesity, and smoking) and included an additional add-on for COPD related mortality. Finally, we derived the mortality rate in the U compartment by multiplying the mortality rate of normal weight diabetic healthy ever-smokers with COPD by a cachexia scaling factor between 1 and 2 (similar to the age scaling factor in Model 2). We chose a maximum of 2 since it was a relatively conservative estimate that corresponded with the age-varying mortality rate. This enabled us to directly compare causal mechanisms without making assumptions about the relative rates for age compared with cachexia associated mortality. Thus, the maximum mortality rate of normal weight diabetic unhealthy ever-smokers was equal to 80%. See Table 2.1 for all parameter ranges.

A.14 Person-Time for Simulated Studies

Even though the CM parameters were in units of years, the time steps for our model were in days. This is due to the fact that the sampling of both the initial conditions and parameter values could potentially lead to very fast, transient dynamics at the beginning of the simulation. For instance, if the transition rates out of the obese compartment are very fast and the initial conditions place the majority of individuals in the obese compartment, there will be a rapid decline in the numbers of obese individuals in the early stages of the simulation. This is not realistic especially in older age groups. Therefore, dividing our one-year time step into days enabled us to get a more precise estimate of the mean person-time spent in each compartment.

We calculated person-time by taking the daily average number of individuals in each disease state for the study. Specifically, we approximated the number of individuals in each state at each timestep using the life table method [205] (which is analogous to the trapezoidal rule) in which for a given timestep t the number of people in a state at time t and time $t+1$ is averaged. See Appendix Equation A.4 for details. We then added all person-days and converted to person-years by dividing the sum by 365 $\frac{days}{year}$ to get a final value in person-years. For slower dynamics or a longer run study, we could have calculated person years without averaging over each day.

A.15 Trapezoidal Rule for Person-Time Calculation

We calculated person time for a given time step, t, in simulations using the following equation i.e. the trapezoidal rule, equivalent to the life table method [205] in which all withdrawals or deaths are assumed to happen at the midpoint of each interval:
Equation for Person Time

$$PersonTime = Nr_t - \frac{Nr_t - Nr_{t+1}}{2} = \frac{Nr_{t+1} + Nr_t}{2}, \tag{A.4}$$

where 'Nr_t' is number of individuals in a given compartment at time t and 'Nr_{t+1}' is number of individuals in the same compartment at time $t+1$.

A.16 Incident Mortality for Simulated Studies

Next, we calculated incident mortality, the outcome, according to the following equation:

$$MD = M_{t+1} - M_t, \tag{A.5}$$

where MD is the incident mortality, M_t is cumulative number of deaths in a given compartment at time t and M_{t+1} is the cumulative number of deaths in the same compartment at time $t+1$. The cumulative number of deaths by compartment was quantified by adding extra equations with only the death rate for given compartment multiplied by the number of people in that compartment. We calculated the total incident mortality numbers for the entire year by each disease state. We next split our dataset into ever-smoking diabetics and never-smoking diabetics. We calculated a unique MRR for each strata. This accounts for the role of smoking as a confounder. See Appendix Section A.19 for an example simulated dataset for a single study.

A.17 Obesity Paradox Mortality Rate Parameterization

We can approximately back-calculate the mortality rate for a given CM compartment (i.e. the rate determined by LHS) by taking the total number of deaths over the course of simulation for that compartment (Equation A.5) divided by the person time approximation for that compartment (Equation A.4). For instance, in Table A.2, the mortality rate of normal weight never-smokers used in the model is just $\frac{25}{1500}$) = 0.017 deaths per-year. We can also calculate the mortality rate ratio of normal weight compared to obese never-smokers by hand from the simulated dataset. For instance, in Table A.2, we can just divide the mortality rate of normal weight individuals (i.e. $\frac{25}{1500}$) = 0.017 deaths per-year) by the mortality rate of obese individuals (i.e. $\frac{19}{988}$) = 0.019 deaths per-year) to get an MRR of 0.895. The final MRRs obtained from our analysis are simply the ratio of CM mortality rates.

Alternatively, if we want to parameterize a CM from real-world data, we can use MRRs by taking the exponentiated beta estimates from a Poisson model. For instance, if the MRR of normal weight, never-smoking diabetics compared to obese never-smoking diabetics is 1.5 we know that the ratio of mortality rates among these two compartments is 1.5 (e.g. they could be 0.3 to 0.2). Now, if our model among never-smoking diabetics is the same as the crude Poisson model equation above (See equation A.6), we can also take the exponentiated $\hat{\beta}_0$ to obtain the mortality rate among normal weight, never-smoking diabetics and then use the MRR to determine the mortality rate among obese never-smoking diabetics.

A.18 Poisson Model

Here, we show how to run a Poisson regression model. This was not used in our analysis because we didn't incorporate any type of sampling error into our model. However, one may want to simulate sampling using a multinomial draw in which case a Poisson regression model would be appropriate and could be used to derive confidence intervals to account for sampling error.

To run a standard Poisson regression model on our simulated dataset from Model 1 (see Table A.2 for example data), we can calculate mortality rate ratios representing the effect of normal weight compared to obese individuals on mortality:

$$\log(\hat{\mu}) = \log(T) + \hat{\beta}_0 + \hat{\beta}_1 NW \tag{A.6}$$

We ran this model for ever-smokers and then separately for never-smokers, where $\log(T)$ is $\log(person - time)$ and is also the offset term which accounts for unequal follow up times between compartments and allows us to model the rate. The outcome, $\log(\hat{\mu})$, is the estimated incident mortality rate and $\hat{\beta}_0$ is log(incident mortality rate) among obese diabetics (i.e. when normal weight (NW) is equal to 0). Finally, $e^{\hat{\beta}_1}$ is the mortality rate ratio comparing mortality among normal weight individuals to mortality among obese individuals. It is also the multiplicative effect on the mortality rate of being obese compared to being normal weight.

Note that it is also possible to individuate our simulated population by sampling according to a standard population e.g. the census and then to run a different type of regression model for count data, e.g. Cox proportional hazards, using individual (not compartment) level data. However, the Poisson regression model is simpler to implement and an appropriate choice for count data which is a commonly assumed in epidemiological studies.

A.19 Example Simulated Dataset

Table A.2 shows an example dataset generated from the CM output among never-smokers. For each characteristic, 'yes' is coded as 1 and 'no' is coded as 0. The first row represents individuals in the obese diabetic never-smoking compartment while the second row represents individuals in the normal weight diabetic never-smoking compartment. The MRR comparing normal weight never-smokers to their obese counterparts for this given parameter set is therefore $\frac{25/1500}{19/988}$.

Table A.2: Example Simulated Dataset Among Never-Smokers

Diabetic	Smoker	Obese	Normal Weight	Deaths	Person-Years
1	0	1	0	19	988
1	0	0	1	25	1500

APPENDIX B

Appendix for Chapter 3

B.1 Features of Generated Networks

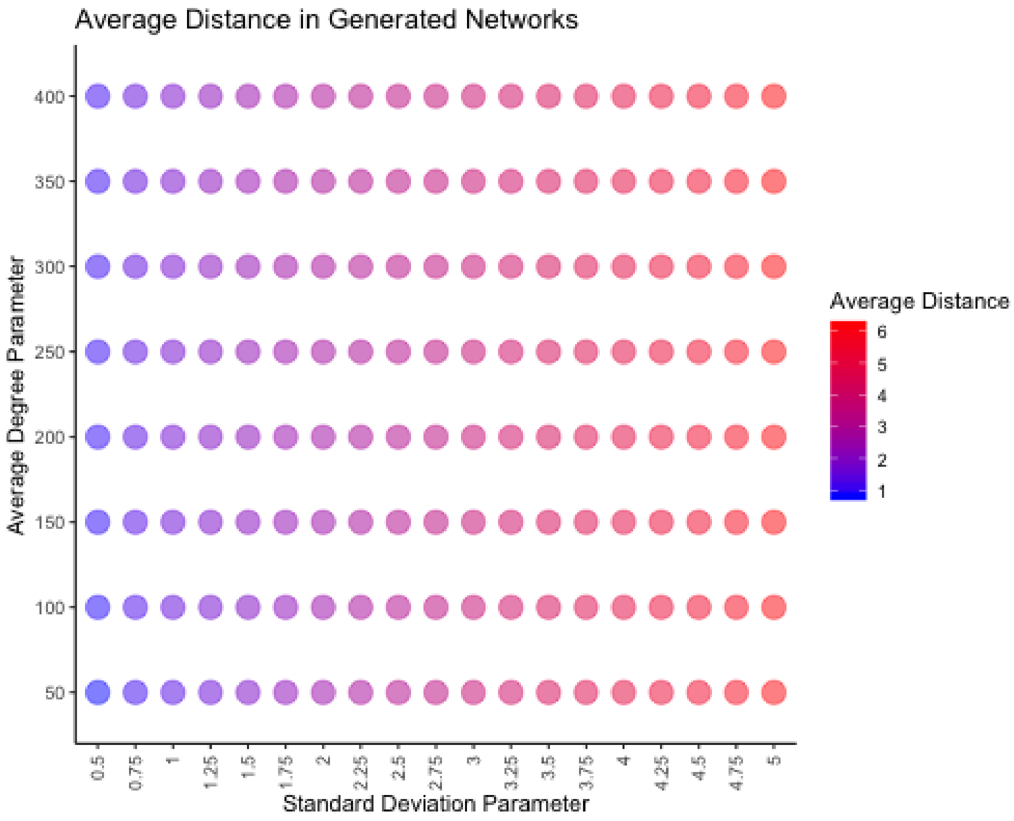


Figure B.1: Average Connection Radius: We generated a wide array of networks average connection radius (from 0.5 to 5) to examine variation in community contact e.g. long range connections vs. short range clustered connections. These plots demonstrate the parameters we used to specify the networks and the actual calculated metrics on the generated networks. For instance, specifying a high average degree with low average connection radius results in a network with a average connection radius.

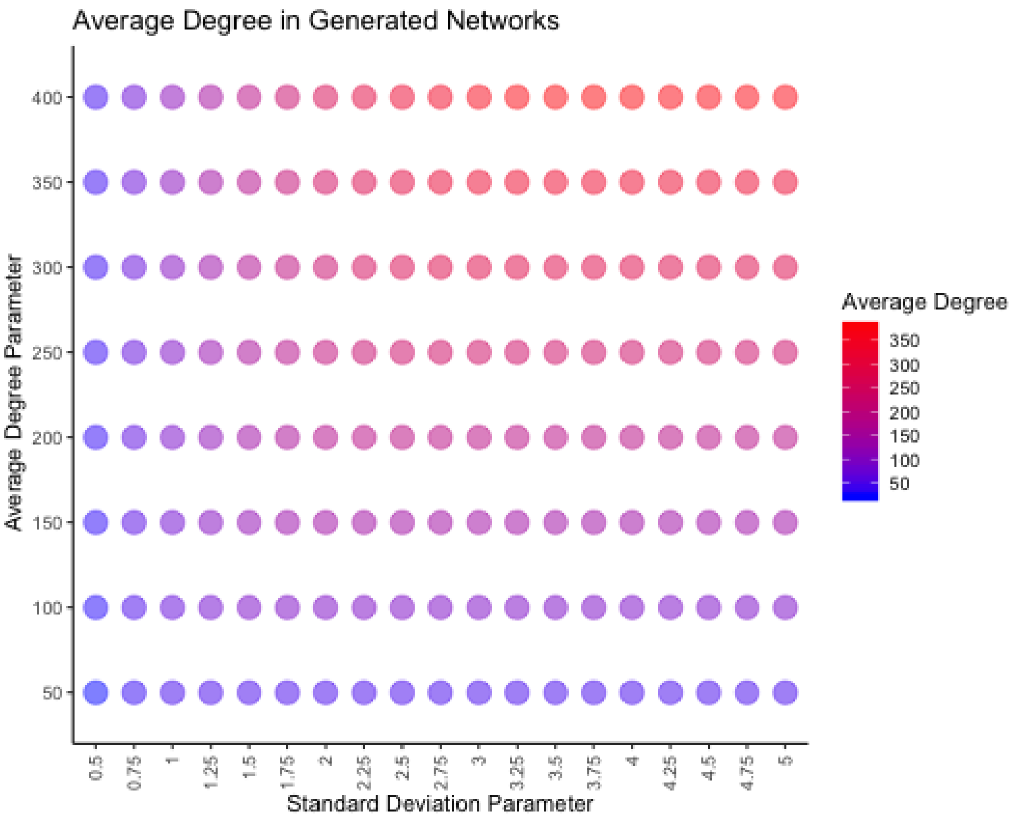


Figure B.2: Network Average Degree: We generated a wide array of networks varying average degree (from 50 to 450) to examine variation in community contact e.g. many community contacts vs. few community contacts. These plots demonstrate the parameters we used to specify the networks and the actual calculated metrics on the generated networks. For instance, specifying a high average degree with low average connection radius results in a network with a low average degree.

B.2 Distribution of Incidence Levels by Model Run

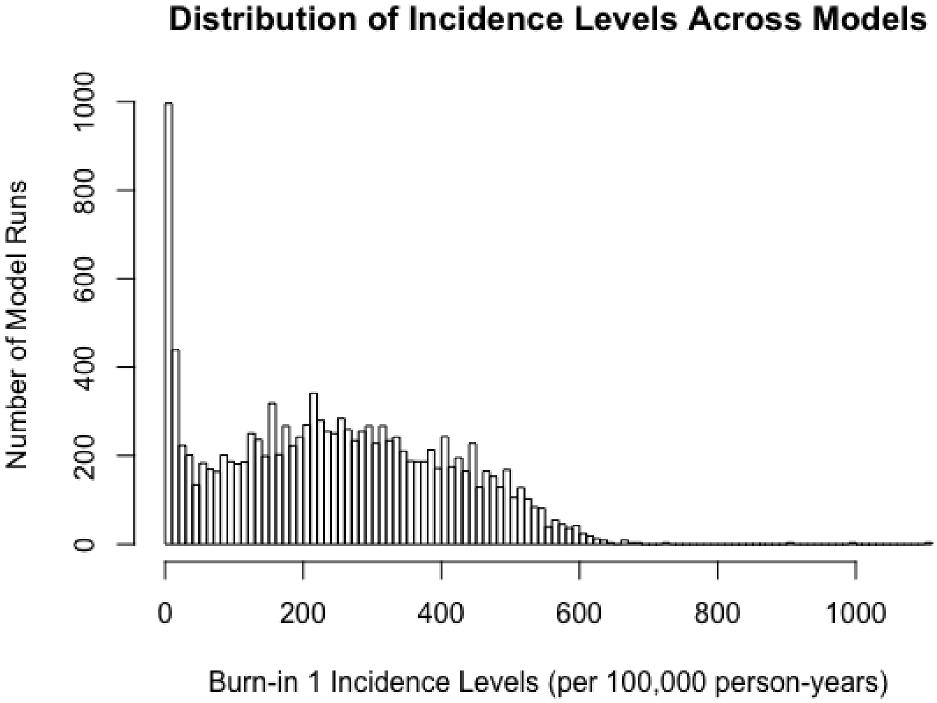


Figure B.3: Distribution of incidence levels across all model runs.

B.3 Protection Conferred by Single Interventions by Community Transmission Rate: Control Intervention as Reference Group

We fit splines to the relationship between β_{uC} and all RRs within the 2.5^{th} to 97.5^{th} percentiles, and plotted the results among different screening scenarios. See Figure B.4 for results with the control intervention as reference group.

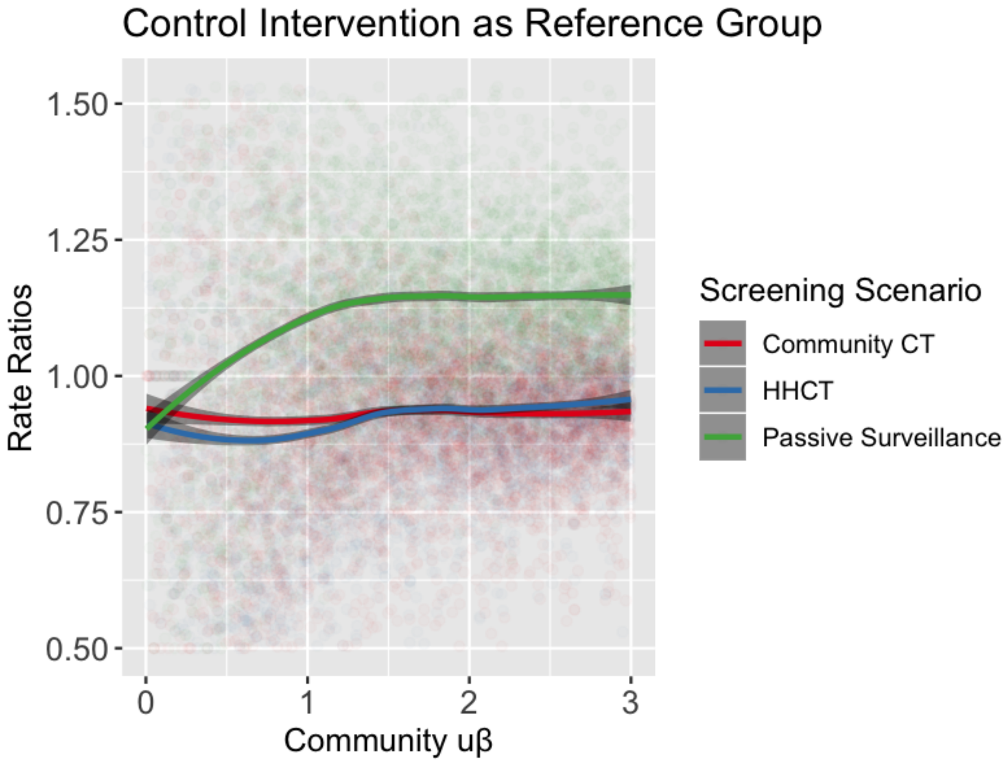


Figure B.4: Fitted splines representing relationship between the unscaled community transmission rate (β_{uC}) and RR (among RRs within the 2.5^{th} to 97.5^{th} percentiles). Passive surveillance only is the reference group. Dots represent individual model runs colored by screening scenario and lines are the splines (with 95% confidence intervals in shaded regions) which were calculated using the loess method in R [4].

B.4 Protection Conferred by Single Interventions by Network and Incidence Strata: Control Intervention as Reference Group

We calculated the mean RRs (among RRs within the 2.5^{th} to 97.5^{th} percentiles) and standard deviation (SD) within strata of network parameters and burn-in 1 incidence. For each network parameter or incidence level strata, we defined 'high' as >median across all model runs and 'low' as <median across all model runs. See

Table B.1 for performance of screening interventions with the control intervention as a reference group ranked in order of most effective scenario to least effective scenario.

Table B.1: RRs of Screening Interventions–Control Intervention as Reference Group

Screening Intervention	Connection Radius Strata	Average Degree Strata	Incidence Strata	Mean RR (SD)
Community CT	low	low	low	0.86 (0.2)
HHCT	low	low	low	0.87 (0.2)
Community CT	low	low	high	0.91 (0.12)
HHCT	high	high	low	0.91 (0.19)
HHCT	low	high	low	0.92 (0.19)
Community CT	high	low	low	0.92 (0.21)
HHCT	high	low	low	0.92 (0.18)
Community CT	high	low	high	0.93 (0.11)
HHCT	low	low	high	0.93 (0.11)
HHCT	low	high	high	0.95 (0.1)
Community CT	high	high	low	0.95 (0.18)
Community CT	low	high	high	0.95 (0.1)
HHCT	high	high	high	0.96 (0.09)
Community CT	low	high	low	0.96 (0.2)
HHCT	high	low	high	0.97 (0.1)
Community CT	high	high	high	0.98 (0.08)
Passive Surveillance	high	high	low	1.07 (0.2)
Passive Surveillance	high	low	low	1.09 (0.21)
Passive Surveillance	low	high	low	1.1 (0.21)
Passive Surveillance	low	low	low	1.12 (0.22)
Passive Surveillance	high	high	high	1.13 (0.11)
Passive Surveillance	high	low	high	1.14 (0.12)
Passive Surveillance	low	high	high	1.14 (0.11)
Passive Surveillance	low	low	high	1.16 (0.13)

Contact tracing interventions performed better than the passive surveillance only intervention. Furthermore, screening interventions performed better in lower incidence, lower average degree, and lower connection radii strata however, there was substantial variability.

The mean incidence was 174.4 (SD: 116.7) per 100,000 person-years among the top third (of 24 total) most effective screening scenarios and 245 (SD: 143.1) per 100,000 person-years among the bottom third. The mean community average degree was 126.3 (SD: 79.5) among the top third and 169.3 (SD: 91.6) among the bottom third. Finally, the trend was less apparent among connection radii in which the mean radius was 3.5 (SD: 1.5) and 3.7 (SD: 1.4) for the top and bottom thirds, respectively.

When examining results only among HHCT and community CT, the mean of all RRs in the high incidence strata is 0.95 (SD: 0.1), while it is only 0.9 (SD: 0.2) in the low incidence strata. Next, the mean of all RRs in the high average degree strata is 0.95 (SD: 0.15), while it is only 0.9 (SD: 0.17) in the low average degree strata. Finally, the mean of all RRs in the high average connection radius strata is 0.95 (SD: 0.14), while it is only 0.91 (SD: 0.17) in the low average degree strata. Therefore, within strata of network parameters, lower network standard deviation and lower average degree results in slightly better performance of screening scenarios. Again, there is substantial variability within strata.

B.5 Protection Conferred by Single and Combined Interventions by Community Transmission Rate: Control Intervention as Reference Group

We fit splines to the relationship between β_{uC} and all RRs within the 2.5^{th} to 97.5^{th} percentiles, and plotted the results among different screening scenarios. See Figure

B.5 for results with the control intervention as reference group.

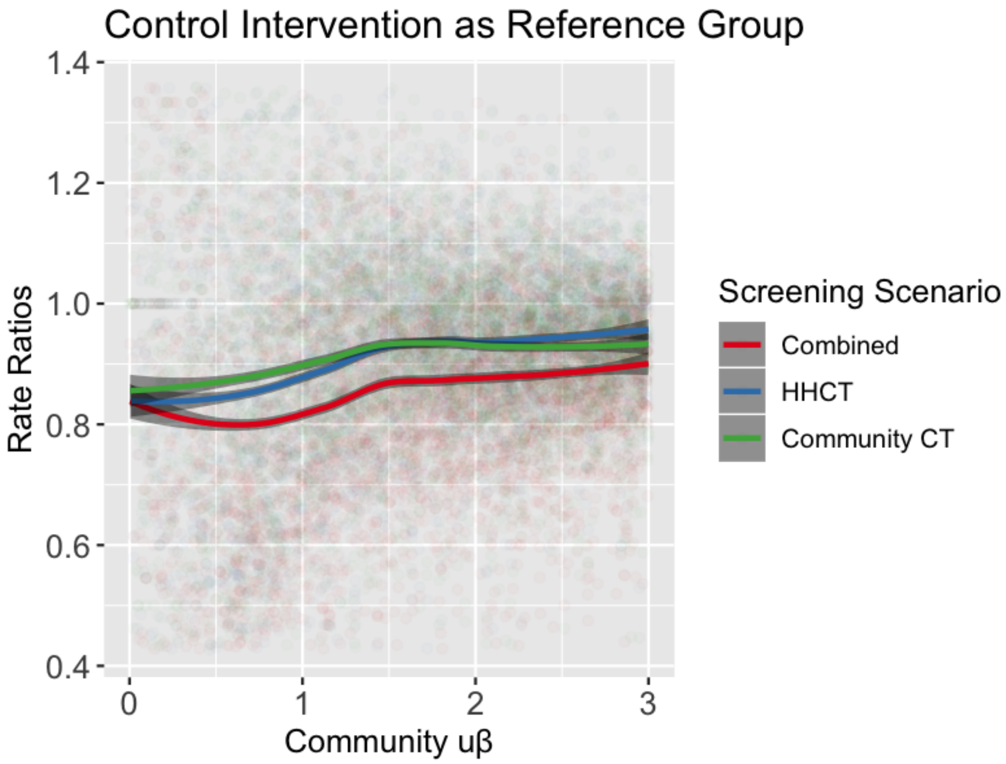


Figure B.5: Fitted splines representing relationship between the unscaled community transmission rate (β_{uC}) and RR (among RRs within the 2.5^{th} to 97.5^{th} percentiles). Passive surveillance only is the reference group. Dots represent individual model runs colored by screening scenario and lines are the splines (with 95% confidence intervals in shaded regions) which were calculated using the loess method in R [4].

Table B.2 shows the RERIs using the control intervention as a reference group.

There appears to be greater negative interactive effects in strata of low incidence. However, we do not see any trends in average degree or average connection radius.

Table B.2: RERIs of Combined Screening Interventions–Control Intervention as Reference Group

Connection Radius Strata	Average Degree Strata	Incidence Strata	Mean RERI (SD)
high	high	low	-0.08 (0.44)
low	low	low	-0.04 (0.51)
high	low	low	-0.03 (0.49)
low	high	high	-0.02 (0.24)
high	high	high	-0.01 (0.21)
low	low	high	0 (0.28)
low	high	low	0 (0.49)
high	low	high	0.01 (0.25)

APPENDIX C

Appendix for Chapter 4

C.1 Transmission Model for Daycare centers and Schools

Transmission occurs either directly through person-to-person contact or indirectly through fomite-mediated pathways i.e., shedding and pickup of virus in shared environments. Individuals start as either susceptible S, partially immune P, or fully recovered R depending on acquired immunity and innate resistance status (Figure 4.1). Susceptible and partially immune individuals become infected according to the force of infection $\lambda(t)$, which is based on: (1) the number of symptomatic (I), and asymptomatic (A_1, A_2, and A_3); (2) the pathogen concentration on fomites in the environment (F_1 and F_2); and (3) the human to human and fomite to human transmission rates (β_A, β_{HH}, and β_{FH}).

Force of Infection

$$\lambda(t) = [I + \beta_A(A_1 + A_2 + A_3)]\beta_{HH} + (F_1 + F_2)\beta_{FH} \tag{C.1}$$

where I is the number of symptomatic individuals, and A_1, A_2, and A_3, are the numbers of asymptomatic individuals, and F_1 and F_2 are the numbers of pathogens on contaminated fomites in the environment. The asymptomatic transmission reduction factor, β_A, is a reduction in efficiency of transmission compared with symptomatic individuals. The human to human and fomite to human transmission matrices are β_{HH} and β_{FH}, respectively. Excluded individuals, X, do not contribute to transmission.

Once infected, individuals pass through a gamma distributed latent period i.e. E_1, E_2, and E_3. It is gamma distributed to represent the empirical distribution of the data [39]. After they pass through the latent period, they become symptomatic. After an individual is symptomatic, they pass through a gamma distributed asymptomatic period i.e., A_1, A_2, and A_3 that represents post-symptomatic shedding and exhibits a reduction in shedding by stage (e.g., individuals in A_2 shed less than individual in A_1, see below for details). A proportion of infected individuals who originate in S do not become symptomatic and pass directly from E_3 to A_1. Individuals who start as partially immune can become infected, but do not become symptomatic and move directly to A_1.

A proportion of symptomatic individuals become excluded and move into the X compartment. After their symptoms resolve, they move to A_1 and return to the general population with the normal transmission and shedding rates for the A_1 compartment. Finally, all individuals who become infected eventually progress to the fully recovered state. All symptomatic and asymptomatic individuals (unless excluded) shed pathogen into the environment as follows:

$$\begin{aligned} Shedding &= \alpha_I I + \alpha_I \beta_A (A_1 e^{-\sigma(\frac{1}{\phi})} + A_2 e^{-\sigma(\frac{1}{\phi}+\frac{1}{\rho})} + A_3 e^{-\sigma(\frac{1}{\phi}+\frac{1}{\rho}+\frac{1}{\rho})}) \\ \dot{F_1} &= Shedding - \xi F_1 \\ \dot{F_2} &= \xi F_1 - \xi F_2 \end{aligned} \tag{C.2}$$

where α_I is the shedding rate for symptomatic individuals, and the reduction factors for shedding among asymptomatic individuals is β_A.

The amount of shedding is reduced exponentially as individuals progress across the gamma distributed asymptomatic period by σ for each state transition [171]. The symptomatic period, ϕ, and the recovery rate, ρ (i.e., from A_1 to A_2 etc.), account for the length of time that individuals shed at certain rates.

Viral concentration on fomites is tracked in the venue. Norovirus pathogen decay on fomites occurs in a biphasic pattern with an initial rapid rate of die-off followed by a period of slower die-off [91, 186]. Since are simulating a single outbreak, waning immunity is ignored. See Appendix Section C.3 for the full model equations, Table C.1 for initial condition ranges, and see Table 4.1 for parameter ranges.

C.2 Model Features

We incorporated different model features to examine mechanisms that can recreate the explosive outbreaks and low ARs characteristic of norovirus. We considered the following models (see Figure 4.1 for reference):

- **Baseline Model**: In this scenario, we simulated a fully susceptible population, with no individual exclusion. All individuals started in the susceptible, S, compartment.

- **Immunity Model**: In this scenario, we simulated a partially immune population with no individual exclusion. Because there is strain-dependent variation in the amount of protection innate resistance provides [175], and due to the fact that there is not an established correlate of norovirus protection that would be able to quantify acquired partial immunity, we examined different proportions of immunity. We assumed that those with innate resistance could not become infected at all and started as fully immune (in the R compartment), while those with acquired immunity started as partially immune (in P). Individuals in P could become infected, but not diseased. Non-diseased individuals were assumed to not be detectable during norovirus outbreaks and therefore were not counted in the numerator of the attack rate. Twenty percent of the population started in the R compartment i.e., with innate resistance [175], and we varied the total number with acquired immunity (P). We chose to vary the percentage starting with acquired immunity, because there is not a well established correlate of protection [206]. Finally, we calibrated the proportion of individuals with acquired immunity to the data by sweeping over a broad range of Latin Hypercube Sampling (LHS) [115] values (Table 4.1).

- **Individual Exclusion Model**: In this scenario, we simulated a fully susceptible population (i.e., all individuals started in the S compartment) with individual exclusion. During the simulation, a proportion of diseased individuals were removed from normal mixing and shedding, i.e., excluded. Excluded individuals do not contribute to transmission.

- **Combined Model**: In this scenario, we simulated a partially immune population, with individual exclusion.

Each of the above approaches were simulated both stochastically and deterministically. The stochastic simulation is a tau leaping version of the model [78] based on the Gillespie algorithm in which the stochastic model is approximate, but more efficient and updates at each large predefined time step (the time interval is denoted τ). We then ran the model 3 times using different random number generator seeds for each parameter set and population size to account for stochastic variation.

C.3 Model equations

Force of Infection

$$\begin{aligned} N_{inf} &= I + \beta_A(A_1 + A_2 + A_3) \\ \lambda &= N_{inf}\beta_{HH} + (F_1 + F_2)\beta_{FH} \end{aligned} \tag{C.3}$$

Human Transmission Model

$$\begin{aligned} \dot{S} &= -\lambda S \\ \dot{E}_1 &= \lambda S - \mu E_1 \\ \dot{E}_2 &= \mu E_1 - \mu E_2 \\ \dot{E}_3 &= \mu E_2 - \theta\mu E_3 - (1-\theta)\mu E_3 \\ \dot{I} &= (1-\theta)\mu E_3 - \phi I - \upsilon I \\ \dot{X} &= \upsilon I - \frac{1}{\frac{1}{\phi} - \frac{1}{\upsilon}} X \\ \dot{A}_1 &= \phi I - \rho A_1 + \lambda P + \theta\mu E_3 + \frac{1}{\frac{1}{\phi} - \frac{1}{\upsilon}} X \\ \dot{A}_2 &= \rho A_1 - \rho A_2 \\ \dot{A}_3 &= \rho A_2 - \rho A_3 \\ \dot{P} &= -\lambda P \\ \dot{R} &= \rho A_3 \end{aligned} \tag{C.4}$$

$$\begin{aligned} Shedding &= \alpha_I I + \alpha_I \beta_A (A_1 e^{-\sigma(\frac{1}{\phi})} + A_2 e^{-\sigma(\frac{1}{\phi}+\frac{1}{\rho})} + A_3 e^{-\sigma(\frac{1}{\phi}+\frac{1}{\rho}+\frac{1}{\rho})}) \\ \dot{F_1} &= Shedding - \xi F_1 \\ \dot{F_2} &= \xi F_1 - \xi F_2 \end{aligned} \tag{C.5}$$

C.4 Initial Conditions

In the baseline model, we start all individuals as Susceptible (S). On the other hand, in the immunity and combined models, 20% of individuals start as fully recovered (R) and some proportion of individuals start with partial immunity (P). This proportion is randomly sample between 0 and 80%. Finally, 10 million pathogens start in the F_1 compartment to initiate the outbreak. However, this number was varied from 0 to 100 million in a sensitivity analysis. Another sensitivity analysis seeded the outbreak with a single infectious individual.

Table C.1: Initial Condition Values and Uncertainty Ranges

State	Description	Value
S	Susceptible	Total population randomly sampled from NORS data – (Number Partially Immune + Number with innate resistance).
E_1 to E_3	Exposed	0 (people)
I	Symptomatic	0 (people) in main analysis and 1 person in the sensitivity analysis seeding with an infectious individual
A_1 to A_3	Asymptomatic	0 (people)
R	Recovered	0 (people) in the baseline scenario and 20% of individuals (representing innate resistance) in the immunity and combined models [175, 207]
P	Partially Immune	0 (people) in the baseline scenario and we the prevalence of acquired immunity from 0 to 80% in the immunity and combined scenarios (people) [99, 175, 207]
X	Excluded	0 (people)
F_1 and F_2	Contaminated Fomite Tracking Compartments	10 million pathogens on F_1 in the main analysis and we varied this from 0 pathogens to 100 million pathogens in the sensitivity analysis examining different initial seeding

C.5 NORS Calibration Ranges

We calibrated our models to the NORS data below. In total, there were 228 daycare outbreaks and 686 school outbreaks.

Table C.2: Calibration Ranges from NORS Data

Metric	Median (5^{th} to 95^{th} percentiles) [Mean]
Population sizes of daycare venue	75 people (7, 410) [94.3]
Population sizes of school venue	420 people (6, 6486) [447.3]
Attack rate within daycare venue	21.6% (4.6%, 69.2%) [25.5%]
Attack rate within school venues	15.3% (4.6%, 68.4%) [20.4%]
Outbreak duration within daycare venue	13 days (2, 40) [14.7]
Outbreak duration within school venues	8 days (1, 32) [10.8]

C.6 Attack Rates vs. Outbreak Duration Stratified by NORS Population Sizes

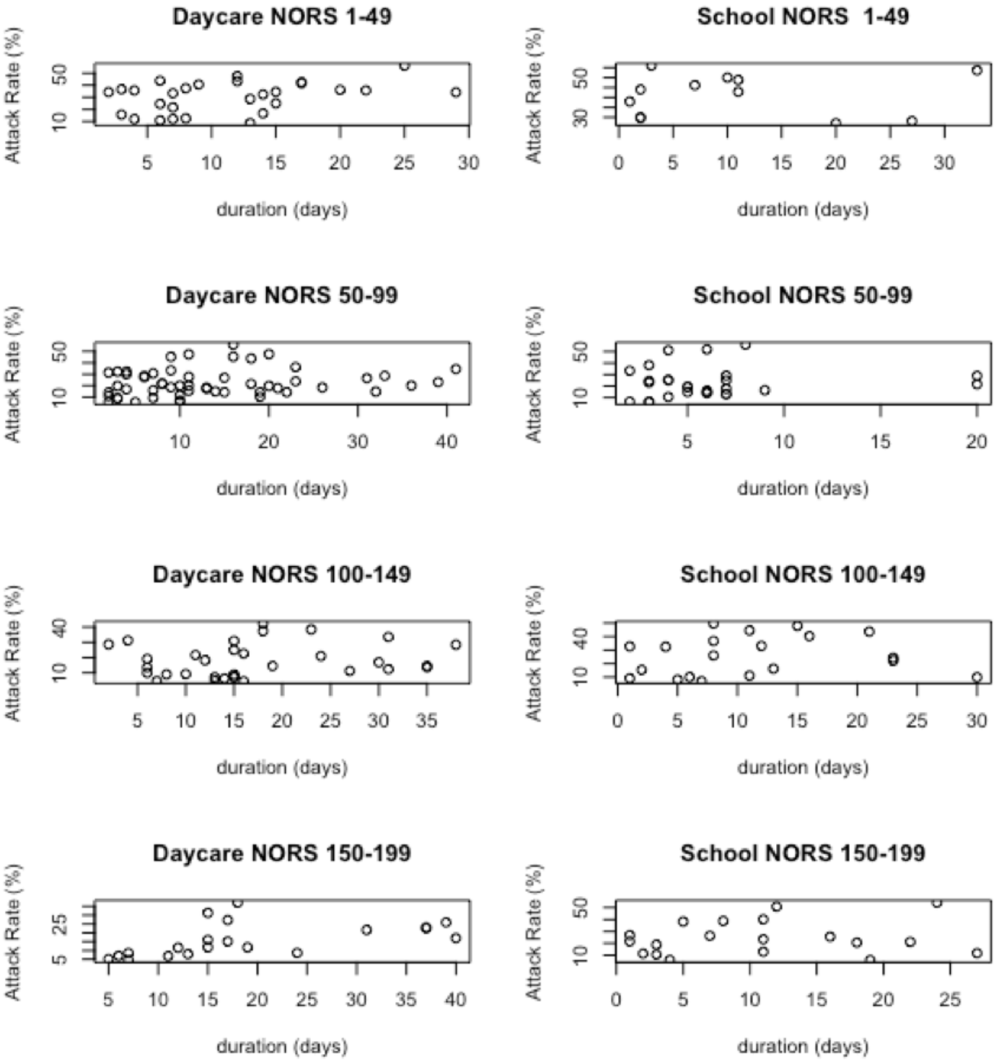


Figure C.1: NORS data: Attack rates vs. outbreak duration stratified by exposed population size.

C.7 Durations and Attack Rates from Model Runs

Below are venue-specific attacks rates and outbreaks durations for the models (i.e., baseline, immunity, individual exclusion and combined) and the NORS data.

Table C.3: Venue-specific Attack Rates for All Models: Median (95% CI) [Mean]

Model	Daycare	School
Baseline	67.6% (3.6%, 84.6%) [65%]	65.3% (0%, 83.3%) [60.3%]
Immunity	21.9% (2.2%, 49.2%) [23.1%]	16.7% (1.1%, 53.2%) [20.3%]
Individual Exclusion	63.3% (2.7%, 83.3%) [53.9%]	58.3% (0%, 81%) [45%]
Combined	20.9% (1.8%, 50.6%) [22.6%]	14.5% (0.3%, 53.3%) [19%]
NORS	21.6% (4.6%, 69.2%) [25.5%]	15.3% (4.6%, 68.4%) [20.4%]

Table C.4: Venue-specific Outbreak Durations for All Models: Median (95% CI) [Mean]

Model	Daycare	School
Baseline	4 days (2, 12) [4.6]	4 days (0, 10) [4.4]
Immunity	4 days (1, 13) [4.7]	5 days (1, 14) [5.7]
Individual Exclusion	5 days (2, 21) [6.9]	5 days (0, 26) [7]
Combined	5 days (1, 21) [6.4]	5 days (1, 27) [7.1]
NORS	13 days (2, 40) [14.7]	8 days (1, 32) [10.8]

C.8 Results from Calibration for Each Model with NORS data

Below are pairwise scatter plots examining joint distributions of attack rate (%), outbreak durations (days), and population sizes (people). The NORS data is in the upper left corner and all models are displayed with individual points colored by the log of the number of Times Calibrated. Points in white were not resampled.

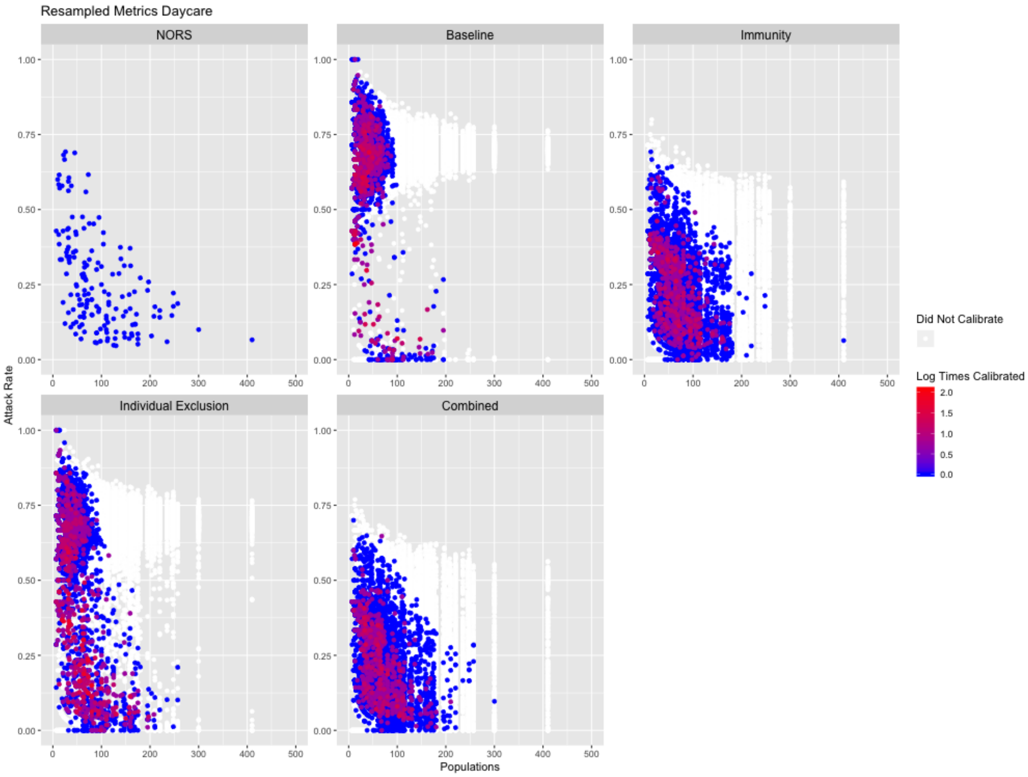


Figure C.2: Daycare Model Runs: Attack rates vs. population sizes results from resampled parameter and initial conditions. NORS data is in the top left. Points correspond to parameter sets and are colored by the amount of times they were resampled.

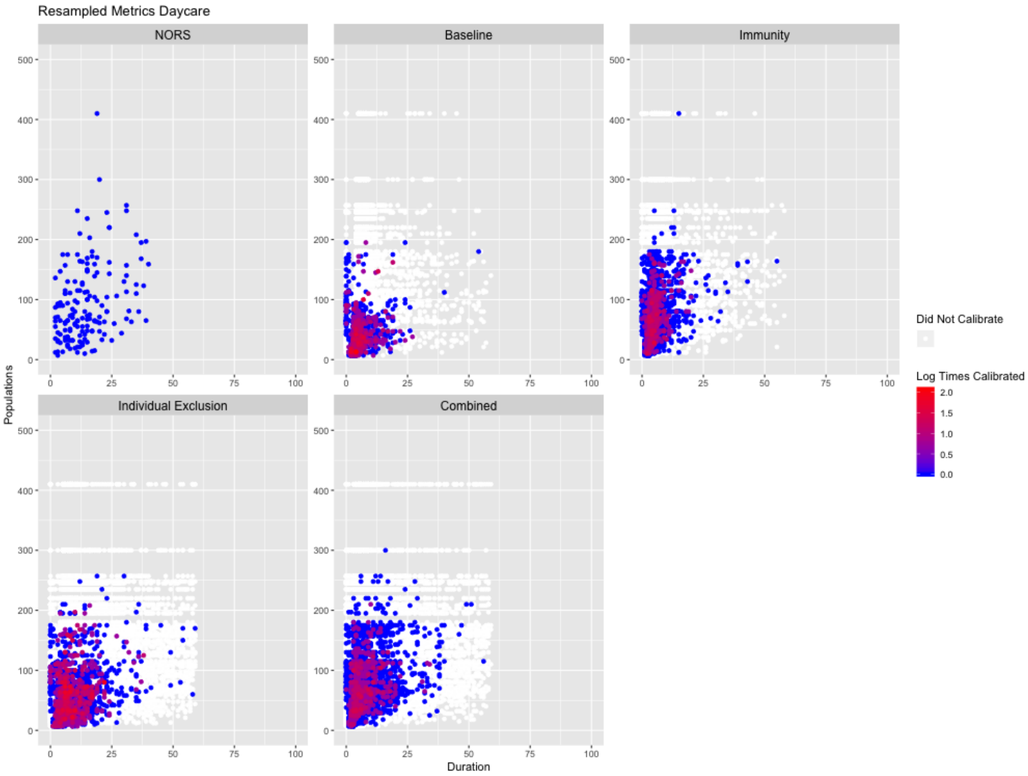


Figure C.3: Daycare Model Runs: Population sizes vs. outbreak durations results from resampled parameter and initial conditions. NORS data is in the top left. Points correspond to parameter sets and are colored by the amount of times they were resampled.

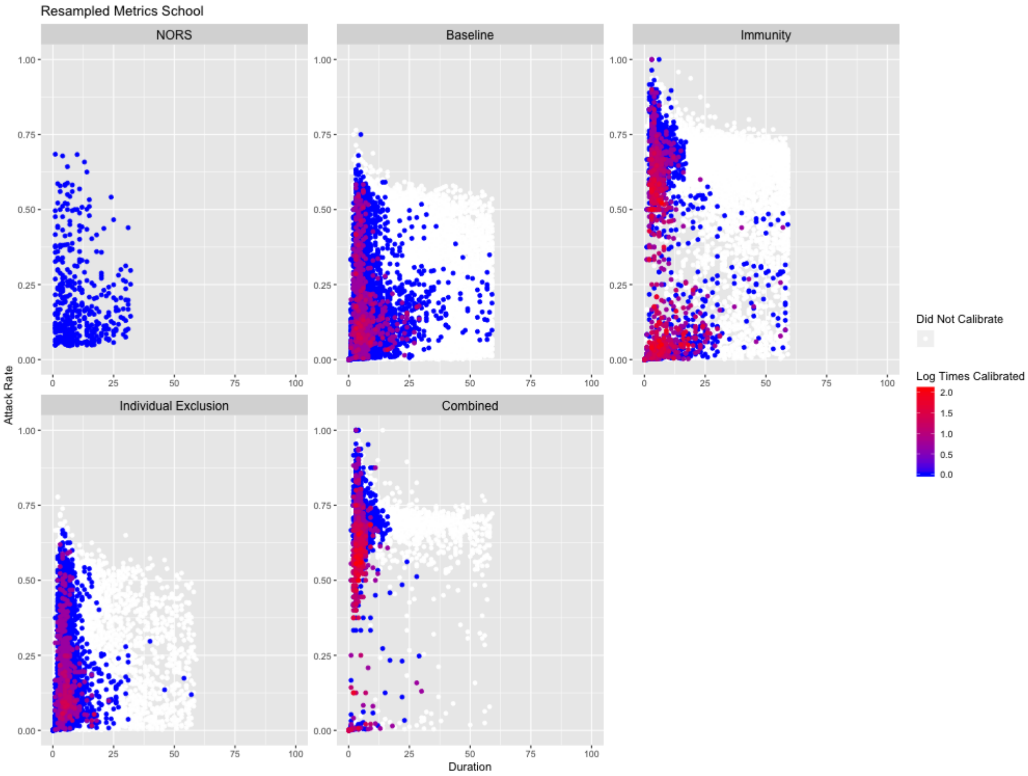


Figure C.4: School Model Runs: Attack rates vs. duration results from resampled parameter and initial conditions. NORS data is in the top left. Points correspond to parameter sets and are colored by the amount of times they were resampled.

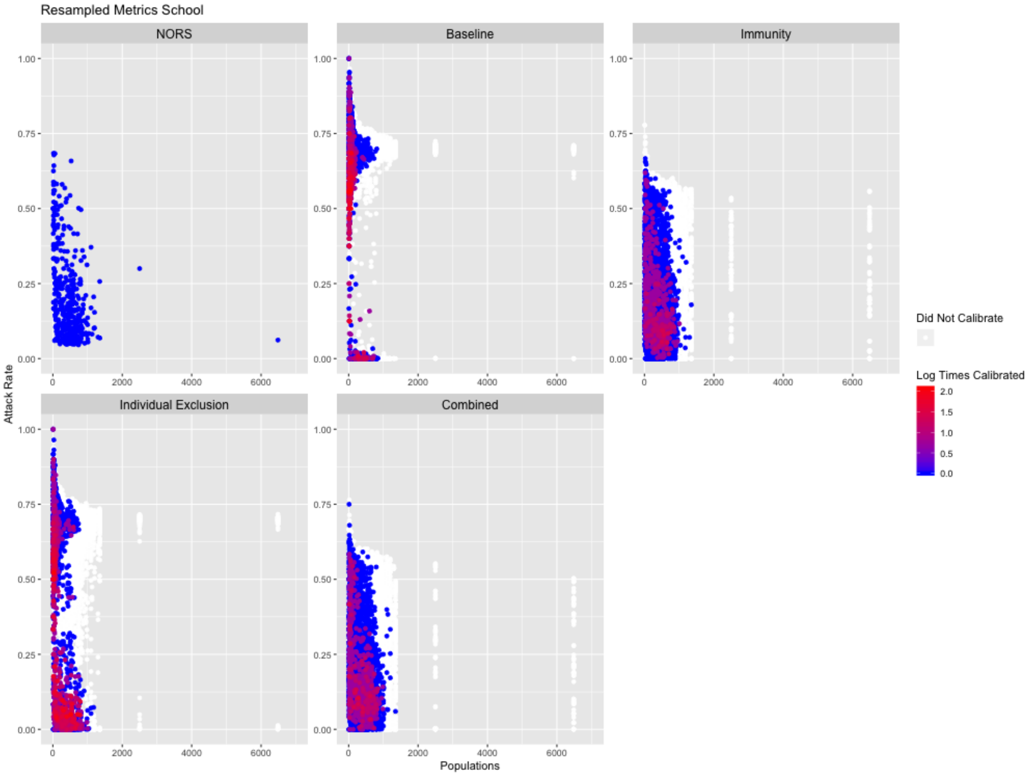


Figure C.5: School Model Runs: Attack rates vs. population sizes results from resampled parameter and initial conditions. NORS data is in the top left. Points correspond to parameter sets and are colored by the amount of times they were resampled.

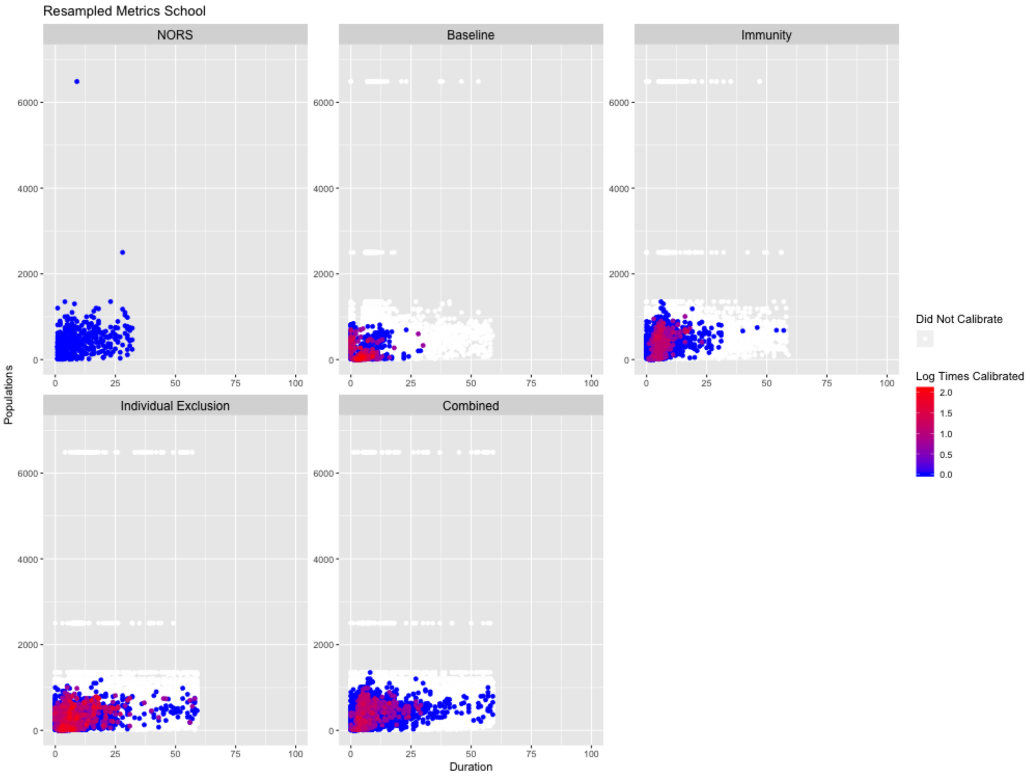


Figure C.6: School Model Runs: Population sizes vs. outbreak durations results from resampled parameter and initial conditions. NORS data is in the top left. Points correspond to parameter sets and are colored by the amount of times they were resampled.

C.9 Sensitivity Analyses

We conducted sensitivity analyses to ensure that our model results were robust to key simplifying assumptions. To ensure the duration calculations were not affected by the choice of initial conditions in different compartments, we conducted two sensitivity analyses. First, we ran the model varying the number of pathogens starting in the environment from 0 to 100 million and second, we seeded the model with a single infectious individual.

C.9.1 Sensitivity Analysis Results:

According to the KL divergence, all sensitivity analyses performed fairly well and were relatively close to the original combined model. Overall, the combined was lower according to KL divergence in the school model, but seeding with an infected individual had a lower KL divergence in the daycare model. See Table 4.2 for KL divergence values of the main analyses and Appendix Table C.7 for KL divergence values of the deterministic models.

Below are attack rates and durations from seeding scenario the sensitivity analyses.

Table C.5: Venue-specific Attack Rates for Sensitivity Analyses: Median (95% CI) [Mean]

Model	Daycare	School
Seeding: Varying Pathogens in Environment	21.4% (2.7%, 48.5%) [22.6%]	15% (0.8%, 52.7%) [19.1%]
Seeding: Diseased Individual	21.8% (0%, 50%) [22.7%]	13.4% (0%, 52.5%) [17.7%]
NORS	21.6% (4.6%, 69.2%) [25.5%]	15.3% (4.6%, 68.4%) [20.4%]

Table C.6: Venue-specific Outbreak Durations for Sensitivity Analyses: Median (95% CI) [Mean]

Model	Daycare	School
Seeding: Varying Pathogens in Environment	4 days (1, 17) [5]	5 days (1, 25) [6.5]
Seeding: Diseased Individual	5 days (0, 25) [7.3]	5 days (0, 24) [6.6]
NORS	13 days (2, 40) [14.7]	8 days (1, 32) [10.8]

C.10 Deterministic Model Calibration Sensitivity Analysis

We calibrated deterministic versions of each model to determine the whether stochastic extinction could lead to rapid outbreaks with lower attack rates. In general, deterministic models performed worse according to KL divergence. All models performed the same relative to each other with the combined and individual exclusion models performing the best.

Model	Daycare	School
Deterministic Simulations		
Baseline	19.3	16.87
Immunity	11.34	11.57
Individual Exclusion	10.57	10.19
Combined	10.45	10.59

Table C.7: Kullback Leibler (KL) divergence for each model compared with the NORS data kernel density estimated distribution. Smaller KL divergence indicates a more similar distribution to NORS (i.e. less information difference between the NORS and the model distribution).

Below are the attack rates and durations from the deterministic model runs. Overall, deterministic model runs resulted in less variation in attack rates and durations with slightly shorter median durations.

Table C.8: Venue-specific Attack Rates for All Models: Median (95% CI) [Mean]

Model	Daycare	School
Baseline	70% (70%, 70%) [70%]	70% (70%, 70%) [69.9%]
Immunity	22.6% (3.8%, 47.5%) [23.4%]	15.7% (2.6%, 52.6%) [19.9%]
Individual Exclusion	70% (5.3%, 70%) [54.2%]	55.5% (0.5%, 70%) [40.4%]
Combined	20.5% (3.2%, 47.9%) [22.3%]	13.4% (0.9%, 52.5%) [17.9%]
NORS	21.6% (4.6%, 69.2%) [25.5%]	15.3% (4.6%, 68.4%) [20.4%]

Table C.9: Venue-specific Outbreak Durations for All Models: Median (95% CI) [Mean]

Model	Daycare	School
Baseline	4 days (3, 5) [4.1]	4 days (3, 5) [4.3]
Immunity	4 days (2, 5) [3.8]	5 days (2, 6) [4.4]
Individual Exclusion	5 days (1, 16) [6.1]	6 days (0, 26) [7.9]
Combined	4 days (0, 13) [5.1]	5 days (0, 24) [7.1]
NORS	13 days (2, 40) [14.7]	8 days (1, 32) [10.8]

C.11 Staff and Students Model

We added staff into the model, to understand whether or not they can affect how norovirus is spread within venues.

C.11.1 Students and Staff Model for Daycare and School

Thus, in the staff and student model, there is a staff age group and a student age group. To derive the human-to-human transmission rates, we assume that the younger age group (i.e., the students) transmit at higher rates than the older age group (i.e., the staff) due to both contact rates [208] and susceptibility decreasing with age (e.g. represented by levels of norovirus antibody titers [209]). Specifically, the human-to-human transmission matrix is derived by taking the β_{HH} from the students only model and setting that to the student to student transmission rate. Next, we assume that the inter-age transmission rates (i.e., staff to student and student to staff transmission are equal) and calculate that by multiplying the student to student transmission rate by a randomly sampled reduction factor between [0,1]. Finally, the staff-to-staff transmission rate is calculated by multiplying the inter-age transmission rate by a randomly sampled reduction factor between [0,1] (this factor is also used to derive the fomite-to-staff transmission rate.

Next, for fomite-to-human transmission there are two rates, one for students and one for staff. The fomite-to-student transmission rate is calculated in the same way as the student only model i.e., β_{HH} multiplied by a randomly sampled parameter between [0, 2]. The fomite-to-staff transmission rate is derived by multiplying the fomite-to-student rate by the same factor used to derive the staff-to-staff transmission rate (mentioned above) between [0,1]. Overall, the force of infection for the staff and students model is as follows:

$$\begin{aligned}\lambda(t)_{tot} &= [I + \beta_A(A_1 + A_2 + A_3)]\beta_{HH} \\ \lambda(t)_1 &= \lambda(t)_1 + (F_1 + F_2)\beta_{Wk} \\ \lambda(t)_2 &= \lambda(t)_2 + (F_1 + F_2)\beta_{Wa}\end{aligned} \tag{C.6}$$

where $\lambda(t)_{tot}$ is the total force of infection (and is a vector representing the student force of infection as the first element and the staff force of infection as the second element. $\lambda(t)_1$ and $\lambda(t)_2$ are added to the force of infection for students and staff, respectively. Finally, β_{Wk} and β_{Wa} are the fomite-to-human transmission rates for students and staff, respectively.

Finally, with respect to shedding, staff and students shed into a single shared environment. Thus, the shedding and fomite tracking equations are as follows:

$$\begin{aligned}Shedding &= \alpha_I I + \alpha_I \beta_A(A_1 e^{-\sigma(\frac{1}{\phi})} + A_2 e^{-\sigma(\frac{1}{\phi}+\frac{1}{\rho})} + A_3 e^{-\sigma(\frac{1}{\phi}+\frac{1}{\rho}+\frac{1}{\rho})}) \\ \dot{F_1} &= \sum Shedding - \xi F_1 \\ \dot{F_2} &= \xi F_1 - \xi F_2\end{aligned} \tag{C.7}$$

where the sum of shedding across both age groups is added to the F_1 compartment because there is a single environmental compartment in each venue.

For the immunity and combined models we assumed that staff had higher rates of partial immunity than children [209].

All other model equations are the same as the student only model, see Appendix Section C.3 for details.

C.11.2 Students and Staff Model Likelihood Calculation

To derive an overall likelihood for a given venue, we took the NORS KDE values which corresponded to a given AR and population size for students from the model and multiplied by the NORS KDE values which corresponded to a given AR and population size for staff from the model and finally, multiplied by the NORS KDE values which corresponded to a given duration from the model. More details can be found in Section 4.4.3.

C.11.3 Students and Staff Model Results

Table C.10: Venue-specific Attack Rates for Students and Staff Model: Median (95% CI) [Mean]

Model	Daycare Students	Daycare Staff	Daycare Pooled	School Students	School Staff	School Pooled
Baseline	63.6% (0%, 75%) [59.3%]	50% (0%, 80%) [49.2%]	61.5% (0%, 72.7%) [57.4%]	57.4% (0%, 75%) [40.6%]	0% (0%, 60.9%) [21.9%]	54.5% (0%, 71.7%) [38.9%]
Immunity	24.6% (4.7%, 53.1%) [25.6%]	11.1% (0%, 40%) [13.4%]	22.4% (4.3%, 47.7%) [23.3%]	18.3% (3%, 51.3%) [21.4%]	5.3% (0%, 29.9%) [7.5%]	16.7% (2.9%, 47.7%) [19.5%]
Individual Exclusion	51.2% (1.9%, 72.9%) [42.9%]	23.1% (0%, 75%) [27.8%]	45.6% (2%, 70.5%) [40%]	9.6% (0%, 71.3%) [21%]	2.6% (0%, 54.5%) [8.9%]	7.9% (0%, 68.6%) [19.1%]
Combined	23.1% (3.1%, 52.3%) [24.4%]	8% (0%, 35.7%) [10.5%]	20.8% (2.8%, 45.8%) [21.7%]	17% (1.7%, 51.2%) [20.2%]	3.1% (0%, 24.4%) [5.3%]	15.1% (1.3%, 46.9%) [18.1%]
NORS	21.6% (4.6%, 69.2%) [25.5%]			15.3% (4.6%, 68.4%) [20.4%]		

Table C.11: Venue-specific Outbreak Durations for Students and Staff Model: Median (95% CI) [Mean]

Model	Daycare	School
Baseline	5 days (0, 17) [6.3]	4 days (0, 13) [4.4]
Immunity	4 days (2, 10) [4.6]	5 days (2, 9) [5.4]
Individual Exclusion	8 days (2, 30) [10.7]	7 days (0, 30) [9.6]
Combined	6 days (2, 23) [8]	6 days (2, 27) [8.6]
NORS	13 days (2, 40) [14.7]	8 days (1, 32) [10.8]

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Tyler J VanderWeele. Commentary: Resolutions of the birthweight paradox: competing explanations and analytical insights. International journal of epidemiology, 43(5):1368–1373, 2014.

[2] Samuel H Preston and Andrew Stokes. Obesity paradox: conditioning on disease enhances biases in estimating the mortality risks of obesity. Epidemiology (Cambridge, Mass.), 25(3):454, 2014.

[3] John C Lang, Hans De Sterck, Jamieson L Kaiser, and Joel C Miller. Analytic models for sir disease spread on random spatial networks. Journal of Complex Networks, 6(6):948–970, 2018.

[4] John Fox and Sanford Weisberg. An R companion to applied regression. Sage Publications, 2018.

[5] Molly K Steele, Justin V Remais, Manoj Gambhir, John W Glasser, Andreas Handel, Umesh D Parashar, and Benjamin A Lopman. Targeting pediatric versus elderly populations for norovirus vaccines: a model-based analysis of mass vaccination options. Epidemics, 17:42–49, 2016.

[6] Miguel A Hernán and James M Robins. Estimating causal effects from epidemiological data. Journal of Epidemiology & Community Health, 60(7):578–586, 2006.

[7] Kamyar Kalantar-Zadeh. Cardiovascular and survival paradoxes in dialysis patients: What is so bad about reverse epidemiology anyway? In Seminars in dialysis, volume 20, pages 593–601. Wiley Online Library, 2007.

[8] Joseph NS Eisenberg, James C Scott, and Travis Porco. Integrating disease control strategies: balancing water sanitation and hygiene interventions to reduce diarrheal disease burden. American Journal of Public Health, 97(5):846–852, 2007.

[9] Suzanne M Marks, Zachary Taylor, Noreen L Qualls, Robin J Shrestha-Kuwahara, Maureen A Wilce, and Cristy H Nguyen. Outcomes of con-

tact investigations of infectious tuberculosis patients. American journal of respiratory and critical care medicine, 162(6):2033–2038, 2000.

[10] Joint Tuberculosis Committee of the British Thoracic Society et al. Control and prevention of tuberculosis in the united kingdom: code of practice 2000. Thorax, 55(11):887–901, 2000.

[11] Helen Ayles, Monde Muyoyeta, Elizabeth Du Toit, Ab Schaap, Sian Floyd, Musonda Simwinga, Kwame Shanaube, Nathaniel Chishinga, Virginia Bond, Rory Dunbar, et al. Effect of household and community interventions on the burden of tuberculosis in southern africa: the zamstar community-randomised trial. The Lancet, 382(9899):1183–1194, 2013.

[12] Greg J Fox, Nguyen V Nhung, Dinh N Sy, Nghiem LP Hoa, Le TN Anh, Nguyen T Anh, Nguyen B Hoa, Nguyen H Dung, Tran N Buu, Nguyen T Loi, et al. household-contact investigation for detection of tuberculosis in vietnam. New England Journal of Medicine, 378(3):221–229, 2018.

[13] Gavin J Churchyard, Katherine L Fielding, James J Lewis, Leonie Coetzee, Elizabeth L Corbett, Peter Godfrey-Faussett, Richard J Hayes, Richard E Chaisson, and Alison D Grant. A trial of mass isoniazid preventive therapy for tuberculosis control. New England Journal of Medicine, 370(4):301–310, 2014.

[14] Colleen F Hanrahan, Bareng AS Nonyane, Lesego Mmolawa, Nora S West, Tsundzukani Siwelana, Limakatso Lebina, Neil Martinson, and David W Dowdy. Contact tracing versus facility-based screening for active tb case finding in rural south africa: A pragmatic cluster-randomized trial (kharitode tb). PLoS medicine, 16(4):e1002796, 2019.

[15] Suzanne Verver, Robin M Warren, Zahn Munch, Madalene Richardson, Gian D van der Spuy, Martien W Borgdorff, Marcel A Behr, Nulda Beyers, and Paul D van Helden. Proportion of tuberculosis transmission that takes place in households in a high-incidence area. The Lancet, 363(9404):212–214, 2004.

[16] Judith R Glynn, José Afonso Guerra-Assunção, Rein MGJ Houben, Lifted Sichali, Themba Mzembe, Lorrain K Mwaungulu, J Nimrod Mwaungulu, Ruth McNerney, Palwasha Khan, Julian Parkhill, et al. Whole genome sequencing shows a low proportion of tuberculosis disease is attributable to known close contacts in rural malawi. PloS one, 10(7):e0132840, 2015.

[17] Jonathan L Zelner, Megan B Murray, Mercedes C Becerra, Jerome Galea, Leonid Lecca, Roger Calderon, Rosa Yataco, Carmen Contreras, Zibiao

Zhang, Bryan T Grenfell, et al. Age-specific risks of tuberculosis infection from household and community exposures and opportunities for interventions in a high-burden setting. American journal of epidemiology, 180(8):853–861, 2014.

[18] Jon Zelner, Megan Murray, Mercedes Becerra, Jerome Galea, Leonid Lecca, Roger Calderon, Rosa Yataco, Zibiao Zhang, and Ted Cohen. Protective effects of household-based tb interventions are robust to neighbourhood-level variation in exposure risk in lima, peru: a model-based analysis. International journal of epidemiology, 2017.

[19] David W Dowdy, Jonathan E Golub, Richard E Chaisson, and Valeria Saraceni. Heterogeneity in tuberculosis transmission and the role of geographic hotspots in propagating epidemics. Proceedings of the National Academy of Sciences, 109(24):9557–9562, 2012.

[20] Z Munch, SWP Van Lill, CN Booysen, HL Zietsman, DA Enarson, and N Beyers. Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. International journal of tuberculosis and lung disease, 7(3):271–277, 2003.

[21] Stephanie M Karst. Pathogenesis of noroviruses, emerging rna viruses. Viruses, 2(3):748–781, 2010.

[22] Ben Lopman, Paul Gastanaduy, Geun Woo Park, Aron J Hall, Umesh D Parashar, and Jan Vinjé. Environmental transmission of norovirus gastroenteritis. Current opinion in virology, 2(1):96–102, 2012.

[23] Scott P Grytdal, Emilio DeBess, Lore E Lee, David Blythe, Patricia Ryan, Christianne Biggs, Miriam Cameron, Mark Schmidt, Umesh D Parashar, and Aron J Hall. Incidence of norovirus and other viral pathogens that cause acute gastroenteritis (age) among kaiser permanente member populations in the united states, 2012–2013. PloS one, 11(4):e0148395, 2016.

[24] C Karsten, S Baumgarte, AW Friedrich, C Von Eiff, K Becker, W Wosniok, A Ammon, J Bockemühl, H Karch, and H-I Huppertz. Incidence and risk factors for community-acquired acute gastroenteritis in north-west germany in 2004. European journal of clinical microbiology & infectious diseases, 28(8):935, 2009.

[25] Sarah J O'brien, Anna L Donaldson, Miren Iturriza-Gomara, and Clarence C Tam. Age-specific incidence rates for norovirus in the community and presenting to primary healthcare facilities in the united kingdom. The Journal of infectious diseases, 213(suppl_1):S15–S18, 2015.

[26] Aron J Hall, Mary E Wikswo, Karunya Manikonda, Virginia A Roberts, Jonathan S Yoder, and L Hannah Gould. Acute gastroenteritis surveillance through the national outbreak reporting system, united states. Emerging infectious diseases, 19(8):1305, 2013.

[27] George Rebane and Judea Pearl. The recovery of causal poly-trees from statistical data. arXiv preprint arXiv:1304.2736, 2013.

[28] Sander Greenland, Judea Pearl, and James M Robins. Causal diagrams for epidemiologic research. Epidemiology, pages 37–48, 1999.

[29] M Maria Glymour. Using causal diagrams to understand common problems in social epidemiology. Methods in social epidemiology, pages 393–428, 2006.

[30] NL Fleischer and AV Diez Roux. Using directed acyclic graphs to guide analyses of neighbourhood health effects: an introduction. Journal of Epidemiology & Community Health, 62(9):842–846, 2008.

[31] Tyler J VanderWeele, Miguel A Hernán, and James M Robins. Causal directed acyclic graphs and the direction of unmeasured confounding bias. Epidemiology (Cambridge, Mass.), 19(5):720, 2008.

[32] Sarah F Ackley, Elizabeth Rose Mayeda, Lee Worden, Wayne TA Enanoria, M Maria Glymour, and Travis C Porco. Compartmental model diagrams as causal representations in relation to dags. Epidemiologic Methods, 6(1), 2017.

[33] Michael Joffe, Manoj Gambhir, Marc Chadeau-Hyam, and Paolo Vineis. Causal diagrams in systems epidemiology. Emerging themes in epidemiology, 9(1):1, 2012.

[34] and. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 115(772):700–721, 1927.

[35] Suresh H Moolgavkar and Alfred G Knudson. Mutation and cancer: a model for human carcinogenesis. JNCI: Journal of the National Cancer Institute, 66(6):1037–1052, 1981.

[36] Petre Stoica and TORSTEN SÖDERSTRÖM. On the parsimony principle. International Journal of Control, 36(3):409–418, 1982.

[37] Howard Howie Weiss. The sir model and the foundations of public health. Materials matematics, pages 0001–17, 2013.

[38] Shweta Bansal, Bryan T Grenfell, and Lauren Ancel Meyers. When individual behaviour matters: homogeneous and network models in epidemiology. Journal of the Royal Society Interface, 4(16):879–891, 2007.

[39] Jonathan L Zelner, Aaron A King, Christine L Moe, and Joseph NS Eisenberg. How infections propagate after point-source outbreaks!" an analysis of secondary norovirus transmission". Epidemiology, pages 711–718, 2010.

[40] Simeone Marino, Ian B Hogue, Christian J Ray, and Denise E Kirschner. A methodology for performing global uncertainty and sensitivity analysis in systems biology. Journal of theoretical biology, 254(1):178–196, 2008.

[41] Sonia Hernández-Díaz, Enrique F Schisterman, and Miguel A Hernán. The birth weight "paradox" uncovered? American journal of epidemiology, 164(11):1115–1120, 2006.

[42] Joel C Kleinman, Mitchell B Pierre Jr, Jennifer H Madans, Garland H Land, and Wayne F Schramm. The effects of maternal smoking on fetal and infant mortality. American Journal of Epidemiology, 127(2):274–282, 1988.

[43] J Yerushalmy. The relationship of parents' cigarette smoking to outcome of pregnancy—implications as to the problem of inferring causation from observed associations. International journal of epidemiology, 43(5):1355–1366, 2014.

[44] Brian MacMahon, Marc Alpert, Eva J Salber, et al. Infant weight and parental smoking habits. American Journal of Epidemiology, 82(3):247–61, 1965.

[45] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. Annual review of sociology, 27(1):415–444, 2001.

[46] Juan Pablo Aparicio and Mercedes Pascual. Building epidemiological models from r 0: an implicit treatment of transmission in networks. Proceedings of the Royal Society B: Biological Sciences, 274(1609):505–512, 2006.

[47] Alden S Klovdahl. Social networks and the spread of infectious diseases: the aids example. Social science & medicine, 21(11):1203–1216, 1985.

[48] Matt J Keeling and Ken TD Eames. Networks and epidemic models. Journal of the Royal Society Interface, 2(4):295–307, 2005.

[49] Odo Diekmann, MCM De Jong, and Johan Anton Jacob Metz. A deterministic epidemic model taking account of repeated contacts between the same individuals. Journal of Applied Probability, 35(2):448–462, 1998.

[50] Denis Mollison. Spatial contact models for ecological and epidemic spread. Journal of the Royal Statistical Society: Series B (Methodological), 39(3):283–313, 1977.

[51] Duncan J Watts and Steven H Strogatz. Collective dynamics of ‘small-world’networks. nature, 393(6684):440, 1998.

[52] Lauren Ancel Meyers, Babak Pourbohloul, Mark EJ Newman, Danuta M Skowronski, and Robert C Brunham. Network theory and sars: predicting outbreak diversity. Journal of theoretical biology, 232(1):71–81, 2005.

[53] Mark EJ Newman. Spread of epidemic disease on networks. Physical review E, 66(1):016128, 2002.

[54] Matthew J Silk, Darren P Croft, Richard J Delahay, David J Hodgson, Nicola Weber, Mike Boots, and Robbie A McDonald. The application of statistical network models in disease research. Methods in Ecology and Evolution, 8(9):1026–1041, 2017.

[55] Jonathan M Read and Matt J Keeling. Disease evolution on networks: the role of contact structure. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270(1516):699–708, 2003.

[56] Victoria J Cook, Sumi J Sun, Jane Tapia, Stephen Q Muth, D Fermín Argüello, Bryan L Lewis, Richard B Rothenberg, and Peter D McElroy. Transmission network analysis in tuberculosis contact investigations. The Journal of infectious diseases, 196(10):1517–1527, 2007.

[57] Collette N Classen, Robin Warren, Madeleine Richardson, John H Hauman, Robert P Gie, James HP Ellis, Paul D van Helden, and Nulda Beyers. Impact of social interactions in the community on the transmission of tuberculosis in a high incidence area. Thorax, 54(2):136–140, 1999.

[58] Christopher Dye, Suzanne Scheele, Paul Dolin, Vikram Pathania, Mario C Raviglione, et al. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. Jama, 282(7):677–686, 1999.

[59] Rein MGJ Houben and Peter J Dodd. The global burden of latent tuberculosis infection: a re-estimation using mathematical modelling. PLoS medicine, 13(10):e1002152, 2016.

[60] Troels Lillebaek, Asger Dirksen, Inga Baess, Benedicte Strunge, Vibeke Ø Thomsen, and Åse B Andersen. Molecular evidence of endogenous reactivation of mycobacterium tuberculosis after 33 years of latent infection. The Journal of infectious diseases, 185(3):401–404, 2002.

[61] C Robert Horsburgh Jr. Priorities for the treatment of latent tuberculosis infection in the united states. New England Journal of Medicine, 350(20):2060–2067, 2004.

[62] Marcel A Behr, Paul H Edelstein, and Lalita Ramakrishnan. Revisiting the timetable of tuberculosis. Bmj, 362:k2738, 2018.

[63] C Huang, MC Becerra, R Calderon, C Contreras, J Galea, L Grandjean, L Lecca, R Yataco, Z Zhang, and M Murray. Isoniazid preventive therapy protects against tuberculosis among household contacts of isoniazid-resistant patients. 2018.

[64] Mukund Uplekar, World Health Organization, et al. The stop tb strategy: Building on and enhancing dots to meet the tb-related millennium development goals. 2006.

[65] World Health Organization et al. Recommendations for investigating contacts of persons with infectious tuberculosis in low-and middle-income countries. World Health Organization, 2012.

[66] D Behera. Implementing the who stop tb strategy: A handbook for national tuberculosis control programmes. Indian Journal of Medical Research, 130(1):95–97, 2009.

[67] Jonathan L Zelner, Megan B Murray, Mercedes C Becerra, Jerome Galea, Leonid Lecca, Roger Calderon, Rosa Yataco, Carmen Contreras, Zibiao Zhang, Bryan T Grenfell, et al. Bacillus calmette-guérin and isoniazid preventive therapy protect contacts of patients with tuberculosis. American journal of respiratory and critical care medicine, 189(7):853–859, 2014.

[68] Greg J Fox, Peter J Dodd, and Ben J Marais. Household contact investigation to improve tuberculosis control. The Lancet Infectious Diseases, 19(3):235–237, 2019.

[69] Gregory J Fox, Simone E Barry, Warwick J Britton, and Guy B Marks. Contact investigation for tuberculosis: a systematic review and meta-analysis. European Respiratory Journal, 41(1):140–156, 2013.

[70] Parastu Kasaie, Jason R Andrews, W David Kelton, and David W Dowdy. Timing of tuberculosis transmission and the impact of household contact tracing. an agent-based simulation model. American journal of respiratory and critical care medicine, 189(7):845–852, 2014.

[71] Sheng Li, Joseph NS Eisenberg, Ian H Spicknall, and James S Koopman. Dynamics and control of infections transmitted from person to person through

the environment. American journal of epidemiology, 170(2):257–265, 2009.

[72] Herbert W Hethcote. The mathematics of infectious diseases. SIAM review, 42(4):599–653, 2000.

[73] Joseph H Tien and David JD Earn. Multiple transmission pathways and disease dynamics in a waterborne pathogen model. Bulletin of mathematical biology, 72(6):1506–1533, 2010.

[74] Otso Ovaskainen and Baruch Meerson. Stochastic models of population extinction. Trends in ecology & evolution, 25(11):643–652, 2010.

[75] Richard Frankham, David A Briscoe, and Jonathan D Ballou. Introduction to conservation genetics. Cambridge university press, 2002.

[76] Chunyan Ji and Daqing Jiang. Threshold behaviour of a stochastic sir model. Applied Mathematical Modelling, 38(21-22):5067–5079, 2014.

[77] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. The journal of physical chemistry, 81(25):2340–2361, 1977.

[78] Daniel T Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. The Journal of Chemical Physics, 115(4):1716–1733, 2001.

[79] Mark S Riddle and Richard I Walker. Status of vaccine research and development for norovirus. Vaccine, 34(26):2895–2899, 2016.

[80] Aron J Hall, Mariana Rosenthal, Nicole Gregoricus, Sharon A Greene, Jeana Ferguson, Olga L Henao, Jan Vinjé, Ben A Lopman, Umesh D Parashar, and Marc-Alain Widdowson. Incidence of acute gastroenteritis and role of norovirus, georgia, usa, 2004–2005. Emerging infectious diseases, 17(8):1381, 2011.

[81] Kari Debbink, Lisa C Lindesmith, and Ralph S Baric. The state of norovirus vaccines. Clinical Infectious Diseases, 58(12):1746–1752, 2014.

[82] JE Matthews, BW Dickey, RD Miller, JR Felzer, BP Dawson, AS Lee, JJ Rocks, J Kiel, JS Montes, CL Moe, et al. The epidemiology of published norovirus outbreaks: a review of risk factors associated with attack rate and genogroup. Epidemiology & Infection, 140(7):1161–1172, 2012.

[83] Karin Nygård, Maria Torvén, Camilla Ancker, Siv Britt Knauth, Kjell-Olof Hedlund, Johan Giesecke, Yvonne Andersson, and Lennart Svens-

son. Emerging genotype (ggiib) of norovirus in drinking water, sweden. Emerging infectious diseases, 9(12):1548, 2003.

[84] Linda Verhoef, Joanne Hewitt, Leslie Barclay, Sharia Ahmed, Rob Lake, Aron J Hall, Ben Lopman, Annelies Kroneman, Harry Vennema, et al. Norovirus genotype profiles associated with foodborne transmission, 1999 ?" 2012. Emerging infectious diseases, 21(4):592, 2015.

[85] JS Cheesbrough, J Green, CI Gallimore, PA Wright, and DWG Brown. Widespread environmental contamination with norwalk-like viruses (nlv) detected in a prolonged hotel outbreak of gastroenteritis. Epidemiology & Infection, 125(1):93–98, 2000.

[86] Mary E Wikswo and Aron J Hall. Outbreaks of acute gastroenteritis transmitted by person-to-person contact—united states, 2009–2010. Morbidity and Mortality Weekly Report: Surveillance Summaries, 61(9):1–12, 2012.

[87] Robert L Atmar, Antone R Opekun, Mark A Gilger, Mary K Estes, Sue E Crawford, Frederick H Neill, Sasirekha Ramani, Heather Hill, Jennifer Ferreira, and David Y Graham. Determination of the 50% human infectious dose for norwalk virus. The Journal of infectious diseases, 209(7):1016–1022, 2013.

[88] MO Milbrath, IH Spicknall, JL Zelner, CL Moe, and JNS Eisenberg. Heterogeneity in norovirus shedding duration affects community risk. Epidemiology & Infection, 141(8):1572–1584, 2013.

[89] Thomas A Parrino, David S Schreiber, Jerry S Trier, Albert Z Kapikian, and Neil R Blacklow. Clinical immunity in acute gastroenteritis caused by norwalk agent. New England Journal of Medicine, 297(2):86–89, 1977.

[90] Philip C Johnson, John J Mathewson, Herbert L DuPont, and Harry B Greenberg. Multiple-challenge study of host susceptibility to norwalk gastroenteritis in us adults. Journal of Infectious Diseases, 161(1):18–21, 1990.

[91] S Fallahi and K Mattison. Evaluation of murine norovirus persistence in environments relevant to food production and processing. Journal of food protection, 74(11):1847–1851, 2011.

[92] Miho Kobayashi, Shima Yoshizumi, Sayaka Kogawa, Tomoko Takahashi, Yo Ueki, Michiyo Shinohara, Fuminori Mizukoshi, Hiroyuki Tsukagoshi, Yoshiko Sasaki, Rieko Suzuki, et al. Molecular evolution of the capsid gene in norovirus genogroup i. Scientific reports, 5:13806, 2015.

[93] Du-Ping Zheng, Tamie Ando, Rebecca L Fankhauser, R Suzanne Beard,

Roger I Glass, and Stephan S Monroe. Norovirus classification and proposed strain nomenclature. Virology, 346(2):312–323, 2006.

[94] Benjamin A Lopman, Duncan Steele, Carl D Kirkwood, and Umesh D Parashar. The vast and varied global burden of norovirus: prospects for prevention and control. PLoS medicine, 13(4):e1001999, 2016.

[95] Robert L Atmar, David I Bernstein, Clayton D Harro, Mohamed S Al-Ibrahim, Wilbur H Chen, Jennifer Ferreira, Mary K Estes, David Y Graham, Antone R Opekun, Charles Richardson, et al. Norovirus vaccine against experimental human norwalk virus illness. New England Journal of Medicine, 365(23):2178–2187, 2011.

[96] J Joukje Siebenga, Harry Vennema, Bernadet Renckens, Erwin de Bruin, Bas van der Veer, Roland J Siezen, and Marion Koopmans. Epochal evolution of ggii. 4 norovirus capsid proteins from 1995 to 2006. Journal of virology, 81(18):9932–9941, 2007.

[97] Roger I Glass, Umesh D Parashar, and Mary K Estes. Norovirus gastroenteritis. New England Journal of Medicine, 361(18):1776–1785, 2009.

[98] Kirsten Simmons, Manoj Gambhir, Juan Leon, and Ben Lopman. Duration of immunity to norovirus gastroenteritis. Emerging infectious diseases, 19(8):1260, 2013.

[99] Kirsi Nurminen, Vesna Blazevic, Leena Huhti, Sirpa Räsänen, Tiia Koho, Vesa P Hytönen, and Timo Vesikari. Prevalence of norovirus gii-4 antibodies in finnish children. Journal of medical virology, 83(3):525–531, 2011.

[100] Ruta Kulkarni, Kavita Lole, and Shobha D Chitambar. Seroprevalence of antibodies against gii. 4 norovirus among children in pune, india. Journal of medical virology, 88(9):1636–1640, 2016.

[101] Johan Nordgren, Sumit Sharma, Anita Kambhampati, Ben Lopman, and Lennart Svensson. Innate resistance and susceptibility to norovirus infection. PLoS pathogens, 12(4):e1005385, 2016.

[102] Gustaf E Rydell, Elin Kindberg, Göran Larson, and Lennart Svensson. Susceptibility to winter vomiting disease: a sweet matter. Reviews in medical virology, 21(6):370–382, 2011.

[103] Pengwei Huang, Tibor Farkas, Séverine Marionneau, Weiming Zhong, Nathalie Ruvoën-Clouet, Ardythe L Morrow, Mekibib Altaye, Larry K Pickering, David S Newburg, Jacques LePendu, et al. Noroviruses bind to human

abo, lewis, and secretor histo-blood group antigens: identification of 4 distinct strain-specific patterns. The Journal of infectious diseases, 188(1):19–31, 2003.

[104] Pengwei Huang, Tibor Farkas, Weiming Zhong, Ming Tan, Scott Thornton, Ardythe L Morrow, and Xi Jiang. Norovirus and histo-blood group antigens: demonstration of a wide spectrum of strain specificities and classification of two major binding groups among multiple binding patterns. Journal of virology, 79(11):6714–6722, 2005.

[105] Haruko Shirato, Satoko Ogawa, Hiromi Ito, Takashi Sato, Akihiko Kameyama, Hisashi Narimatsu, Zheng Xiaofan, Tatsuo Miyamura, Takaji Wakita, Koji Ishii, et al. Noroviruses distinguish between type 1 and type 2 histo-blood group antigens for binding. Journal of virology, 82(21):10756–10767, 2008.

[106] Mercedes R Carnethon, Peter John D De Chavez, Mary L Biggs, Cora E Lewis, James S Pankow, Alain G Bertoni, Sherita H Golden, Kiang Liu, Kenneth J Mukamal, Brenda Campbell-Jenkins, et al. Association of weight status with mortality in adults with incident diabetes. Jama, 308(6):581–590, 2012.

[107] Emmanuel Aja Oga and Olabimpe Ruth Eseyin. The obesity paradox and heart failure: a systematic review of a decade of evidence. Journal of obesity, 2016, 2016.

[108] Lisa Oesch, Turgut Tatlisumak, Marcel Arnold, and Hakan Sarikaya. Obesity paradox in stroke–myth or reality? a systematic review. PloS one, 12(3):e0171334, 2017.

[109] Miguel A Hernán, Sonia Hernández-Díaz, and James M Robins. A structural approach to selection bias. Epidemiology, 15(5):615–625, 2004.

[110] Lorenzo Richiardi, Rino Bellocco, and Daniela Zugna. Mediation analysis in epidemiology: methods, interpretation and bias. International journal of epidemiology, 42(5):1511–1519, 2013.

[111] Hailey R Banack and Jay S Kaufman. From bad to worse: collider stratification amplifies confounding bias in the “obesity paradox”. European journal of epidemiology, 30(10):1111–1114, 2015.

[112] Alban De Schutter, Carl J Lavie, Sergey Kachur, Dharmendrakumar A Patel, and Richard V Milani. Body composition and mortality in a large cohort with preserved ejection fraction: untangling the obesity paradox. In Mayo Clinic Proceedings, volume 89, pages 1072–1079. Elsevier, 2014.

[113] Lisanne Schenkeveld, Michael Magro, Rohit M Oemrawsingh, Mattie Lenzen, Peter de Jaegere, Robert-Jan van Geuns, Patrick W Serruys, and Ron T van Domburg. The influence of optimal medical treatment on the ‘obesity paradox’, body mass index and long-term mortality in patients treated with percutaneous coronary intervention: a prospective cohort study. BMJ open, 2(1):e000535, 2012.

[114] Luigi Marzio Biasucci, Francesca Graziani, Vittoria Rizzello, Giovanna Liuzzo, Caterina Guidone, Alberto Ranieri De Caterina, Salvatore Brugaletta, Gertrude Mingrone, and Filippo Crea. Paradoxical preservation of vascular function in severe obesity. The American journal of medicine, 123(8):727–734, 2010.

[115] Michael Stein. Large sample properties of simulations using latin hypercube sampling. Technometrics, 29(2):143–151, 1987.

[116] & Meyer J. A. Howden, L. M. Age and sex composition: 2010. US Census Bureau, (22), 2011.

[117] Carole Willi, Patrick Bodenmann, William A Ghali, Peter D Faris, and Jacques Cornuz. Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. Jama, 298(22):2654–2664, 2007.

[118] Masayuki Itoh, Takao Tsuji, Kenji Nemoto, Hiroyuki Nakamura, and Kazutetsu Aoshiba. Undernutrition in patients with copd and its treatment. Nutrients, 5(4):1316–1335, 2013.

[119] Stephan von Haehling and Stefan D Anker. Cachexia as a major underestimated and unmet medical need: facts and numbers. Journal of cachexia, sarcopenia and muscle, 1(1):1–5, 2010.

[120] David M Mannino, D Thorn, A Swensen, and F Holguin. Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in copd. European Respiratory Journal, 32(4):962–969, 2008.

[121] Jason L Loeppky, Jerome Sacks, and William J Welch. Choosing the sample size of a computer experiment: A practical guide. Technometrics, 51(4):366–376, 2009.

[122] Nyi Nyi Naing. Easy way to learn standardization: direct and indirect methods. The Malaysian journal of medical sciences: MJMS, 7(1):10, 2000.

[123] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2017. R package version 3.3.3.

[124] KER Soetaert, Thomas Petzoldt, and R Woodrow Setzer. Solving differential equations in r: package desolve. Journal of Statistical Software, 33, 2010.

[125] Vojtech Hainer and Irena Aldhoon-Hainerová. Obesity paradox does exist. Diabetes care, 36(Supplement 2):S276–S281, 2013.

[126] Timo Koski and John Noble. Bayesian networks: an introduction, volume 924. John Wiley & Sons, 2011.

[127] Frank K Friedenberg, Derek M Tang, Thais Mendonca, and Vishwas Vanar. Predictive value of body mass index at age 18 on adulthood obesity: results of a prospective survey of an urban population. The American journal of the medical sciences, 342(5):371–382, 2011.

[128] World Health Organization et al. Global tuberculosis report 2018. In Global tuberculosis report 2018. 2018.

[129] Ewa Augustynowicz-Kopeć, Tomasz Jagielski, Monika Kozińska, Kristin Kremer, Dick van Soolingen, Jacek Bielecki, and Zofia Zwolska. Transmission of tuberculosis within family-households. Journal of Infection, 64(6):596–608, 2012.

[130] Jonathan E Golub, Wendy A Cronin, Olugbenga O Obasanjo, William Coggin, Kristina Moore, Diana S Pope, Deidre Thompson, Timothy R Sterling, Susan Harrington, William R Bishai, et al. Transmission of mycobacterium tuberculosis through casual contact with an infectious case. Archives of internal medicine, 161(18):2254–2258, 2001.

[131] Justin Lessler, Andrew S Azman, Heather S McKay, and Sean M Moore. What is a hotspot anyway? The American journal of tropical medicine and hygiene, 96(6):1270–1273, 2017.

[132] ELN Maciel, W Pan, R Dietze, RL Peres, SA Vinhas, FK Ribeiro, M Palaci, RR Rodrigues, E Zandonade, and JE Golub. Spatial patterns of pulmonary tuberculosis incidence and their relationship to socio-economic status in vitoria, brazil. The international journal of tuberculosis and lung disease, 14(11):1395–1402, 2010.

[133] Saroochi Agarwal, Duc T Nguyen, Larry D Teeter, and Edward A Graviss. Spatial-temporal distribution of genotyped tuberculosis cases in a county with active transmission. BMC infectious diseases, 17(1):378, 2017.

[134] Ian Haase, Sherry Olson, Marcel A Behr, Ian Wanyeki, Louise Thibert, Allison Scott, Alice Zwerling, Nancy Ross, Paul Brassard, Dick Menzies, et al. Use of geographic and genotyping tools to characterise tuberculosis transmission

in montreal. The International Journal of Tuberculosis and Lung Disease, 11(6):632–638, 2007.

[135] William R Bishai, Neil MH Graham, Susan Harrington, Diana S Pope, Nancy Hooper, Jacqueline Astemborski, Laura Sheely, David Vlahov, Gregory E Glass, and Richard E Chaisson. Molecular and geographic patterns of tuberculosis transmission after 15 years of directly observed therapy. Jama, 280(19):1679–1684, 1998.

[136] Peter M Small, Philip C Hopewell, Samir P Singh, Antonio Paz, Julie Parsonnet, Delaney C Ruston, Gisela F Schecter, Charles L Daley, and Gary K Schoolnik. The epidemiology of tuberculosis in san francisco–a population-based study using conventional and molecular methods. New England Journal of Medicine, 330(24):1703–1709, 1994.

[137] Fabíola Karla Correa Ribeiro, William Pan, Adelmo Bertolde, Solange Alves Vinhas, Renata Lyrio Peres, Lee Riley, Moisés Palaci, and Ethel Leonor Maciel. Genotypic and spatial analysis of mycobacterium tuberculosis transmission in a high-incidence urban setting. Clinical Infectious Diseases, 61(5):758–766, 2015.

[138] Wayner Vieira Souza, Maria de Fátima Militão Albuquerque, Cristhovam Castro Barcellos, Ricardo Arraes de Alencar Ximenes, and Marília Sá Carvalho. Tuberculosis in brazil: construction of a territorially based surveillance system. Revista de saude publica, 39(1):82–89, 2005.

[139] Gabriel Chamie, Midori Kato-Maeda, Devy M Emperador, Bonnie Wandera, Olive Mugagga, John Crandall, Michael Janes, Carina Marquez, Moses R Kamya, Edwin D Charlebois, et al. Spatial overlap links seemingly unconnected genotype-matched tb cases in rural uganda. PloS one, 13(2):e0192666, 2018.

[140] Edine W Tiemersma, Marieke J van der Werf, Martien W Borgdorff, Brian G Williams, and Nico JD Nagelkerke. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in hiv negative patients: a systematic review. PloS one, 6(4):e17601, 2011.

[141] Google data. https://www.google.com/publicdata/explore?ds=d5bncppjof8f9_&met_y=sp_dyn_le00_in&idim=country:USA&dl=en&hl=en&q=averagelifeexpectancy#!ctype=l&strail=false&bcs=d&nselm=h&met_y=sp_dyn_le00_in&scale_y=lin&ind_y=false&rdim=country&idim=country:PER&ifdim=country&hl=en_US&dl=en&ind=false;. Accessed: 2018-03-05.

[142] E Vynnycky and PEM Fine. The natural history of tuberculosis: the im-

plications of age-dependent risks of disease and the role of reinfection. Epidemiology & Infection, 119(2):183–201, 1997.

[143] Dermot Maher, Pierre Chaulet, Sergio Spinaci, A Harries, et al. Treatment of tuberculosis: guidelines for national programmes. Treatment of tuberculosis: guidelines for national programmes. Second edition., (Ed. 2):1–77, 1997.

[144] Maria Elvira Balcells, Sara L Thomas, Peter Godfrey-Faussett, and Alison D Grant. Isoniazid preventive therapy and risk for resistant tuberculosis. Emerging infectious diseases, 12(5):744, 2006.

[145] S Woldehanna and J Volmink. Treatment of latent tuberculosis infection in hiv infected persons. The Cochrane database of systematic reviews, (1):CD000171–CD000171, 2004.

[146] Jason R Andrews, Farzad Noubary, Rochelle P Walensky, Rodrigo Cerda, Elena Losina, and C Robert Horsburgh. Risk of progression to active tuberculosis following reinfection with mycobacterium tuberculosis. Clinical infectious diseases, 54(6):784–791, 2012.

[147] Ted Cohen, Caroline Colijn, Bryson Finklea, and Megan Murray. Exogenous re-infection and the dynamics of tuberculosis epidemics: local effects in a network model of transmission. Journal of the Royal Society Interface, 4(14):523–531, 2006.

[148] Jennifer Ho, Greg J Fox, and Ben J Marais. Passive case finding for tuberculosis is not enough. International journal of mycobacteriology, 5(4):374–378, 2016.

[149] John D Walley, M Amir Khan, James N Newell, and M Hussain Khan. Effectiveness of the direct observation component of dots for tuberculosis: a randomised controlled trial in pakistan. The lancet, 357(9257):664–669, 2001.

[150] Anders Ahlbom and Lars Alfredsson. Interaction: a word with two meanings creates confusion. European journal of epidemiology, 20(7):563–564, 2005.

[151] Mirjam J Knol, Tyler J VanderWeele, Rolf HH Groenwold, Olaf H Klungel, Maroeska M Rovers, and Diederick E Grobbee. Estimating measures of interaction on an additive scale for preventive exposures. European journal of epidemiology, 26(6):433–438, 2011.

[152] Nicky McCreesh and Richard G White. An explanation for the low proportion

of tuberculosis that results from transmission between household and known social contacts. Scientific reports, 8(1):5382, 2018.

[153] Natalia Blanco, Marisa C Eisenberg, Terri Stillwell, and Betsy Foxman. What transmission precautions best control influenza spread in a hospital? American journal of epidemiology, 183(11):1045–1054, 2016.

[154] Rachel E Gicquelais, Betsy Foxman, Joseph Coyle, and Marisa C Eisenberg. Hepatitis c transmission in young people who inject drugs: Insights using a dynamic model informed by state public health surveillance. Epidemics, 2019.

[155] Chen-Yuan Chiang and Lee W Riley. Exogenous reinfection in tuberculosis. The Lancet infectious diseases, 5(10):629–636, 2005.

[156] Gregg S Gonsalves, J Tyler Copple, Tyler Johnson, A David Paltiel, and Joshua L Warren. Bayesian adaptive algorithms for locating hiv mobile testing services. BMC medicine, 16(1):155, 2018.

[157] Aron J Hall, Ben A Lopman, Daniel C Payne, Manish M Patel, Paul A Gastañaduy, Jan Vinjé, and Umesh D Parashar. Norovirus disease in the united states. Emerging infectious diseases, 19(8):1198, 2013.

[158] ELIZABETH M HEUN, RICHARD L VOGT, PAUL J HUDSON, STEVE PARREN, and G WILLIAM GARY. Risk factors for secondary transmission in households after a common-source outbreak of norwalk gastroenteritis. American journal of epidemiology, 126(6):1181–1186, 1987.

[159] PJ Marks, IB Vipond, FM Regan, K Wedgwood, RE Fey, and EO Caul. A school outbreak of norwalk-like virus: evidence for airborne transmission. Epidemiology & Infection, 131(1):727–736, 2003.

[160] Hannelore Götz, Karl Ekdahl, Johan Lindbäck, Birgitta de Jong, Kjell Olof Hedlund, and Johan Giesecke. Clinical spectrum and transmission characteristics of infection with norwalk-like virus: findings from a large community outbreak in sweden. Clinical Infectious Diseases, 33(5):622–628, 2001.

[161] Christian JPA Hoebe, Harry Vennema, Ana Maria de Roda Husman, and Yvonne THP van Duynhoven. Norovirus outbreak among primary schoolchildren who had played in a recreational water fountain. The Journal of infectious diseases, pages 699–705, 2004.

[162] Elmira T Isakbaeva, Sandra N Bulens, R Suzanne Beard, Susan Adams, Stephan S Monroe, Sandra S Chaves, Marc-Alain Widdowson, and Roger I

Glass. Norovirus and child care: challenges in outbreak control. The Pediatric infectious disease journal, 24(6):561–563, 2005.

[163] Seiji Morioka, Tooru Sakata, Atsuko Tamaki, Takahide Shioji, Akira Funaki, Yasuo Yamamoto, Hiroomi Naka, Fumio Terasoma, Kenji Imai, and Koji Matsuo. A food-borne norovirus outbreak at a primary school in wakayama prefecture. Japanese journal of infectious diseases, 59(3):205, 2006.

[164] ZA Marsh, SP Grytdal, JC Beggs, E Leshem, PA Gastañaduy, B Rha, M Nyaku, BA Lopman, and AJ Hall. The unwelcome houseguest: secondary household transmission of norovirus. Epidemiology & Infection, 146(2):159–167, 2018.

[165] Jonathan L Adler and Raymond Zickl. Winter vomiting disease. The Journal of infectious diseases, 119(6):668–673, 1969.

[166] Ellen L Jones, Adam Kramer, Marlene Gaither, and Charles P Gerba. Role of fomite contamination during an outbreak of norovirus on houseboats. International journal of environmental health research, 17(2):123–131, 2007.

[167] Meirion Rhys Evans, R Meldrum, W Lane, D Gardner, CD Ribeiro, CI Gallimore, and D Westmoreland. An outbreak of viral gastroenteritis following environmental contamination at a concert hall. Epidemiology & Infection, 129(2):355–360, 2002.

[168] Doris H D'Souza, Arnie Sair, Karen Williams, Efstathia Papafragkou, Julie Jean, Christina Moore, and LeeAnn Jaykus. Persistence of caliciviruses on environmental surfaces and their transfer to food. International journal of food microbiology, 108(1):84–91, 2006.

[169] JS Cheesbrough, L Barkess-Jones, and DW Brown. Possible prolonged environmental survival of small round structured viruses. Journal of Hospital Infection, 35(4):325–326, 1997.

[170] Stefanie Clay, Sunil Maherchandani, Yashpal S Malik, and Sagar M Goyal. Survival on uncommon fomites of feline calicivirus, a surrogate of noroviruses. American journal of infection control, 34(1):41–43, 2006.

[171] PFM Teunis, FHA Sukhrie, Harry Vennema, Jolanda Bogerman, MFC Beersma, and MPG Koopmans. Shedding of norovirus in symptomatic and asymptomatic infections. Epidemiology & Infection, 143(8):1710–1717, 2015.

[172] Hui Xu, Qin Lin, Cong Chen, Jiantao Zhang, Huili Zhang, and Chao Hao. Epi-

demiology of norovirus gastroenteritis outbreaks in two primary schools in a city in eastern china. American journal of infection control, 41(10):e107–e109, 2013.

[173] Mikko Paunio, Heikki Peltola, Martti Valle, Irja Davidkin, Martti Virtanen, and Olli P Heinonen. Explosive school-based measles outbreak: intense exposure may have resulted in high risk, even among revaccinees. American journal of epidemiology, 148(11):1103–1110, 1998.

[174] William F Wells, Mildred W Wells, Theodore S Wilder, et al. The environmental control of epidemic contagion. i. an epidemiologic study of radiant disinfection of air in day schools. American Journal of Hygiene, 35(1):97–121, 1942.

[175] Rebecca L Currier, Daniel C Payne, Mary A Staat, Rangaraj Selvarangan, S Hannah Shirley, Natasha Halasa, Julie A Boom, Janet A Englund, Peter G Szilagyi, Christopher J Harrison, et al. Innate susceptibility to norovirus infections influenced by fut2 genotype in a united states pediatric population. Clinical Infectious Diseases, 60(11):1631–1638, 2015.

[176] Nada M Melhem. Norovirus vaccines: correlates of protection, challenges and limitations. Human vaccines & immunotherapeutics, 12(7):1653–1669, 2016.

[177] Aron J Hall, Jan Vinjé, Benjamin Lopman, Geun Woo Park, Catherine Yen, Nicole Gregoricus, and Umesh Parashar. Updated norovirus outbreak management and disease prevention guidelines. Morbidity and Mortality Weekly Report: Recommendations and Reports, 60(3):1–15, 2011.

[178] L Barclay, GW Park, E Vega, A Hall, U Parashar, J Vinjé, and B Lopman. Infection control for norovirus. Clinical microbiology and infection, 20(8):731–740, 2014.

[179] Matthew James Keeling and Joshua V Ross. On methods for studying stochastic disease dynamics. Journal of The Royal Society Interface, 5(19):171–181, 2008.

[180] Donald B Rubin. Using the sir algorithm to simulate posterior distributions. Bayesian statistics, 3:395–402, 1988.

[181] Solomon Kullback and Richard A Leibler. On information and sufficiency. The annals of mathematical statistics, 22(1):79–86, 1951.

[182] Nors. `https://www.cdc.gov/NORS/about.html`. Accessed: 2018-08-23.

[183] Nors. https://www.cdc.gov/nors/downloads/guidance.pdf. Accessed: 2018-11-23.

[184] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2016.

[185] Robert L Atmar, Antone R Opekun, Mark A Gilger, Mary K Estes, Sue E Crawford, Frederick H Neill, and David Y Graham. Norwalk virus shedding after experimental human infection. Emerging infectious diseases, 14(10):1553, 2008.

[186] Katharina Verhaelen, Martijn Bouwknegt, Froukje Lodder-Verschoor, Saskia A Rutjes, and Ana Maria de Roda Husman. Persistence of human norovirus gii. 4 and gi. 4, murine norovirus, and human adenovirus on soft berries as compared with pbs at commonly applied storage conditions. International journal of food microbiology, 160(2):137–144, 2012.

[187] Rachel M Lee, Justin Lessler, Rose A Lee, Kara E Rudolph, Nicholas G Reich, Trish M Perl, and Derek AT Cummings. Incubation periods of viral gastroenteritis: a systematic review. BMC infectious diseases, 13(1):446, 2013.

[188] Barry Rockx, Matty de Wit, Harry Vennema, Jan Vinjé, Erwin de Bruin, Yvonne van Duynhoven, and Marion Koopmans. Natural history of human calicivirus infection: a prospective cohort study. Clinical Infectious Diseases, 35(3):246–253, 2002.

[189] KAM Gaythorpe, Caroline Louise Trotter, B Lopman, M Steele, and AJK Conlan. Norovirus transmission dynamics: a modelling review. Epidemiology & Infection, 146(2):147–158, 2018.

[190] Henry Deng and Hadley Wickham. Density estimation in r. Electronic publication, 2011.

[191] Tarn Duong et al. ks: Kernel density estimation and kernel discriminant analysis for multivariate data in r. Journal of Statistical Software, 21(7):1–16, 2007.

[192] M Elizabeth Halloran and Claudio J Struchiner. Causal inference in infectious diseases. Epidemiology, pages 142–151, 1995.

[193] Ian Shrier and Robert W Platt. Reducing bias through directed acyclic graphs. BMC medical research methodology, 8(1):70, 2008.

[194] James M Robins. Data, design, and background knowledge in etiologic inference. Epidemiology, 12(3):313–320, 2001.

[195] Hans Heesterbeek, Roy M Anderson, Viggo Andreasen, Shweta Bansal, Daniela De Angelis, Chris Dye, Ken TD Eames, W John Edmunds, Simon DW Frost, Sebastian Funk, et al. Modeling infectious disease dynamics in the complex landscape of global health. Science, 347(6227):aaa4339, 2015.

[196] Christina E Mills, James M Robins, and Marc Lipsitch. Transmissibility of 1918 pandemic influenza. Nature, 432(7019):904, 2004.

[197] Dennis T Villareal, Caroline M Apovian, Robert F Kushner, and Samuel Klein. Obesity in older adults: technical review and position statement of the american society for nutrition and naaso, the obesity society. Obesity, 13(11):1849–1863, 2005.

[198] M Sue Kirkman, Vanessa Jones Briscoe, Nathaniel Clark, Hermes Florez, Linda B Haas, Jeffrey B Halter, Elbert S Huang, Mary T Korytkowski, Medha N Munshi, Peggy Soule Odegard, et al. Diabetes in older adults. Diabetes care, 35(12):2650–2664, 2012.

[199] Richard Edwards, Kristie Carter, Jo Peace, and Tony Blakely. An examination of smoking initiation rates by age: results from a large longitudinal study in new zealand. Australian and New Zealand journal of public health, 37(6):516–519, 2013.

[200] Barbara A Forey, Alison J Thornton, and Peter N Lee. Systematic review with meta-analysis of the epidemiological evidence relating smoking to copd, chronic bronchitis and emphysema. BMC pulmonary medicine, 11(1):36, 2011.

[201] Yiqing Song, Anna Klevak, JoAnn E Manson, Julie E Buring, and Simin Liu. Asthma, chronic obstructive pulmonary disease, and type 2 diabetes in the women's health study. Diabetes research and clinical practice, 90(3):365–371, 2010.

[202] Centers for Disease Control, Prevention (CDC, et al. Chronic obstructive pulmonary disease among adults–united states, 2011. MMWR. Morbidity and mortality weekly report, 61(46):938, 2012.

[203] Sumbul Ali and Jose M Garcia. Sarcopenia, cachexia and aging: diagnosis, mechanisms and therapeutic options-a mini-review. Gerontology, 60(4):294–305, 2014.

[204] National Center for Health Statistics et al. Worktable 23r death rates by 10-year age groups: United states and each state, 2007, 2010.

[205] David G Kleinbaum, Lawrence L Kupper, and Hal Morgenstern.

Epidemiologic research: principles and quantitative methods. John Wiley & Sons, 1982.

[206] Sasirekha Ramani, Mary K Estes, and Robert L Atmar. Correlates of protection against norovirus infection and disease—where are we now, where do we go? PLoS pathogens, 12(4):e1005334, 2016.

[207] Lisa Lindesmith, Christine Moe, Severine Marionneau, Nathalie Ruvoen, XI Jiang, Lauren Lindblad, Paul Stewart, Jacques LePendu, and Ralph Baric. Human susceptibility and resistance to norwalk virus infection. Nature medicine, 9(5):548, 2003.

[208] Joël Mossong, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS medicine, 5(3):e74, 2008.

[209] Noelia Carmona-Vicente, Manuel Fernández-Jiménez, Juan M Ribes, Carlos J Téllez-Castillo, Parisá Khodayar-Pardo, Jesús Rodríguez-Diaz, and Javier Buesa. Norovirus infections and seroprevalence of genotype gii. 4-specific antibodies in a spanish population. Journal of medical virology, 87(4):675–682, 2015.