Data-driven Approaches to Understand the Implications of Social Processes for Infectious Disease Risk

by

Ramya Naraharisetti

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in the University of Michigan
2023

Doctoral Committee:

Professor Joseph NS Eisenberg, Co-chair
Associate Professor Jon Zelner, Co-chair
Associate Professor Philip S Boonstra
Associate Professor Alexis J Handal

Ramya Naraharisetti
rsetti@umich.edu
ORCID iD: 0000-0001-7324-8195

DEDICATION

This dissertation is dedicated to the perfect exemplar: ‘Abdu’l-Bahá.

ACKNOWLEDGMENTS

Within the field, I would like to acknowledge my committee: Drs. Jon Zelner, Joseph Eisenberg, Alexis Handal and Phil Boonstra as well as several other mentors who were pivotal to the writing of this thesis including Drs. Marcia Caldas de Castro, Gwen O. Lee, Jim Trostle, Marisa C. Eisenberg, Emily T. Martin and Bruce G. Link. Their commitment to serving students with such a high degree of morality and ethics is an example for future scholars in the field. These analyses would not be possible without the collaborative environment in both the EpiBayes and EcoDess labs. Specifically those who supported with data collection and management: Dr. Krzysztof Sakrejda, Dr. Robert Neale Trangucci, Dr. Nina B. Masters, Dr. Michael AL Hayashi, Ms. Jessica Uruchima along with Ms. Andrea Paola Anchundia Ortega and Mr. Willian Cevallos in Ecuador. My growth as a doctoral student would not be possible without a department of wonderful educators, staff and students including: Dr. Nancy L. Fleischer, Dr. Linda Lysabeth, Ms. Elvira Rivera, Dr. Delvon Mattingly and Dr. Mia Peng. Within my family, I would like to acknowledge the support I received from my bedrock of a family: My steadfast and kind husband Tucker D. Myers, my son who I gave birth to at the beginning of my journey as a doctoral student, Ezme Kazim MyerSetti, our little soul growing in my womb who accompanied me throughout the last months of writing and defending, my mother, Sudha Naraharisetti, and sister, Sruthi Naraharisetti along with her partner Neal Anderson, my in-laws Karen and Joel Myers, my late father who absolutely loved learning, Gautam Naraharisetti and my late uncle Ramesh Pentapati, who passed away during the Delta variant wave of the COVID-19 pandemic in India, who both inspire me to pursue questions of justice and health. And most importantly, the coherence I receive from studying, acting and reflecting on the Bahá'í Writings alongside others who are looking to build a new more just world.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

LIST OF ACRONYMS

COVID-19 Coronavirus Disease 2019

IR incidence rate

CFR case fatality ratio

SIR Susceptible Infected Recovered

WHO World Health Organization

WaSH water sanitation and hygiene

RCTs Randomized Control Trials

MDSS Michigan Disease Surveillance System

MDHHS Michigan Department of Health and Human Services

NPI Non-Pharmaceutical Intervention

PCR Polymerase Chain Reaction

ACS American Community Survey

US United States

CrI credible interval

NCDs Non-Communicable Diseases

TB Tuberculosis

STIs Sexually Transmitted Infections

GPS Geographic Position System

GIS Geographic Information Systems

LMICs low-middle income countries

EcoDess Environmental Change and Diarrheal Disease

GIM General Important Matters

PV People Visited

IWD individual with disabilities

ID Infectious Disease

ABSTRACT

Generally, social determinants or social factors have been studied in the context of non-communicable diseases yet there is a critical need to elucidate the mechanisms behind social and environmental factors which affect communicable disease spread in different spatio-temporal and geographic contexts with the explicit purpose of informing policy, intervention and study design. Social determinants of health frameworks are useful for infectious diseases but lack the ability to translate said mechanisms onto processes that are specific to pathogen spread such as environmentally mediated, vector-borne and direct transmission, transmission cycles and exposure routes in a meaningful way. The focus of most transmission models which investigate social factors, have been within long-term infections and far fewer studies exist of social factors which drive acute infections such as influenza and enteric disease. Given the nature of health inequities within the SARS-CoV-2 pandemic recently, there has been a renewed call for social variables to be incorporated into the study of infectious disease to generate a coherent framework. What is lacking then is an overarching framework which can account for the unique ways social and environmental factors impact systems of pathogenic spread across long-term and acute infections. The overarching goal of my research is to elucidate the mechanisms behind social and environmental factors which affect pathogen spread in different spatio-temporal and geographic contexts to inform policy, intervention and study design to inform such a framework. Each aim within this thesis delves into the heterogeneity of social and environmental mechanisms over time and across space. The first paper emphasizes the role of changing social mechanisms over time to examine COVID-19 disparities. The second utilizes space as a dimension of variability and seeks to identify clusters of infectious disease outcomes across two time periods. The third adds more complexity and asks how social context and spatial context interact to increase infectious disease risk. Each of my papers also delves into the heterogeneity of social mechanisms across different geographic contexts. My second paper looks at social mechanisms at the marco level within the low-middle income country of India, my first paper looks at meso-level mechanisms in the state of Michigan within the United States and the last paper looks at micro-level mechanisms within six communities in the northwestern region of Ecuador. Overall, the exploration of social mechanisms over time, spatial mechanisms over time and the interaction of social and spatial mechanisms across different geographic contexts advances the concept of a socially and environmentally informed framework for

infectious disease systems.

CHAPTER 1

Introduction

1.1 What are the social determinants of infectious disease risk?

Historically, the social determinants of health have been studied primarily with respect to Non-Communicable Diseases (NCDs) (Figure 1.1), with comparatively little theoretical work focused on the social patterning of infection outcomes. Nonetheless, infectious diseases are the leading cause of death in low-income countries, with lower respiratory infections and diarrheal disease topping the list [5]. In middle and high income countries, infectious disease continues to concentrate in lower resourced and vulnerable populations. The most common and accepted framework for understanding infectious disease transmission is that of the dynamic Susceptible Infected Recovered (SIR) model and its variants. There is a long history of work on the social determinants of certain infectious diseases within a dynamic framework such as with Sexually Transmitted Infections (STIs) and HIV/AIDS in particular [6–9], Tuberculosis (TB) [2, 10–12] and diarrheal diseases [13, 14] such as cholera [15, 16]. What is lacking is a well articulated framework for the social determinants of infectious disease which translates the great advances over the past several decades, within the study of social determinants for NCDs [17], into the context of infection outcomes. Given the nature of health inequities within the SARS-CoV-2 pandemic recently, there has been a renewed call for social variables to be incorporated into the study of infectious disease to generate a coherent framework [18–20]. In this dissertation, I pursue three research questions which seek to advance our understanding of social factors in dynamic systems of infectious disease. In this introduction I share two frameworks, related to COVID-19 and enteric disease respectively, to serve as a basis for the research questions I pursue in Aims 1-3.

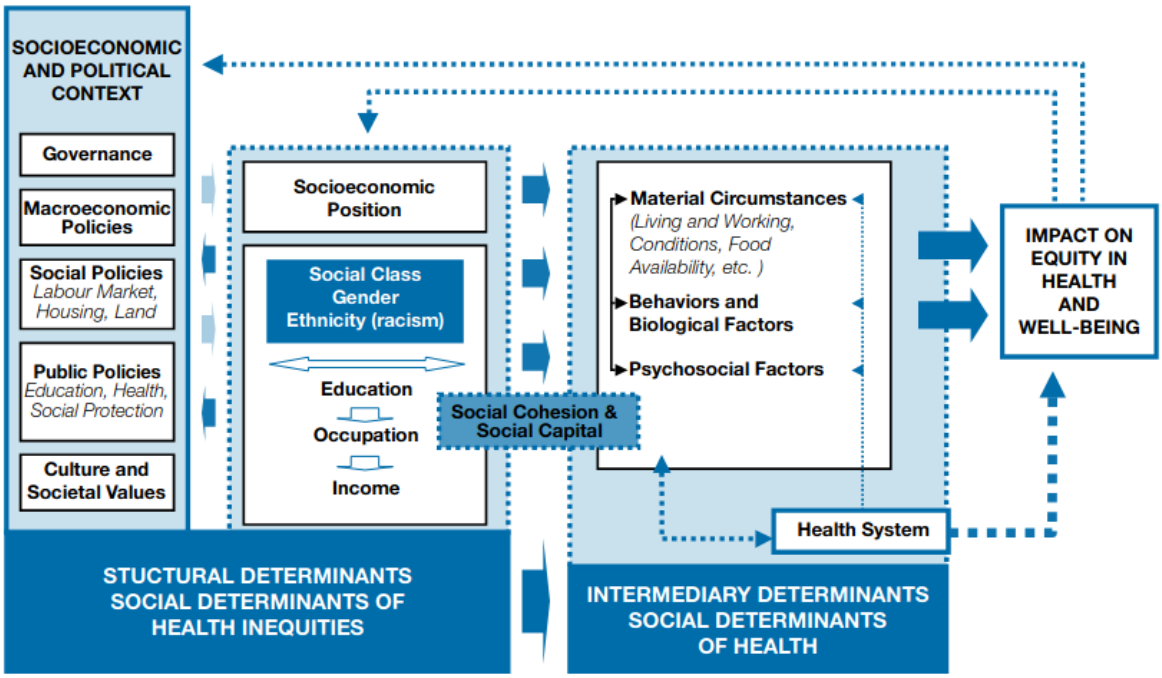

Figure 1.1: World Health Organization (WHO) updated framework for the social determinants of health [1]

.

1.1.1 Aim 1: Have the determinants of SARS-CoV-2 mortality disparity changed over the course of the COVID-19 pandemic?

As of early 2023, SARS-CoV-2 infections have resulted in 761 million cases and 6.9 million deaths since the beginning of the global pandemic [21]. Social factors have come to the fore as important predictors of morbidity and mortality. The role of social factors in the spread of SARS-CoV-2 over time in 2020 in the state of Michigan within the United States (US) is pursued in Aim 1. Primarily, we identify that privileged groups are able to practice Non-Pharmaceutical Intervention (NPI)s early in the pandemic and therefore only experience high incidence once the clinical management of COVID-19 has greatly improved. If Black residents of Michigan were afforded the same privilege, 35 % of deaths within this group could have been avoided. In developing this research question, it was important to have a working framework for how structural racism could function within the spread of SARS-CoV-2. The advanced understanding of social factors in disease transmission models within TB [2] research were helpful in generating a similar working framework for SARS-CoV-2 (Figure 1.2). Parameters from Dr. Marisa Eisenberg's mathematical model for COVID-19 in Michigan [3] provided a skeleton for the dynamics portion of the framework. The various elements of the framework (organized in Table 1.1) are based on an informal literature review of structural racism, socio-economic status and COVID-19 conducted in 2021. This conceptualization served as the basis for the research questions pursued in Aim 1 (Chapter 3).

1.1.2 Aims 2: Identifying population-level trends for public health planning

Collecting infectious disease surveillance data in low-middle income countries (LMICs) can be especially difficult given resource constraints, yet high levels of morbidity and mortality are concentrated within the large proportion of impoverished population within such countries. While, population-level data is difficult to collect, most data sets utilize cluster sampling processes to estimate infectious disease parameters within a nation or state but often these estimates lack the granularity to make decisions at the local level. Census data allows us to study population-level processes without such estimation and modeling. The difficulty of census data although is the limited set of questions most governments ask of their populations. One set of questions which give us insight into infectious disease processes are those related to disability. Disability, especially in LMICs are often driven by infectious disease processes and details related to the type of disability provide insight into which pathogens might be affecting large portions of the population. In this paper we utilize a socially and environmentally informed framework of infectious disease to advance a way of analysing infectious disease data in resource constrained settings. We apply spatial analysis approaches to census data on disability across two data collection waves to identify clusters of certain disability types, such as hearing and sight impairment, which are related to

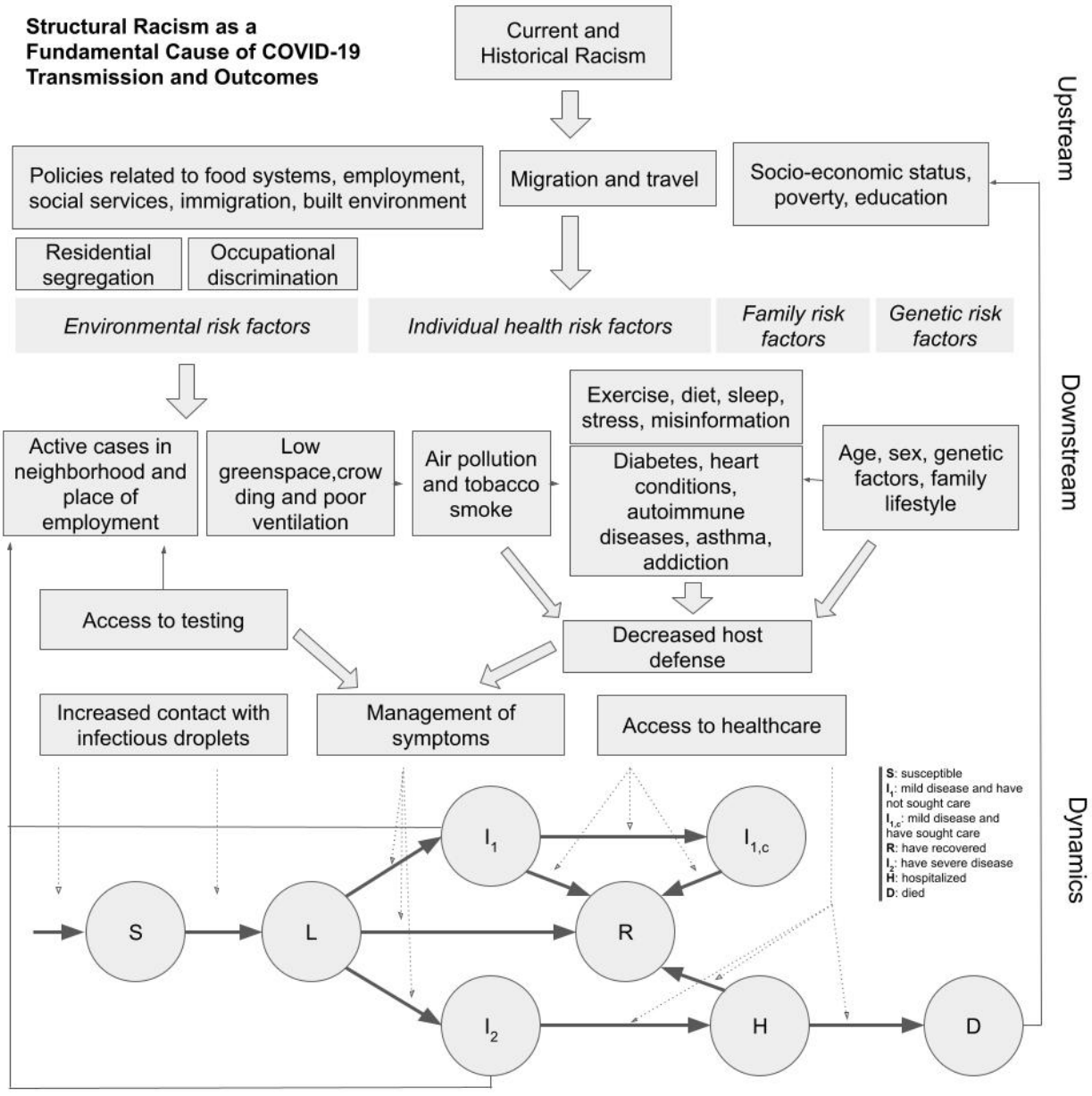

Figure 1.2: Adaptation of Pedrazzoli et al. 2017 [2] to the SARS-CoV-2 pandemic utilizing parameters from Dr. Marisa Eisenberg's [3] mathematical model for COVID-19 in Michigan. This framework illustrates the upstream, downstream and transmission dynamics from a social determinants perspective. The factors within this framework are based on a literature review of structural racism, socio-economic factors and Coronavirus Disease 2019 (COVID-19) in 2021.

Type	**Determinant**
Upstream	- Racism within social, economic and political systems - Migration, community patterns and travel - Economic, social and environmental policies - Health system access - Food system access - Employment policies - Transportation system access - Poverty/low socio-economic status
Downstream *Environmental*	- Active cases in immediate community - Cases at place of employment or commuting locations - Tobacco smoke, air pollution - Crowding in household
Downstream *Behavioral*	- Pre-existing conditions - Social distancing - Mask wearing
Downstream *Genetic*	- Age - Sex - Genetic factors leading to co-morbid conditions
Transmission Recovery Dynamics	- Impaired immune defense - High level contact with infectious droplets - Access to healthcare (including testing)

Table 1.1: Types of social determinants within a structural racism framework for COVID-19. This list was generated after a literature review of publications on COVID-19 and social determinants in 2021.

congenital and post-natal diseases. The identification of statistically significant clusters allow us to inform targeted prevention and programming in low-resource settings.

1.1.3 Aims 3: Addressing key measurement issues in the social determinants of acute infection risk

Respiratory illness and diarrheal disease are the top two leading causes of death among children worldwide [1]. Enteric disease and respiratory illness is primarily thought to be caused by environmental and social factors [4, 22–25]. An examination of the role of social factors—as measured by geographic and social network variables—in the spread of enteric disease is pursued in Aim 2. We ask the question of whether spatial density (household geographic closeness) and social density (number of contacts) interact to create multiplicative and/or additive interaction. This research helps bring together the environmental, behavioral and social factors related to disease spread and seeks to advance a more coherent understanding of their interplay at the individual level.

To characterize the social context of enteric disease spread (as an example) we conducted an informal literature search in PubMed of MeSH terms to determine published articles on "socioeconomic factors" and "dysentery" and located 121 results. None of these articles had mathematical models and all of them used empirical data to draw conclusions about socioeconomic factors and their relationship to various outcomes related to diarrheal disease. From a synthesis of this literature we were able to generate a list of factors which were used in regression models (they are highlighted in Table 1.2 below). For example, the role of community social connectivity [23], poverty in the U.S [26, 27] and governmental revenue [28] seem to be associated with enteric disease outcomes.

We conceptualize how the social determinants of enteric disease could inform dynamic models by generating a working framework, see Figure 1.3. This framework is adapted from Pedrazzoli et al. 2017 and water sanitation and hygiene (WaSH) factors highlighted in Einsenberg et al. 2006. This framework is not meant to be comprehensive but rather illustrates that social determinants could plausibly play a role in disease transmission dynamics. It highlights the complex interplay between social, behavioral and environmental factors which may cause diarrheal disease in a child or adult. Overall this conceptual model illustrates dynamics which could be taken into account when considering a mathematical model of diarrheal disease which includes social factors. This framework also highlights which factors may be relevant for empirical studies of enteric disease and social determinants or Randomized Control Trials (RCTs) of WaSH. This conceptualization served as the basis for the research questions pursued in Aim 2 (Chapter 5)).

Type	**Determinant**
Upstream	- Economic, social and environmental policy which disadvantages the poor - Globalization causing unfair labor practices - Migration patterns for employment - Health system access - Human and social capital - Community social connectivity - Transportation system access - Poverty/low socio-economic status
Downstream *Environmental*	- High levels of environmental pathogens - Active cases in community - Community hygiene and sanitation
Downstream *Behavioral*	- Inappropriate health seeking - Within household sanitation and hygiene behavior - Breastfeeding
Transmission Recovery Dynamics	- Impaired immune defense - High level contact with infectious droplets - Management of symptoms - Access to healthcare

Table 1.2: Types of social determinants within an enteric disease framework with social determinants of health. This list was generated after a literature review of publications on enteric disease and social determinants in 2021.

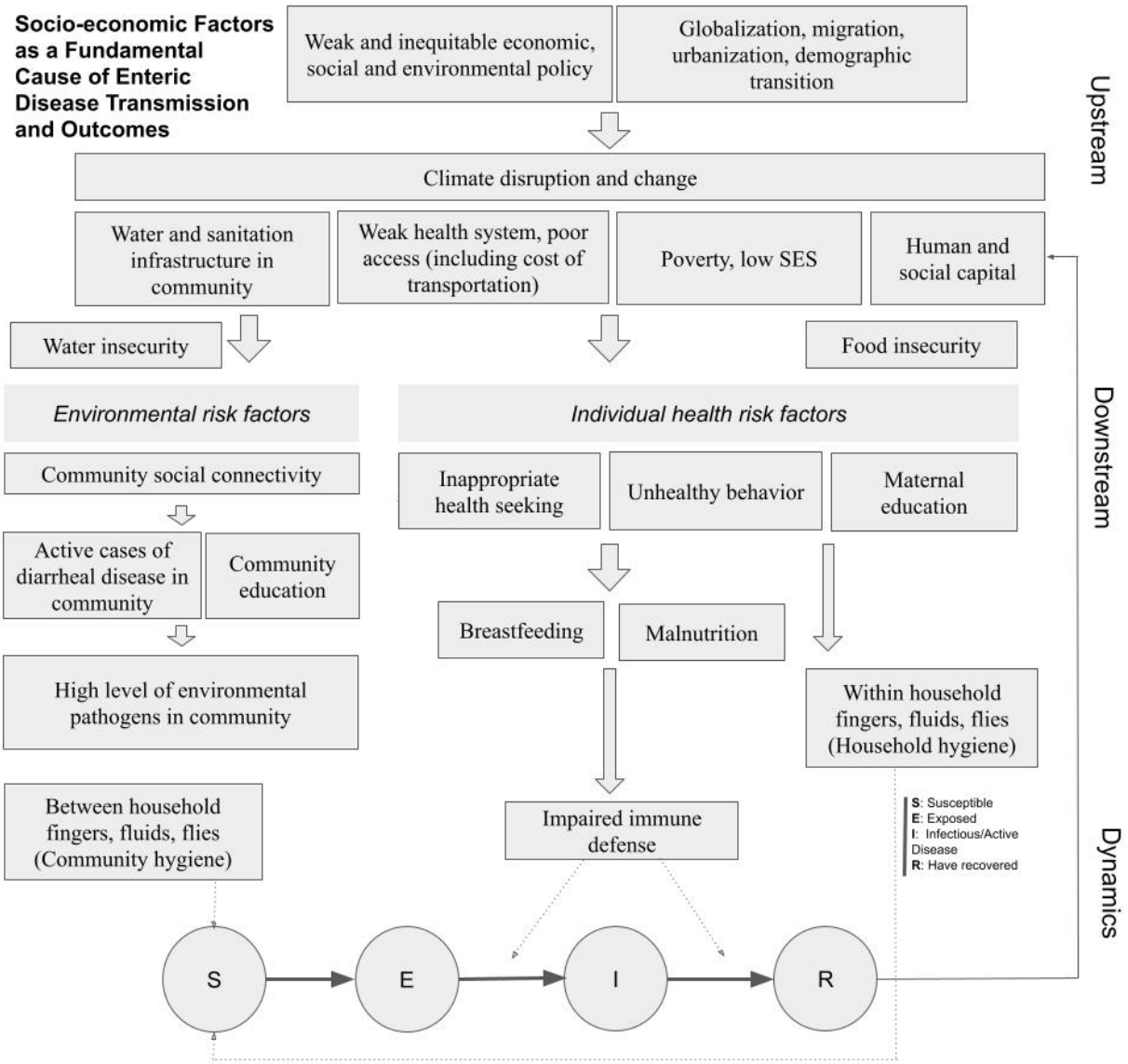

Figure 1.3: This framework is an adaption of Pedrazzoli et al. 2017 [2] to the enteric disease context with some factors from Eisenberge et al. 2006 [4]. This framework illustrates the upstream, downstream and transmission dynamics from a social determinants perspective. The parameters for this framework are based on a literature review of socio-economic factors and enteric diseases outcomes in 2021.

1.2 Conclusion

In conclusion, social determinants of infectious disease are complex and specific to a disease context. Further, they are defined by the particular social, economic, political and historical realities of a population under study. It has become clear although—especially after the SARS-CoV-2 pandemic—that social factors are not merely relegated to traditionally social infectious diseases but rather play a vital role in our understanding of transmission dynamics for many, if not all, communicable disease. Yet, the lack of an overarching framework for social determinants of infectious disease makes it challenging to synthesize literature around this topic. Conceptualizing social determinants of health within a disease-specific dynamic framework is therefore helpful in the formation of research questions. Though various forms of data have been introduced in the recent past to help answer questions around social factors and infectious disease, including detailed demographic data, mobility data and geographic surveillance, conceptual models of how social factors affect disease dynamics can prevent the inappropriate use of such data to advance purely behavioral and cultural explanation of infectious disease disparities without the consideration of structural and economic factors [20, 29]. In the following aims, I attempt to answer several questions using these frameworks as the basis for data analysis and research question formation.

CHAPTER 2

Specific Aims

2.1 Introduction

The COVID-19 pandemic has underscored the urgent need to better understand how social processes interact with the biology of infectious disease transmission to create the patterns of infection and mortality we see in the world. Although social and economic factors are understood to impact (either positively or negatively) infection risk, the mechanisms by which these occur are under-studied and therefore fail to properly inform randomized trial design, policy and programming. My proposed dissertation is organized into three projects, outlined below. Each project is focused on elucidating the mechanisms behind social factors which affect pathogen spread in different social and geographic contexts.

2.2 Aim 1: Timing of infection was a key driver of race/ethnic disparities in COVID-19 mortality during the pre-vaccine period

Introduction Disparities in COVID-19 mortality rates are driven by inequalities in group-specific incidence rates (IRs) and case fatality rates (CFRs), and their interaction. For emerging infections, such as SARS-CoV-2, group-specific IRs and CFRs change on different timescales, and inequities in these measures may reflect different social and medical mechanisms. To be useful tools for public health surveillance and policy, analyses of changing mortality rate disparities must independently address changes in IRs and CFRs. However, this is rarely done.

Objective In this analysis, we examine the separate contributions of disparities in the timing of infection - reflecting differential participation in essential work and other infection risk factors - and declining CFRs over time on mortality disparities by race/ethnicity in the U.S. state of Michigan.

Methods We used detailed case data to decompose race/ethnic group-specific mortality rates into their age-specific incidence and CFR components during each of three periods from March-December 2020.

Results We used these estimates in a counterfactual simulation model to estimate that 35% (95% Credible Interval (CrI)=30%-40%) of Black resident deaths could have been prevented if Black Michigan residents were infected at the same pace as White residents, resulting in a 67% (95% Credible Interval (CrI)=61%-72%) reduction in the mortality rate gap between Black and White Michigan residents during 2020.

Conclusion These results clearly illustrate why differential power to 'wait out' infection during an infectious disease emergency – a function of structural racism - is a key, underappreciated, driver of inequality in disease and death from emerging infections.

2.3 Aim 2: Spatio-temporal patterns of Disability in India: 2001-2011

Introduction Eighty percent of individuals with disability (IWD) live in the global South, and India has one of the largest concentrations. The need to understand the spatial distribution of disability in a low- and middle-income country (LMIC) is twofold: (1) to identify potential sources of environmental, social, and economic causes of disease, and (2) to characterize the socio-economic and medical needs of IWD to inform treatment, programming, and policy. **Objective** Identify spatial clustering of disability in India, as well as change in these patterns over the period from 2011-2011.

Methods We used the 2001 and 2011 Population Census of India and apply population weights at the district-level to compare across years. Local indicators of spatial autocorrelation with correction for multiple testing allow for the identification of high prevalence clusters throughout the country for each of the five categories of disability: speech, sight, hearing, mental (cognitive), and movement (physical).

Results All types of disability are experienced heterogeneously across space in terms of both current and increasing prevalence. We observed particularly large and growing clusters of speech impairment in western-central India, and movement disability in eastern-central India.

Discussion Given the rapid urbanization and industrialization at the local level, temporal changes in spatial patterns of disability at the district level should inform disability policy both nationally and regionally in LMICs. We illustrate that there are noteworthy changes over time in terms of where high prevalence disability clusters are located in India both for disability overall and across specific types of disability.

Conclusion The spatial heterogeneity of current and increasing disability prevalence in India illustrates the need for further research into the identification of such large clusters of specific disability types in other LMICs. Country-level identification of local clusters could help prioritize funding and shape the agenda of disability research, programming, and policy at the national and regional levels.

2.4 Aim 3: Do sociality and spatiality interact? Characterizing the joint impact of spatial proximity and social network connectivity on diarrheal disease risks and intervention efficacy

Introduction Diseases that are spread from person-to-person and via environmental intermediaries (such as enteric diseases) induce a shared geographic risk space between contacts. Within(LMICs, especially those with lower population densities, distance remains a relevant measure of changing environmental context.

Data/Methods Our analysis utilizes recently collected data from a long-standing sociometric study of enteric disease outcomes in western Ecuador. Using both a contact network and core-discussion network, we generate a metric of combined social and spatial density which we denote as *contact density*. This measure serves as an interaction term of social and spatial density and allows us to capture the additional infection risk when both exposures are present.

Results Social density and spatial density predict enteric disease outcomes and display interaction on the multiplicative and additive scales. These variables display effect modification by community social structure (low density versus high density networks).

Discussion Our analysis illustrates that, in relative risk terms, individuals who live in dense neighborhoods with their social contacts are at highest risk for disease low density social network communities (peripheral communities). In absolute risk terms, our analysis shows that efforts to reduce the effects of spatial density on enteric outcomes will see the most benefit among individuals with fewer social contacts.

Conclusion Social and spatial environments interact to create risk and advancing our understanding of this interaction could lead us to improvements in the form of enteric disease morbidity and mortality.

CHAPTER 3

Timing of Infection was a Key Driver of Race/ethnic Disparities in COVID-19 Mortality During the Pre-vaccine Period

3.1 Introduction

A core principle of social epidemiology has been realized in the emergence of SARS-CoV-2 in the United States and around the world. Transmission of a virus that is agnostic about who it might infect has nevertheless been powerfully shaped by the social circumstances it travels within, resulting in robust social inequalities in morbidity and mortality [30–32]. Lying beneath this general principle is a more specific puzzle: exactly how did this social shaping of pandemic morbidity and mortality occur [32]. In this analysis we show that disparities in SARS-CoV-2 outcomes can only be properly understood as the confluence of two critical factors: 1) when in the epidemic members of marginalized populations were at the greatest risk of infection and, 2) how likely were individuals belonging to these groups to die when these infections occurred?

We use detailed individual-level incidence and mortality data from the state of Michigan to estimate age- and race/ethnic group-specific per-capita COVID-19 incidence and case-fatality rates1 over three critical periods during 2020. The core proposition of this analysis is that gaps in mortality between race/ethnic groups during the pre-vaccination period of the COVID-19 pandemic were driven by differences in infection risk early in the pandemic, when life-saving knowledge was sparse and population-wide case-fatality rates were high. These differences in exposure reflect inequalities in social power often [33], though by no means exclusively [29], reflected along lines of race and ethnicity.

Michigan has a large and diverse population and is among the most racially segregated and economically unequal states in the U.S. [34] Michigan also implemented one of the strongest non-pharmaceutical intervention (NPI) policies early in the pandemic, reaching an Oxford Stringency Index (OSI) score of 50, second only in the U.S. to Delaware [35]. It also had a nearly seven-

fold disparity in all-cause excess mortality for Blacks vs. White in 2020, the largest recorded in the U.S [36]. This combination of a comprehensive approach to limiting community spread in a context characterized by stark race/ethnic inequality make Michigan an ideal context in which to investigate the mechanisms by which structural racism contributes to unequal outcomes even under a 'best case' scenario of public health response.

Most analyses of racial/ethnic disparities in COVID-19 mortality have been cross-sectional, focusing on a single period in which policy and medical responses to SARS-CoV-2 infection remained relatively consistent [37] or cumulative infection and death from March 2020 onward [38]. They often draw on the differential role of inequity within stages of susceptibility, exposure and recovery [39, 40]. Some of these analyses show that large racial and socioeconomic disparities in COVID-19 mortality observed in the U.S. are attributable to differential rates of infection, while inequity in case-fatality rates (CFRs) stemming from comorbidities, such as obesity or heart disease and differential access to care, played an important but less pivotal role [37, 41, 42]. These findings ran counter to early scientific and media speculation that inequities in mortality were likely attributable to disparities in comorbidities and care that increased the risk of death following SARS-CoV-2 infection [43].

Longitudinal analyses have shed some light on the dynamics of inequity over the course of the pandemic: Van Dyke et al. [44] analyzed national data covering the period from Summer 2020 to Winter 2021. They determined that the declining relative risk for young minoritized populations ($<$25 years), compared to same-aged White populations, primarily reflected growth in the rate of infection and death among White populations rather than sharp decreases among minoritized populations. This suggests that shrinking disparities throughout 2020 resulted from worsening pandemic conditions rather than reflecting positive policy outcomes. Aschmann et al. [45] recently analyzed all-cause mortality by racial-ethnic group over the past four years and observed that Indigenous groups continue to experience levels of disparity higher than their pre-pandemic levels. Their findings highlight the importance of longitudinal analyses when examining and identifying disparities caused by systems of exposure or structural causes [46].

In this paper, we add to the understanding of the dynamics of inequity in infection and mortality by examining changes in incidence and mortality across three periods during the pre-vaccine phase of the COVID-19 pandemic in 2020. These represent distinct phases of the early pandemic and the policy response to COVID-19 in Michigan:

Period 1: Emergence of SARS-CoV-2 and the most-restrictive lockdown interval (March–June 2020) with an OSI of 73 out of 100. This period includes sparse personal protective equipment (PPE), low supply of ventilators and pulse oximeters, minimal medical knowledge about treatment protocols and marginalized groups more likely to be in essential work categories – meat packing, grocery etc. [47–49] The pandemic hit marginalized communities hard without protection against

death, while other non-marginalized communities benefited from stay-at-home orders for most office-based workplaces [48–52]. For example, an analysis of essential frontline workers in Census data and a US Department of Labor dataset called the Occupational Information Network revealed that over 70% of male workers from one of the following minoritized groups: Latino, Black, Native American, and Pacific Islander were considered frontline [51].

Period 2: Summer of moderately eased contact restrictions accompanied by a suite of NPIs including indoor mask mandates and social distancing policies [53] reflecting a 27% decrease in the OSI measure. This period included incidence spread to other non-marginalized groups, protective gear becoming more available, medical knowledge growing and ventilator supply increased to prevent death in all groups.

Period 3: Further easing of restrictions in the fall (September–December 2020) reflecting a further 21% decrease in the OSI measure. This period includes the availability of the first FDA approved antiviral drug Veklury (remdesivir) for widespread use [54], lifting of nearly all stay-at-home orders for most office-based workplaces, widespread high community transmission within all groups, increased availability of PPE to the general population along with limited vaccination roll-out to high-risk populations beginning in December 2020.

Our goals in this analysis are twofold: 1) characterize disparities in incidence and mortality during each of these periods and 2) estimate the impact of each period on overall patterns of mortality inequality throughout the period from March-December 2020.

Incidence and mortality data. SARS-CoV-2 infection and mortality data were obtained from the Michigan Disease Surveillance System (MDSS) maintained by the Michigan Department of Health and Human Services (MDHHS). The current analysis includes all probable and polymerase chain reaction (PCR)-confirmed SARS-CoV-2 cases recorded in MDSS from March 8-December 31, 2020. Probable cases were determined using the Michigan State and Local Public Health COVID-19 Standard Operating Procedures [55]. Deaths were attributed when COVID-19 was the primary or secondary cause of death on an individual's death certificate. Further information about these data can be found in [37]. For all cases, we obtained the date of case identification, age, sex, and racial/ethnic category. This dataset consists of 508,648 probable and PCR-confirmed COVID-19 cases and 13,078 COVID-19 deaths. We binned case data into 10-year age groups, with individuals 80+ in one group. Race of cases was categorized to match census population groups including Black/African American, Asian/Pacific Islander, Native American, White, and Other (this comprised the census category of "Other" and those who identified as two or more racial/ethnic categories). In Michigan, race (e.g. African-American, Caucasian) and ethnicity (e.g. Latino) were classified separately, like the US Census approach to data collection. In this analysis Latino ethnicity was categorized as any race with Latino ethnicity, therefore this group includes Black/African American Latinos, White Latinos, etc. and those individuals were not included in

their primary race category. Our race and ethnicity groups are therefore mutually exclusive. Please see Appendix Table A.2 for the distribution of races within Latino ethnicity. Therefore, we use the term race/ethnicity to indicate that we are observing differences in terms of race categories and the Latino ethnicity category. We obtained statewide estimates of population size stratified by age, sex, and race/ethnicity using data from the 2018 American Community Survey (ACS) [56].

When comparing risks, we used the White population as the reference group and look at contrasts for each of the racial-ethnic categories identified above for which Census population data were available. Although Michigan has a large Arab-American population, we were unable to include this group in our analyses, the U.S. Census does not recognize Arab populations as a separate race/ethnicity and they are categorized as White [57].

3.2 Methodology

SARS-CoV-2 infection and mortality data were obtained from the Michigan Disease Surveillance System (MDSS)) maintained by the Michigan Department of Health and Human Services (MDHHS). The current analysis includes all COVID-19 cases recorded in MDSS from March 8-December 31, 2020. For all cases, we obtained the date of case identification, age, sex and racial/ethnic category. This data-set consists of 508,648 probable and Polymerase Chain Reaction (PCR)-confirmed COVID-19 cases and 13,078 deaths attributed to COVID-19. Probable cases were determined using the Michigan State and Local Public Health COVID-19 Standard Operating Procedures. [58]

We binned case data into 10-year age groups, with individuals 80+ in a single group. Race/ethnicity of cases was categorized to match census population data for groups including Black/African American, Latino, Asian or Pacific Islander, Native American, White, and Other (this comprised the census category of "Other" and those who identified as two or more racial/ethnic categories). We utilized 2018 American Community Survey (ACS) data from IPUMS [56] to determine demographic counts based on age, sex and race/ethnicity in Michigan. For further details on racial/ethnic categorization.

Our statistical analysis is focused on estimating age- and race/ethnic group-specific incidence rate (IR)s (denoted as λ) and corresponding case fatality ratio (CFR)s (denoted by ρ), which are the composite parts of mortality rates ($\rho\lambda$). Taking this two-step approach allows us to dis-aggregate the relative contribution of the risk of infection versus the risk of death following infection to the period-specific mortality rate for each group.

Incidence Rates For each time period (l)), we estimated per-capita rates of COVID-19 in each age (i), sex (j) and racial/ethnic category (k), using a Poisson regression model with a population offset term, $log(n_{ijk})$, where n_{ijk} is the size of population in each ijk group within the 2018 ACS

data-set. The model included several interaction terms to obtain the full spectrum of potential heterogeneity in the outcome data including age$\times$ sex, sex$\times$ race and race $\times$ age. The observed number of cases in each group are y_{ijk} and the per-capita IR in each bin is denoted by λ_{ijk}. We utilized a direct standardization [59] approach to generate age- and sex-adjusted marginal results in addition to age-specific IRs. This model included a weakly-informative log-Gaussian prior distribution with a mean of zero and standard deviation of 0.1.

Case-fatality Ratios We used a binomial regression model to estimate age-specific CFRs for each group in each time period. Our modeled outcome was the number of deaths (z_{ijk}) as a function of the total number of cases (y_{ijk}) in each age, sex and racial/ethnic stratum, with the CFR for each group denoted as ρ_{ijk}. Therefore, $z_{ijk} \sim Binomial(y_{ijk}, \rho_{ijk})$.

Standardization We present raw and standardized IRs and CFRs in our results to highlight both the aggregate burden of illness and death and to facilitate comparison between groups. IRs were age- and sex-standardized J. Zelner et al. (2020) to reflect group-specific IRs under a scenario of uniform age- and sex distributions across all race/ethnic groups. Standardized CFRs were weighted to present a measure in which the proportion of cases contributed by each racial/ethnic group during 2020 was held constant across time periods to facilitate comparison across epidemic phases [60].

Alternate Infection Timeline Declines in incidence and mortality disparities over time have been presented as indications of progress in addressing causes of disparate infection outcomes during the COVID-19 pandemic [53]. While a narrowing of these gaps may reflect short-term success in reducing inequity, it is critical that we not confuse these changes with substantive improvements to the long-term, structural determinants of infection inequality [20]. During the COVID-19 pandemic, the ability to avoid, or at least delay, infection may have been a function of race/ethnic and socioeconomic privilege, e.g. through the ability to work from home for a sustained period of time. Findings from studies which observed disparity patterns in the first ten weeks of the pandemic highlight this point. Low-income counties experienced much higher rates of infection in the first few weeks of the pandemic compared to high-income counties throughout the US [61]. By April 2020, this relationship had inverted, and wealthier counties began experiencing higher incidence than poor counties. This change coupled with declining CFR across the board over time may conceal an underappreciated dimension of COVID-19 disparity that reflects structural inequalities related to occupational inequality and residential segregation [29, 62, 63].

To assess whether and how much the privilege of waiting to acquire infection later in the pandemic affected disparities in mortality, we developed a counterfactual analysis in which the pacing of infection of White residents was applied to all other racial/ethnic groups while maintaining the overall volume and age/sex distribution of cases. We calculated the difference in the expected number of deaths (y) in each race/ethnic group (i) in the data versus under a counterfactual scenario

(y'), within a period (l), to estimate the proportion of deaths that could have been averted in each race/ethnic group (ζ_i) if all individuals experienced the same infection timeline, as follows:

$$\zeta_i = 1 - \frac{\sum_{l=1}^{n} y'_{ijk}}{\sum_{l=1}^{n} y_{ijk}}$$

Software Bayesian regression analyses were completed in R 4.0.5 using the *rstanarm* package [64], and the *tidybayes* package [65] for post-processing of model results [65]. The analysis pipeline was made reproducible using *Snakemake* [66].

Sensitivity to Missing Data Dropping cases with missing race/ethnicity information, as we did in this analysis, may result in biased estimates if rates of missingness vary across racial-ethnic categories [67]. The need for a sensitivity analysis was informed by a recent analysis of COVID-19 case data that indicated that rates of missing race/ethnicity were in fact higher among African Americans as compared to Whites in early 2020 [68]. To assess the risk that this bias could have impacted both the quantitative and qualitative conclusions of our analysis, we conducted extensive sensitivity analyses. Specifically, we generated numerous synthetic datasets in which race/ethnicity data were dropped for African-Americans and Whites under 1) a baseline scenario of equal likelihood of missingness for both groups (Odds ratio [OR] = 1) as well as scenarios of 2) moderate (OR=2) and 3) extreme (OR=3) differences in missingness for African-Americans as compared to Whites. We also varied the baseline rate of missingness by race/ethnicity to ensure that the impact of these group-specific rates of missingness was robust across different levels of overall missingness. The results of this analysis show that neither the qualitative conclusions nor quantitative results of our analysis are likely to be strongly impacted by non-differential missingness of race/ethnicity data. For additional explanations and figures, please see the Supplement Section II: Sensitivity Analysis (A.6-A.10).

3.3 Results

The MDSS dataset included 68,413, 64,377 and 375,858 probable and PCR-confirmed cases in the first, second and third periods of 2020, respectively. In addition, 6,415, 874 and 5,789 deaths were recorded during the first, second and third period, respectively. Across all three periods, 59,514 cases and 3,173 deaths were among those who identified as African-American or Black, 28,350 cases and 376 deaths were among those identified as Latino, 1,819 cases and 38 deaths were among those identified as Native American or Alaskan Native, 7,392 cases and 134 deaths were among those identified as Asian or Pacific Islander, 291,247 cases and 8,386 deaths were among those identified as White and 20,051 cases and 265 deaths were individuals identified as belonging to any other racial/ethnic group.

This corresponds to unstandardized race/ethnic group-specific COVID-19 mortality rates of 233/100K for Black residents, 75/100K for Latino residents, 82/100K for Native American or Alaskan Native residents, 44/100K for Asian or Pacific Islander residents, 112/100K for White residents and 103/100K for individuals identified as belonging to any other racial/ethnic group. The case fatality rate (CFR) for all SARS-CoV-2 cases reported in Michigan during 2020 was 3

There were 100,275 cases and 706 deaths which did not have an identified racial/ethnic group, which were dropped from subsequent analyses (see Figures A.6-A.10 and Table S11). Cases with missing racial/ethnic group identification varied across periods with 9,564 in the first period, 7,639 in the second period and 83,072 in the third period. This reflects an unstandardized population incidence rate of SARS-CoV-2 infection for all groups during the first period, 232/100K, 217/100K in the second period and 1,277/100K in the third period.

Change in incidence disparities over time. During the first period, age- and sex-standardized analyses show that Black and Latino residents experienced 5.0 (95% CrI = 4.9, 5.1) times and 3.5 (95% CrI = 3.4, 3.6) times the incidence of Whites, respectively (Figure A.1). The values are characterized as incidence rate ratios (IRRs) where the incidence rate of the "exposed" group (in this case those exposed to racism) is compared to that of the "unexposed" group. By the second and third period, most of these differences in incidence were at or near zero.

However, the narrowing of IRRs does not reflect a monotonic decline in incidence. In fact, there is a clear temporal shift in age- and sex-standardized incidence within Black residents from 1,789/100K (95% CrI: 1,764-1,812) in the first period to 497/100K (95% CrI: 486-509) in the second period before rebounding to 2,334/100K (95% CrI: 2,308-2,359) in the last period (Figure 2). A closer look at age-specific incidence in Period 3 among Black residents reveals that there was a marked increase in infection among working age persons—as high as 2,929/100K (95% CrI: 2,849-3,010) among 40–50-year-olds and 3,827/100K (95% CrI: 3,733-3,925) cases among those 30-40-years-old. In contrast, Latino residents experienced a lower age- and sex-standardized incidence of 1,261/100K (95% CrI: 1,223-1,301) that remained relatively constant during the first two time periods, before increasing to 4,085/100K (95% CrI: 4,016-4,157) in the last period. This dramatic increase was not isolated to working-age groups for Latino residents. White and Asian or Islander resident incidence on the other hand was minimal in the first and second period (see Supplement Figure A.1 for age-specific incidence rates in all racial-ethnic categories). White residents experienced incidence levels closer to their Black counterparts during the last period for an age- and sex-standardized rate of 2,946 cases per 100,000 people (95% CrI: 2,934-2,958)

Change in case-fatality rates over time. Age-specific case-fatality rates (CFRs) for all racial/ethnic groups fell sharply after an initial period of high SARS-CoV-2 case-fatality for all groups in the first period (Figure A.3). For example, 50–60-year-old Black residents experienced CFRs of 7.8% (95% CrI = 7.0%-8.7%) in the first period, which decreased to 2.0% (95% CrI =

1.23%-3.0%) and 1.23% (95% CrI = 0.92%-1.58%) in the second and third period, respectively (Figure A.4).

Disparities in the risk of death from COVID-19, as measured by case-fatality risk ratios (CFRRs) were small but did change over time (Figure A.2). For example, Black residents experienced CFRs 1.5, (95% CrI = 1.3-1.8), 1.3 (95% CrI = 1.3-1.4), and 1.5 (95% CrI = 1.3-1.6) times higher than their White counterparts in the first, second and third periods, respectively. However, marked age-specific case-fatality disparities by race/ethnicity were observed across all periods among those 30 and older (see Supplement Figure A.4). For Latino residents, change over time in CFRRs relative to Whites was more dramatic (Figure A.2).

Impact of infection timing on race/ethnic group-specific mortality. The infection trajectory for Black and Latino Michigan residents was different from that of White populations, with these differences in timing most pronounced between Black and White residents (Figure 4). 36.7% of cases among Black residents occurred during the first period, 11.3% in the second period and 52.1% in the third period. By contrast, only 9.6% of cases among Whites occurred during the first period, 13.9% in the second period, with the large majority of 76.5% occurring in the third period when CFRs were at their lowest level.

In our counterfactual simulation analysis, we found that 35% (95% CrI= 30%-40%) of all Black resident deaths could have been avoided under a scenario in which Black residents experienced infection at the same pace as Whites. This corresponds to a 67% (95% CrI=61%-73%) reduction in the difference in per-capita mortality rates between Black and White residents in 2020. Our results also suggest that a portion of Latino deaths (9%; 95% CrI= -5%-21%) could have been prevented as well, although the posterior CrI span the null value of 0. See Supplement Figure A.5 for a more detailed illustration of changes in underlying rates under the observed and counterfactual scenarios.

Sensitivity Analysis. Our sensitivity analysis showed that IRR and CFR estimates were robust to differential rates of missing observations by race/ethnicity. Please see the Sensitivity to Missing Data Analysis section in the Data Materials Section and Section II of the Appendix for additional information.

3.4 Discussion

Our results indicate that although the ratio of COVID-19 incidence for non-White compared to White populations declined precipitously over the course of 2020, a narrow focus on such within-pandemic improvements may mask deeper structural issues that will re-emerge in future epidemics. Our finding that the timing of infection strongly influenced the risk of death and magnitude of disparity between Black and White Michiganders during 2020 highlights a critical but

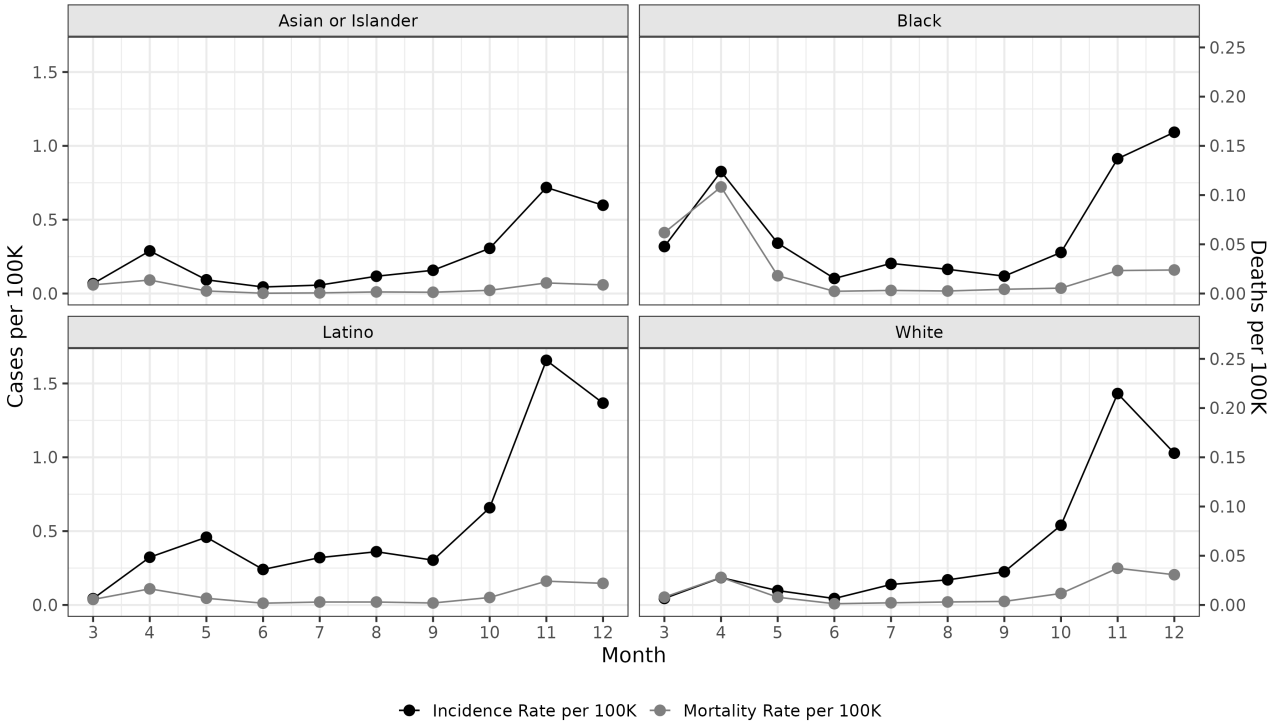

Figure 3.1: Incidence and mortality rate of COVID-19 by race/ethnicity among Michigan residents by month in 2020. The left-side y-axis is on a scale from zero to 1.75 for the cases per 100K (in black). The right-side y-axis is on a scale from zero to 0.25 for the deaths per 100K (in gray). The x-axis shows the month of the year 2020.

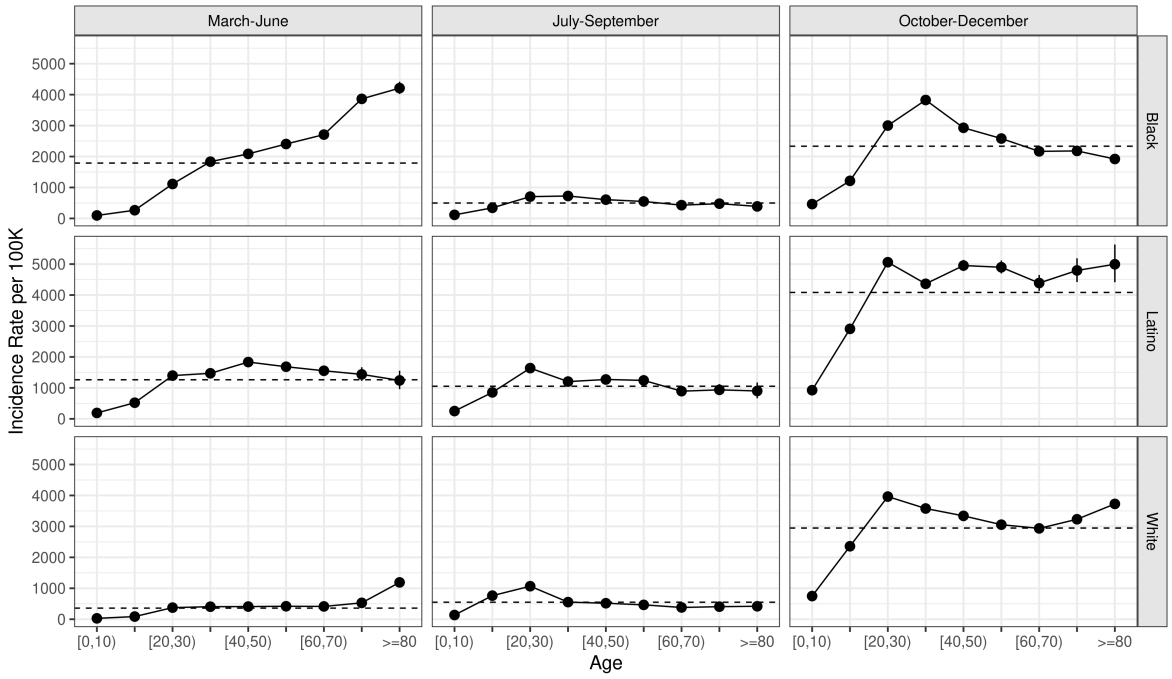

Figure 3.2: Age-specific incidence of COVID-19 by race/ethnicity among Michigan residents over three periods during 2020. Each panel represents age-specific incidence of COVID-19 by race/ethnic group (rows) for each of three time periods during the first year of the SARS-CoV-2 pandemic (columns). Dashed lines in each panel represent the age-standardized incidence of SARS-CoV-2 infection for each race/ethnic group during each time period . The vertical lines around the points indicate the width of the 95% posterior credible interval (CrI) of the age-specific incidence rate.

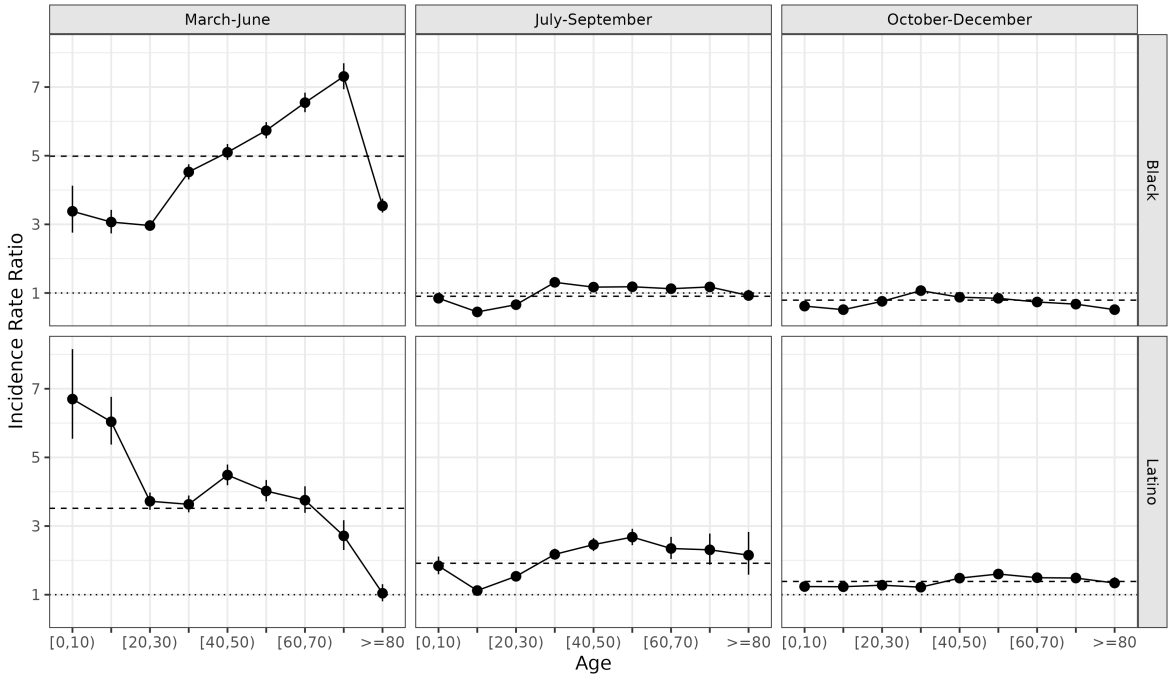

Figure 3.3: Age-specific incidence rate ratios (IRRs) of COVID-19 infection for Black and Latino residents compared to White residents of Michigan over three periods in 2020. Each panel represents age-specific IRR of COVID-19 by race/ethnic group (rows) as compared to Whites, for each of three time periods during the first year of the SARS-CoV-2 pandemic (columns). The dashed line in each panel represents the age-standardized IRR for each race/ethnic group during each time period during 2020. Dotted lines indicates an IRR of one and is provided as a reference for assessing the magnitude of risk disparity between each race/ethnic group and White residents during each time period. Vertical lines around each point indicate the width of the 95% posterior credible interval (CrI) of the age-specific IRR.

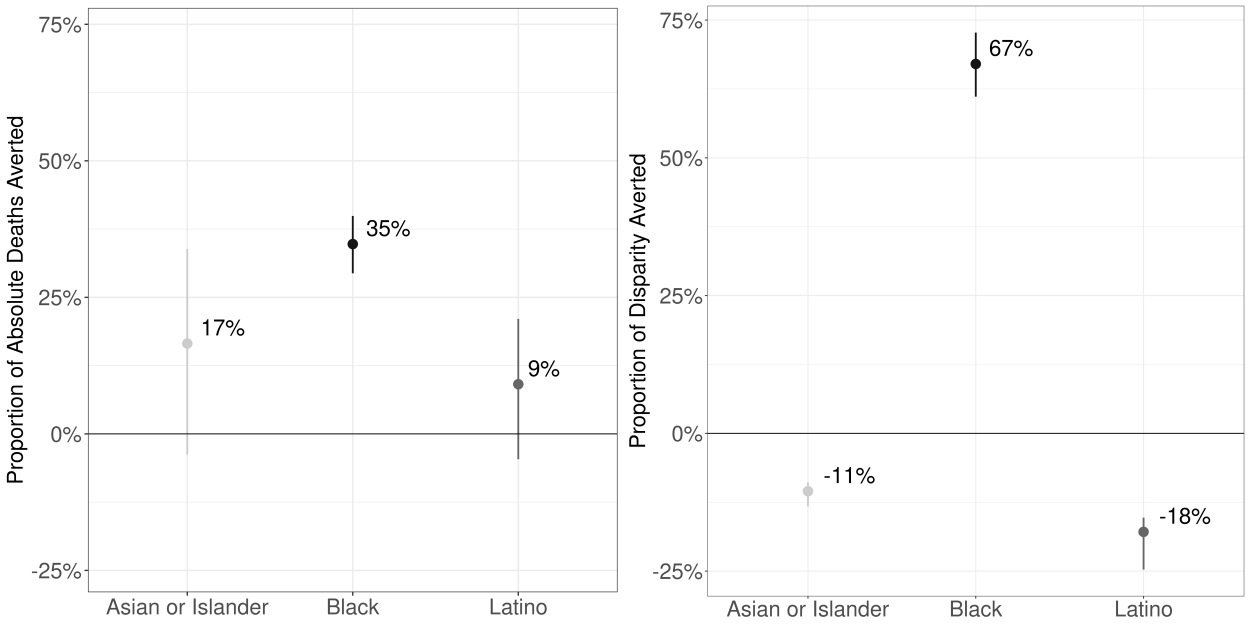

Figure 3.4: (Left) Proportion of deaths that could potentially have been averted if minoritized residents experienced the same pacing of infection as White Michigan residents during 2020. (Right) Proportion of disparity as measured by the mortality rate difference that could potentially have been averted if moinoritized residents experienced the same pacing of infection as White residents during the pre-vaccine period. The vertical lines around the points indicate the width of the 95% posterior credible interval (CrI).

under-recognized potential mechanism of socioeconomic and racial inequality in the COVID-19 pandemic: White Michigan residents appear to have been more likely than Black residents to be able to marshal the resources necessary to delay exposure until the first period of greatest uncertainty and highest mortality risk from SARS-CoV-2 was over. Differentials in economic and social power, reflected for example by the large proportion of minoritized residents as compared to White residents who work in frontline and often precarious jobs and the dynamics of racial residential segregation, may explain these differentials in the ability to 'wait out' infection [48–52, 62].

In this analysis, we drew upon a detailed administrative dataset that included information on the timing, age, and race/ethnicity of cases. However, it is important to highlight some potential limitations when interpreting these results: First, a portion of cases were missing information on race/ethnicity (see Table S11), which resulted in their removal from our final dataset. However, our sensitivity analysis indicated that our IRR and CFR estimates are robust to even large differences in the probability of missing race/ethnicity for cases among Black residents as compared to their White counterparts.

If there were systematic differences in the frailty of individuals (regardless of racial/ethnic group) who died in the first vs. second and third periods, our counterfactual analysis might overestimate the impact of infection pacing on inequalities in mortality. However, the use of age-specific comparisons using narrow age bands should mitigate this risk and facilitate a comparison between groups as a function of time. Our findings also echo findings from the national level which illustrated that the differences between White and African-American case fatality were minimal throughout the first period of our analysis [42]. Nonetheless, our counterfactual analysis should be interpreted as a thought experiment to quantify the potential life lost to inequity in infection timing rather than a definitive estimate of this disparity.

When interpreting our results, it is important to consider that, as compared to other states, Michigan's NPI policies were comprehensive, implemented rapidly, and were sustained over several months. It is possible that the net impact of this approach was to increase mortality inequity above where it would have been otherwise: Since work-from-home and other contact-limiting NPIs were disproportionately accessible to upper-class, professional workers than blue-collar, service-sector workers, such policies may have widened race/ethnic and socioeconomic gaps in infection while they reduced population-level incidence. Conversely, they may have decreased this inequity relative to what it would have been under more lax policies by providing legal and regulatory protections to employees.

Nonetheless, we believe that the overarching insight of our analysis – pre-vaccine race/ethnic disparities in COVID-19 mortality were driven by differential timing of infection rather than differences in age-specific case fatality by race/ethnicity – is likely to apply across state and local contexts, regardless of the extent of their NPI policies. For example, reports from CDC and other

sources from this period echo our findings that case-fatality rates fell dramatically and rapidly during Spring 2020 for all race-ethnic groups and that the temporal differences in these risks are considerably larger than between-group differences within any period [37, 44, 69].

Geographic, socioeconomic, and race/ethnic differences in rates of SARS-CoV-2 testing and case reporting may also have impacted our results, as our dataset reflects only probable and PCR-confirmed cases. During 2020, the availability of COVID-19 testing varied over time with limited testing available in the first part of 2020 and much wider availability towards the end of the year. The high overall case-fatality rates we observed in the first period could potentially reflect the impact of a large number of unobserved positive cases which did not result in death [70, 71]. However, case-fatality was documented to have fallen dramatically nationwide throughout 2020 [71] as a result of improvements in the clinical management of SARS-CoV-2 infection. Geographic and race/ethnic group specific differences in case-ascertainment could also bias our results. However, analyses of testing access by race/ethnicity and social conditions across three major cities in 2020 suggested that per-capita testing in areas characterized by high social deprivation was not meaningfully different from wealthier areas nor those with more minoritized populations compared to White populations [72].

Finally, analyzing race/ethnic disparities in isolation, as we did due to a lack of information on individual SES and other modifiable risk factors, limits the ability to understand the differential contributions of wealth and income inequalities versus structural and interpersonal racism on observed COVID-19 outcomes [29]. For example, Kamis et al. [73] showed that the intensity of the relationship between household crowding and COVID-19 mortality risk within U.S. counties strengthened during Summer 2020, suggesting that despite a narrowing of some race/ethnic gaps in reported rates of infection and mortality, overarching socioeconomic disparities may have in fact widened during this period. Increasing our understanding of the role played by the socioeconomic factors that put people at-risk-of-risk can only be done if such information is routinely collected by public health and medical information systems. Further, our analyses includes Latino ethnicity as a primary group, but there is great heterogeneity within this category by race [74]. In our analysis we analyzed Latino ethnicity as a mutually exclusive category from other races (if Latino ethnicity was indicated for an individual they were included in this category regardless of their race). Nonetheless, future analyses could observe trends within Latino ethnicity given our inability to do so adequately due to small case numbers within this group.

While frameworks of social equity and infectious disease preparedness are strengthening [39, 40], our analysis illustrates the centrality of a socio-structural and historical approach to questions of health equity in infectious disease [29, 32, 46, 75]. Our finding that differential rates of infection over time rather than differences in age-specific case-fatality rates drove disparities suggests that the fundamental causes of inequity have not been addressed: In the early days of the pandemic, it

was the working poor and non-White populations who were disproportionately exposed while more privileged groups were able to wait out the period of highest case-fatality. Our results highlight why it is crucial to not mistake within-pandemic narrowing of disparities for progress against the underlying causes of inequality in mortality from SARS-CoV-2 which are likely to re-emerge in future epidemics and pandemics. To understand why, a simple thought experiment is instructive: If these inequities are downstream effects of racial capitalism, i.e. a system that extracts material benefits from those racialized as non-White for the benefit of those considered to be White [76], can we expect they will not reappear when the next highly virulent and highly-infectious pathogen emerges?

The stakes associated with these types of analytic choices, which are routinely made in the analysis of infection inequity, are high and fall hardest on the most vulnerable: They shape not only our understanding of the causes of inequity in infection mortality but also the universe of policy mechanisms we perceive as capable to address them. For example, analyses of cumulative disparities in COVID-19 incidence and mortality which collapsed the three time periods across 2020 in which infection and mortality risks as well as policy responses were in flux, e.g.[62, 63], may lead us to the conclusion that race/ethnic inequities in COVID-19 mortality were attributable to group-specific differences in the risk of death upon infection, i.e. the case-fatality ratio, and thus reflecting some characteristic of minoritized individuals, i.e. pre-existing poor health, limited access to care, that could not have been addressed quickly enough to prevent these outcomes.

By contrast, our results suggest that race/ethnic differentials in case-fatality throughout 2020 reflect in large part the impact of infection timing on case-fatality. This finding suggests that increased equity in protection from infection, in the form of short-to-medium term income, increased workplace and employments protections for essential workers, and housing support, were accessible policy levers that could have dramatically reduced these differentials. Preventing these outcomes in the future necessitates both more proactive disease surveillance to detect inequity in infection as it emerges, as well as social welfare and workplace policies that could prevent these inequitable exposures in the first instance.

There is no evidence that the structural inequities which set the table for the disparities in death from COVID-19 documented here have lessened since early 2020 [77]. In fact, it is likely that the social and economic inequities which made minoritized communities more vulnerable to SARS-CoV-2 infection in the earliest days of this pandemic have been exacerbated by its effects [78, 79]. Our results suggest that addressing these inequities using every social and public health policy lever at our disposal must be a key focus of both the response to the ongoing COVID-19 pandemic and our preparedness for the future.

CHAPTER 4

Spatio-temporal Patterns of Disability in India: 2001-2011

4.1 Introduction

Globally, 1 billion people–15% of the world's population–live with at least one form of disability, and 80% of these individuals live in the global South, otherwise known as low- and middle-income countries (LMICs) throughout the world [80]. A vast majority of these disabilities are highly preventable. In a recent study, which analyzed the burden of disease among the world's poorest billion people utilizing the Global Burden of Disease framework, it was established that 65% of disability-adjusted life years were attributable to communicable, maternal, neonatal, and nutritional diseases [81]. Further, postnatal disease is the leading cause of disability in children, 95% of which are in LMICs [82]. Declines in postnatal disease mortality have led to more morbidity in the past three decades than ever before [82]. Poverty is highly correlated with disability status [83, 84] and contributes to increased burden within LMICs. The causal factors underlying the disability, such as economic deprivation, environmental contaminant exposure, and infectious disease, are typically distributed heterogeneously across spatial areas [85, 86].

In LMICs, childhood vaccination rates are low [87]. For example in India childhood vaccination rates are currently around 63% but have been lower in past decades [87, 88]. This type of environment is conducive to high rates of communicable disease within children acquired both congenitally and post-natally [89]. When these infections are left untreated they can lead to life-long disability (see Table 4.1 for list of infections which lead to specific types of disability). Congenitally, infections such as toxoplasmosis, cytomegalovirus, and varicella infections can lead to microcephaly, hydrocephaly, growth retardation and blindness [89]. Further, congenital syphilis can lead to deafness and cognitive impairment. In terms of early childhood infections malaria and helminthic diseases after repeated exposures leads to growth impairment and cognitive impairment and 15 to 20 percent of childhood meningitis survivors are left with deafness and cognitive impairments [89]. Japanese Viral Encephalitis can lead to cognitive and motor disabilities and measles

can lead to encephalitis causing cognitive impairment. Polio destroys motor neurons and leads to irreversible paralysis and although it is now eradicated in India, the impact of this illness is visible in adult populations. Lastly, trachoma infections are an example of an illness that can lead to blindness.

Characterizing the spatial distribution of disability within India and other large, geographically heterogeneous countries, provides several advantages: First, these methods are useful for identifying hotspots of disability that could be the result of spatially localized infectious, environmental, social and economic causes of disease. Second, understanding both the extent of spatial clustering of disability and socio-economic and medical issues underlying this clustering is essential for appropriately targeting screening, treatment, and policy. With regard to the former, targeted prevention campaigns have benefited from hotspot detection [90]. Concerning the latter, the unaddressed needs of people with disability–such as those related to school attendance, unemployment, etc.–are often overlooked when generating socio-economic policy [91]. Further, it has been found that education and wealth are inversely related to levels of disability and infectious disease leading to increased economic inequality in both high- and low-income countries due to increasing disability prevalence [92]. It has been proposed that low resource settings produce environments conducive to poor health outcomes and therefore high disability levels lead to low employment rates and therefore a further exacerbation of poverty. This leads to a type of poverty trap phenomenon within a subgroup of those with particularly severe health outcomes [93].

Increasingly, national censuses are collecting information on disability, facilitating evidence-based decision-making around disability policy based on the likely impact both at a regional and population level [94]. Estimates of spatial patterns of total and type specific disability prevalence and incidence are increasingly needed to plan future policy [95].

A lack of information on how disability is spatially distributed throughout LMICs, such as India, makes it challenging to funnel resources and research to achieve a targeted approach. Further, the lack of appropriate methodology for identifying concentrations of disability–for example with the use of prevalence maps–can paint an inaccurate picture of where disability is concentrated. It is important to look for disability hot spots at the local level (within states) using spatial analysis techniques that can guide targeted programs. Lastly, given the diversity of disability types, it becomes incredibly important to disaggregate by impairment to see where particular environmental and social causes are concentrated and would benefit from further investigation.

India has one of the largest populations of individuals with disability (individual with disabilities (IWD)) in the world [96]. The percentage of IWD went from 2.1% (21.9 million) in 2001 to 2.21% (26.8 million) in 2011. Previous work on disability across time in India showed spatial heterogeneity of total disability across districts in 2011 [97], and the existence of differences in disability experience across states [98]. Based on the 1991 and 2001 Census data, a Disability

Table 4.1: Congential and post-natal infections which lead to life-long disability outcomes

		Disability Type					
	Type of Infection	***Cognitive***	***Visual***	***Sight***	***Hearing***	***Movement***	***Mental***
Congenital	Toxoplasmosis	x		x		x	
	Cytomegalovirus (CMV)	x		x		x	
	Varicella	x		x		x	
	Congenital syphilis	x			x		
Post-Natal	Malaria	x				x	
	Helminthic	x				x	
	Meningitis	x			x		
	Japanese Viral Encephalitis	x				x	
	Measles	x					
	Poliomyelitis					x	
	Trachoma						

Index was calculated for 26 states in India showing differences across states [99].

Nevertheless, little is known about the extent to which specific classes of disability vary over space and time. Since the causes and consequences of different types of disability are quite variable, it is important to examine them both individually and collectively. For example, sight and hearing impairments are most commonly associated with infectious causes [96] but looking at the prevalence of these disability types in the aggregate at the state-level would not reveal potential causes of such disability types (which mainly occur at the local level). Further, studies have not compared district-level change over time to determine clusters of districts with heterogeneous increases or decreases. Lastly, the focus on state-level disability prevalence hampers the identification of environmental and social causes of disability that cross state boundaries.

In this paper, we take a spatio-temporal approach to identify spatial clusters of disability over time in India in 2001 and 2011, and the change from 2001 to 2011 at the district level. We use Indian Census data from 2001 and 2011 at the district level and analyze disability prevalence measures. We conduct spatial analysis to statistically quantify the level of spatial clustering of disability, and to identify significant clusters of high and low disability (by type) as well as changes in disability prevalence over time and across space.

4.2 Materials and Methods

Census data on disability prevalence give a unique population-wide perspective on a financially burdensome and programmatically challenging problem. Since the data are not sampled, but rather a complete enumeration, spatial analyses of Census data provide a comprehensive picture of spatial concentrations of high and low disability prevalence. Further, the collection of data at two-time

intervals, ten years apart, allowed us to examine changes in prevalence over time—after accounting for adjustments in district boundaries between the two censuses.

4.2.1 2001 and 2011 Indian Census

We used the 2001 and 2011 Population Census, which for the first time in 2001 included multiple questions on disability. According to the Census of India's methodology section, each person was asked if he/she suffered from a physical or mental disability in each of five categories (see Table B.2 in Appendix): speech, sight, hearing, mental (cognitive), or movement (physical). If a person suffered from two or more types of disabilities only one was recorded, and it was left to the respondent to decide which one they wanted to be classified into. Persons with a temporary cognitive or locomotor inability (due to acute medical conditions) on the date of enumeration were not considered as disabled. In the 2011 Census, there were some changes to the classification of disability, mainly two measures of mental (cognitive disability) were asked: mental retardation and mental illness, and an additional question was added to classify people with multiple disabilities.

4.2.2 Analysis Approaches

Disability data were aggregated to the district level, which is the first-level of administrative units within an Indian state. The average geographic size of a district is 1,313 square miles with an average of 1.1 million people. There were 593 districts in 2001 and 640 districts in 2011 within the 28 states and seven union territories of India. These data were spatially joined with a boundary map of Indian administrative units (retrieved from Harvard GeoSpatial Library, Cambridge, Massachusetts). Data were projected using Kalianpur 1975 India Zone IIb. Data joining, projection, and mapping were done in R version 4.1.2 [100], using the *sf* [101] and *tmap* [102]packages. To account for the changes in districts' boundaries between the 2001 and 2011 Census, we used population density weighting (See Appendix subsection "Additional Methodology"). Forty seven districts were added between the 2001 and the 2011 Census, and seven of them originated from two or more districts.

4.2.3 Measures of Disease

Variables of interest were all summarized at the district level for census years 2001 and 2011: (i) proportion of IWD in the population; (ii) proportion of IWD with speech impairment; (iii) proportion of IWD with seeing impairment; (iv) proportion of IWD with hearing impairment (vii) proportion of IWD with cognitive impairment (collapsed values for mental illness and retardation in 2011 Census) and (viii) proportion IWD with physical impairment. We did not include a separate

category of multiple disabilities because it was not recorded in the 2001 Census. In addition to individual years of disability prevalence, we were interested in the changes in prevalence between 2001 and 2011. We generated a rate difference measure to determine the change in prevalence by comparing the proportion of IWD in the population (at the district level) in 2011 and 2001.

4.2.4 Spatial Analysis

We mapped the prevalence of disability in each district and stratified it by type of disability in 2001 and 2011. We also mapped the change in the prevalence of disability between 2001 and 2011. To assess whether the prevalence of disability and change over time were clustered in space we used the local Moran's I indicator of spatial association (LISA) with 9999 permutations [103]. Clusters were identified using a k-nearest neighbors contiguity neighborhood definition, with k=4 and row standardization where the rows in the weights matrix were standardized to equal one allowing for different models to be readily compared. All islands were removed from the analysis. Results were corrected for multiple testing utilizing the false discovery rate (FDR) approach [104]. Data analysis, mapping, spatial analysis, and FDR correction were done in R version 4.1.2 [100] using the sf [101] and tmap [102] packages. All script was made reproducible with Snakemake software [66].

4.3 Results

The average prevalence of IWD at the district level in 2001 was 2,286 per 100,000 and 2,201 per 100,000 in 2011. In 2001, the highest mean percentage of disability types at the district level was sight (47.5%), followed by movement (27.6%) and cognitive disability (10.1%). Hearing and speech accounted for about 7-8% of IWD each. Twenty-seven percent of IWD resided in urban areas (with some districts not having any IWD residing in urban areas and others having all IWD in urbanized areas).

Figure 4.1 (c) illustrates the change in the pattern of disability prevalence (percentage of the total population with disability) across districts in India from 2001 to 2011. There was a decrease in the prevalence of disability within districts in the state of Arunachal Pradesh in the far east of India, with most districts reporting 3.1-4.7% of the population as IWDs in 2001, and every district having less than 3.1% IWD in 2011. Another decrease occurred in the south of the country in the state of Tamil Nadu, where nearly all districts in 2011 were lower than 2% prevalence of disability. Increases were observed in the states of Jammu and Kashmir and in the center of the country in states such as Maharashtra, Telangana, and Andhra Pradesh from 2001 to 2011.

Regarding spatial clusters of IWD, we found that, in 2011, high prevalence districts of disability

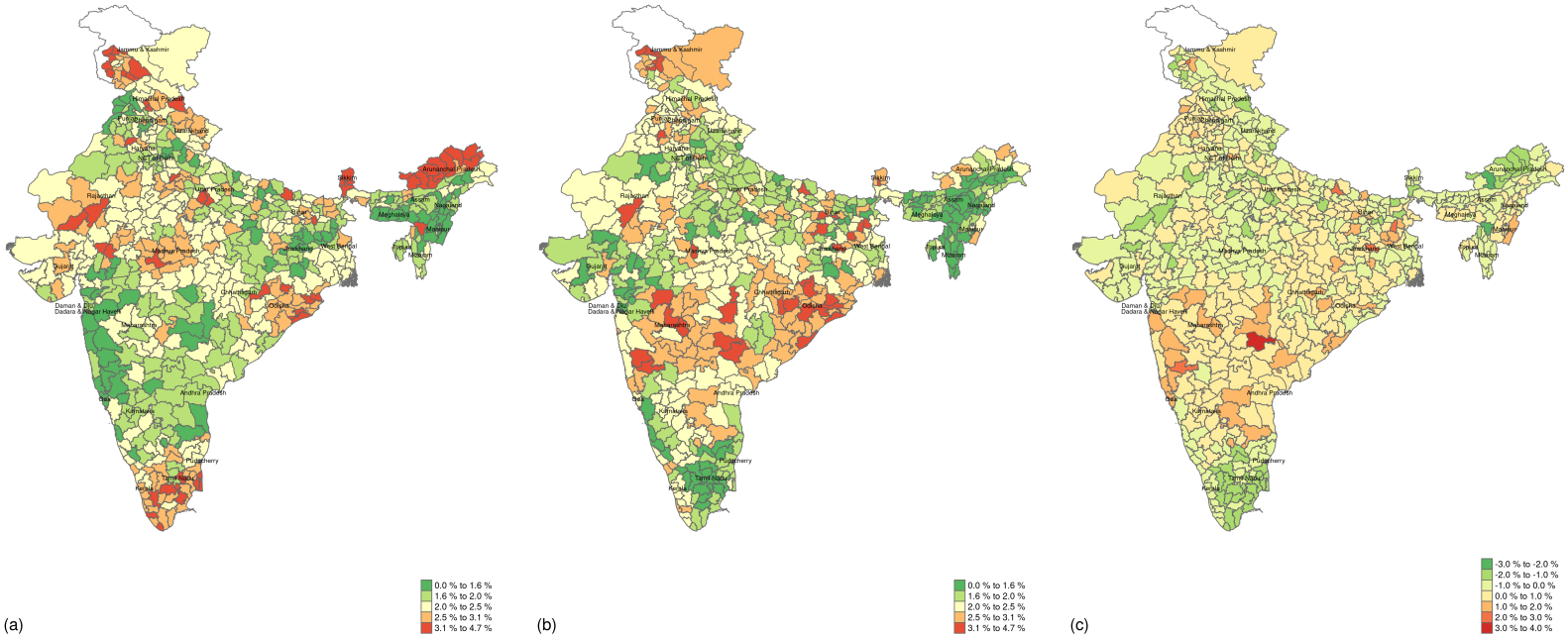

Figure 4.1: Prevalence of disability at the district level in India in (a) 2001 (b) 2011 and (c) change from 2001 to 2011.

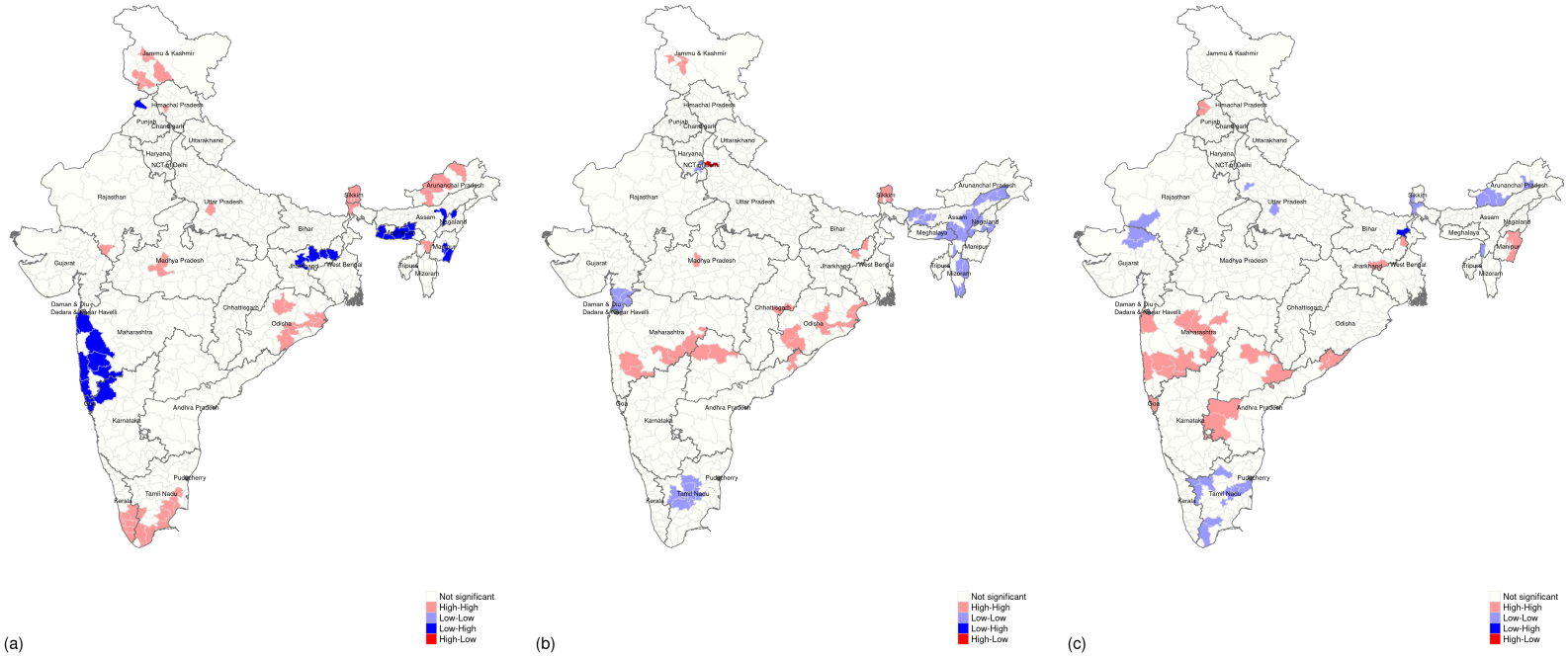

Figure 4.2: Clustering of disability prevalence at the district level in India as measured by the local indicator of spatial autocorrelation (LISA) with correction for multiple testing in (a) 2001 (b) 2011 and (c) change from 2001 to 2011. Change in disability prevalence could be both positive or negative. High-high clusters of change therefore represent those areas that had clusters of increased value (positive change). Low-low clusters of change therefore represent those areas that had clusters of decreased value (negative change). Low-high clusters of change are those that had a decreased value (negative change) surrounded by those that had an increasing value (positive change).

were clustered, even after correcting for multiple testing (Figure 4.2(b)) Some specific disability types had more spatially extensive or large area clusters of high prevalence in 2001 and 2011 such as speech impairment in mid-western India (See Appendix Figure B.1c and B.2c). There were also clusters of different disability types in the same area such as with sight and hearing impairment in eastern India in 2001, intellectual disability and mental impairment in 2011 in the state of Kerala and increasing clusters of speech and sight impairment between 2011 and 2001 in southwestern India (see Appendix Figures B.1, B.2 and B.3). Interestingly, although some low prevalence clusters of disability were identified in 2011 in eastern India (and smaller clusters in the south and western India) nearly no low prevalence clusters were identified for specific disability types in 2011. In other words, when you analyze the disability types individually, we do not see patterns of clusters emerge but analyzed together—several clusters of low prevalence can be identified. Figure 4.3 illustrates that several districts experienced high prevalence clustering in more than one type of disability (with darker shades indicating more than one type of high prevalence disability cluster in that district). In 2011 southeastern India shows the darkest colors, illustrating that these districts had several high prevalence disability type clusters in one district.

Temporal changes in disability prevalence in India from 2001-2011 were spatially heterogeneous. Increases in disability (from 2001 to 2011) were clustered in central India, while declines were clustered in areas of southern and eastern, and western India (Figure 4.2c). Larger (spatially more extensive area) clusters of increasing disability (positive change) were observed within specific disability types: speech impairment and sight impairment in mid-western India (See Appendix Figure 4.3a and c). Smaller clusters of decreasing disability prevalence (negative change) over time were observed for mental impairment in central and eastern India and movement disability in south central India (See Appendix Figure 4.3e and d). Figure 4.3c shows districts with increases in disability prevalence in one or more disability types (with darker shades indicating increasingly more types of disability). When observing increasing prevalence clusters (positive change) of disability across each disability type (see Figure 4.3c), as expected from the individual maps in Appendix Figure B.3, the districts are concentrated in the southeastern and central parts of the country. When looking at specific types of disability we see that sight impairment, whic is known to be related to trachoma, toxoplasmosis, cytomegalovirus and varicella, is clustered in particular regions of India which are somewhat similar to those of all-type disability clusters. Further, clusters of increasing prevalence in hearing impairment, which is related to congenital syphilis and meningitis, are different from all-type disability clusters and very different from sight impairment clusters.

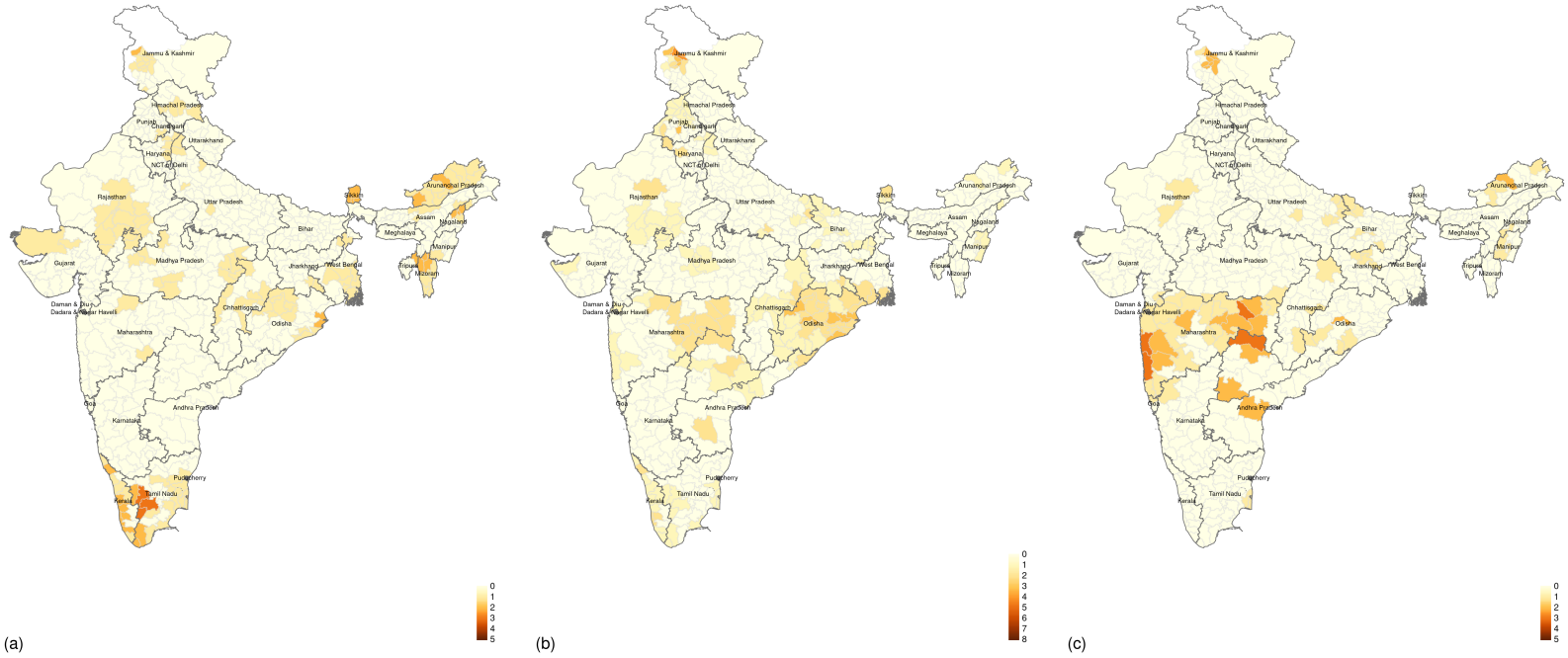

Figure 4.3: Districts with high prevalence clustering in more than one type of disability in India as measured by the local indicator of spatial autocorrelation (LISA) with correction for multiple testing bias in (a) 2001 (b) 2011 and (c) change from 2001 to 2011. Increasing darkness of color indicates more disability types.

4.4 Discussion

In this paper, we sought to identify clustering of district-level rates of overall and type-specific disability in India in 2001 and 2011, as well as the change between the two time periods using local indicators of spatial autocorrelation. We set out to understand the overlap between specific types of disability in terms of where they clustered by generating maps of every disability type in 2001 and 2011. We then compared change in disability prevalence (positive change or negative change) from 2001 to 2011 to learn about clusters of increasing and decreasing prevalence. We corrected for multiple testing to ensure an accurate representation of significant values. Complete enumeration of disability in India with the 2001 and 2011 Census allowed for the detection of clusters of high and low disease prevalence throughout the country. Our findings show that disability prevalence from 2001 to 2011 at the district level is highest in the southern center of India and increasing prevalence clusters (positive change) are also concentrated in this area. When looking across specific disability types from 2001 to 2011, western India shows the most spatially extensive clusters of increasing prevalence (positive change) for both speech and hearing disability. . Spatially clustered increases in specific types of disability over ten years should send an alert to researchers and governments to further investigate the causes of these changes.

The 2006 United Nations Convention on the Rights of Persons with Disability currently has 164 countries that are signatories indicating a willingness on their part to protect IWD citizens in their countries. LMICs make up a large share of the signatories. In India, as a response to the signing of the Convention in 2007, the Rights of Persons with Disabilities Act was enacted in 2016 outlining some of the strongest legal protections for this group in its history. The 2016 Act along with two other Acts in Indian law specifies each disability type's protections and outlines benefits for these groups. For example, currently, anyone between the ages of 6-18 with a disability is provided access to free education. The collection of data in Census records is an important step many LMICs have taken to systematically account for these groups, especially when enumerated by specific disability type. The identification of spatial clusters of IWD prevalence can and should become standard practice to target policy. For example, several states, with alarming clusters of increasing disability prevalence (positive change) for specific types of disability (see Supplement Figure B.3(a), (c) and (d)—including Andhra Pradesh, Maharashtra and the Union territories of Jammy and and Kashmir, Ladakh, Dadra and Nagar Hael and Haman and Diu—have not yet submitted their detailed rules under the Rights of Persons with Disability Act, 2016 [105]. In contrast states that have set up special courts—including Assam, Kerala, Rajasthan and Madhya Pradesh—which were mandated under the Rights of Persons with Disability Act, 2016, have no change or at times decreasing prevalence (negative change) of disability from 2001 to 2011 (see Supplement Figure B.3(a-e)). Although the Act came into accordance after the collection of this data, it illustrates the

willingness (or the lack thereof) on the part of each state to uphold the rights and protect the well-being of persons with disabilities [105]. Beyond protecting those with disability, there is a need for prevention through the identification of environmental and social causes of disability. The Right to Health is another human right, which many LMICs are signatories to under the Convention of the Economic, Social and Cultural Rights. Article 14 of the Convention outlines the importance of national and regional governments to improve disability-adjusted life years (DALYs) in their nations. The identification of spatial patterns of disability, especially those most affected by environmental causes such as hearing and sight, over time, can highlight areas of concern in the country which require further investigation. For example, some states with high levels of poverty also include clusters of increasing disability prevalence (positive change) from 2001 to 2011 including Bihar and Odisha. Therefore the above stated factors makeup a complex interplay of social, political and economic realities—for instance the state of Madhya Pradesh did excel in their willingness to enact the Rights of Persons with Disability Act, 2016, but they house the country's largest tribal populations and the fourth poorest state, and this is reflected in the band of increasing movement disability clusters in the middle of the state [106].

This analysis utilizes a complete enumeration of disability in India at the district level, making it ideal for spatial analysis and interpretation. Nonetheless, several limitations may be considered. First, the Indian Census in 2001 and 2011 likely missed some people with disability due to a lack of reporting or awareness on the part of families given the self-reported nature of Census data collection. This may be particularly true for mental illness and would likely lead to undercounting for some disability types. The National Sample Survey Organization, an independent non-governmental research organization, estimates a 20% larger proportion of the Indian population as having a disability, compared to the Census of India in 2001 [107], while other nationally representative household surveys indicate the Census disability rate was 25.2% higher [98]. It has been stated that these variable sources of disability statistics capture distinct aspects of disability and lead to disparate estimates. For this analysis, we feel that whether disability was overcounted or undercounted in the Census of India, it is likely not deferentially biased across regions. Therefore, the estimates of clustering patterns are likely held. Second, it is challenging to know how and why certain disability types cluster in the way they do in the absence of detailed policy, public health, and environmental data from the local level. The Indian Census includes some socio-economic variables at the district level but publically available data are stratified by disability status. Although this is outside the scope of this paper, future work in this area may seek to identify spatial regions where certain disability policies were enforced, environmental hazards were present or specific diseases proliferated. Third, this analysis utilizes non-age standardized prevalence rates of disability (for those over 15) at the district level and it is possible that the age structure both at the district level and nationally could have changed over time. Such age-standardization is not possi-

ble with publicly available Census records because age is not stratified by disability status nor is disability status stratified by age categories. Given that age standardization of disability prevalence at the district level is not possible with Census data, this is a limitation that should be considered in the interpretation of maps.

4.5 Conclusion

Clusters of disability are observed in districts in India in 2002 and 2011. Considering disability types, some areas have high concentrations of one or more types. Over 10 years, clustering emerged, persisted, or expanded. Our results point to the need for policymakers to consider targeting disability programs and policy both locally and by type of disability.

CHAPTER 5

Do Sociality and Spatiality Interact? Characterizing the Joint Impact of Spatial Proximity and Social Network Connectivity on Acute Infectious Disease Risks and Intervention Efficacy

5.1 Introduction

Despite declines in child mortality over the last decade, global enteric disease and respiratory illness incidence remains high [5]. Respiratory illnesses and enteric infections are still the first and second, respectively, largest contributors to global under-five childhood mortality [108]. Specifically, enteric disease and respiratory infections can be spread both via person-to-person contact and environmental intermediaries including water and fomites. According to the World Health Organization, of the 1.6 billion diarrheal cases over half a million children pass away every year and for those that survive it is the leading cause of malnutrition (and subsequently stunting and wasting) in children under 5 [109]. Further, 700,000 children lose their lives due to preventable respiratory infections. These diseases are concentrated within under resourced communities and are often a part of cycles of infections within these populations, although they are acute in nature. Respiratory illnesses and diarrheal disease make up a handful of diseases categorized as ”diseases of poverty” where 90% of the children who pass away from diarrhea live in the worlds poorest countries [1]. Prevention in the form of improvements in water sanitation and hygiene (WaSH) are cost-effective and highly protective. They aim to reduce both *exposure* to pathogens (in the form of removing pathogens from the environment) and *susceptibility* from infection (in the form of improving the nutrition and health of populations most affected by disease)[110].

Until recently, universal availability of improved WaSH at the household level was seen as the ultimate goal of enteric and respiratory disease management. There is an increasing awareness although that the success of household-based interventions in settings with a high burden of

enteric disease is dependent on community-level incidence and sanitary infrastructure[111]. In other words, it appears that the weakness of household-targeted preventative interventions in high-incidence settings stems from their inability to address community-level environmental, social and behavioral factors that require conceptualizing household-level risk as a function of embeddedness in social and spatial networks that connect individuals and households both within and across communities [112]. Disease mitigation efforts require a more advanced understanding of transmission which includes social and environmental processes together. Community-led and managed latrines and water systems are increasingly being utilized in LMICs [113, 114]. Such programs and interventions illustrate that larger socio-political processes can be effected from the bottom-up in small ways. Some of these efforts have been more successful than others, but what may be required is a more in-depth understanding of how social and environmental processes interact at the local level. For example, where are the best places to initiate these interventions to ensure optimal success? What environmental and social processes work for versus against such interventions?

Social networks exist within geographic space Over the past two decades, great progress has been made in advancing our understanding of how social networks interact with physical environments to influence enteric disease spread [115]. A recent systematic review of social networks and health by Perkins, Subramanian and Christakis in 2015, alerts the field of public health in particular to the importance of integrating space into network analyses [116] and that such integration more realistically captures the reality of a disease exposure network. Studies which focus on the social context of disease spread [23, 26–28] have highlighted the role of social class, geographic context and social connectivity. Several public health researchers in infectious disease have attempted to bring a socio-spatial perspective into observational studies by collecting information on both community social networks (contact) and geographic context (GPS coordinates). The three main sources of data for observational studies within the context of enteric disease are surveys on individual or household WaSH behaviors, Geographic Information Systems (GIS) data (household longitude and latitude) and social network (contact) surveys or kinship maps alongside disease data data. Some analyses of enteric disease have even attempted to disentangle social network process from spatial ones [22, 25]–which we explore further below.

Relatively few studies have directly examined the relative contributions of spatial and social networks to the community-wide distribution of enteric disease. Bates et al. [22] examined the role of social network connectivity and geographic proximity as predictors of enteric disease risk from 2003-2005 in the same study region as our analysis. Social connectedness was measured via two types of contact networks: whether individuals regularly spent time together (causal contact) and if food was shared between households. The findings of this study suggested that the presence of increased social contacts was correlated with increased enteric disease risk. The analysis independently analyzed the role of geography on increase network contact and enteric disease risk.

They found that more geographically dense communities had increased enteric disease risk and a lowered probability of social network connections.

A paper by Emch et al. 2012 [25] was a village level analysis which utilized kinship-based networks and variable spatial buffers to look at clustering across different spatial scales over a twenty year time period. Using a spatial effects-spatial disturbance model they identified that in the context of Bangladesh, the local environment was most important for propagating cholera and shigellosis transmission, though the social network had some influence of lesser magnitude. Density in one's contact network increased risk to enteric disease. The same research group published another study which [117] builds on this analysis, using a similar data-set but utilized road networks and maternal networks based on household migration. The main finding from this work was that clustering of incident cases consistently occurred within the space-based road network and inconsistently clustered within the social network. The last two papers utilize two connectivity matrices, one for spatial connectivity and another for social connectivity, to determine the relative contribution of each to disease spread. While both the Bates (2007) and Emch (2012, 2013) papers look at contact networks, others have demostrated the important of core-dicussion networks as a protective effect against enteric disease. While these studies did not evalute the role of geography in enteric disease spread, they illustrate the nuances of how a network could be measured and its subsequent relationship with enteric disease risk. For example Zelner et al. dentified that increased connectedness, when measured as meaningful emotional interactions or a core-discussion network (CDN), is associated with decreased enteric infection risk (utilizing data from 2007), which runs somewhat counter to the epidemiological intuition that increased connectivity is a driver of risk [23]. Other studies reinforced these findings and illustrated that core discussion networks and a contact networks have joint effects [118].

Sociality and spatiality can interact to create increased risk Although the above mentioned literature sheds light on the relative contribution of social compared to spatial connectivity in enteric disease transmission, a measure of their interaction would shed important light on the socio-spatial way neighborhoods and communities function [119]. The interaction between sociality and spatiality is a well documented feature of environments within ecology, human interaction within social movement science and migration patterns within transportation literature [120–122] but it has yet to be explored within infectious disease epidemiology. Space or spatiality captures geographic processes within a community at multiple levels. It allows us to approximate the degree in which individuals share an environment geographically close to them. This measure seeks to quantify how much individuals share things such as outdoor land, animal husbandary, latrines and soil which could be contaminated with fecal matter around their household. For example, distance-based connectivity as a measure implies that one's geographic proximity to neighbors (regardless of social interaction) informs disease spread. Building on geographic contexts, we look at an-

other measure we call social density which captures how individuals chose to interact in space. For example, social network connectivity (or social density) implies that one's social proximity to individuals (regardless of geographic proximity) informs disease spread. Both of these factors, social density and spatial density, are well characterized risk factors in the spread of infectious disease, as mentioned above, but their interactive effect has yet to be explored [119]. In other words, given that social density increases disease and spatial density increases disease, is there an additional multiplicative and/or additive effect of their interaction which increases risk beyond each individual risk factor alone?

In this paper we explore the interactive effects of social and spatial density on acute infections. Our main research question is: Do sociality and spatiality interact to increase disease risk on the additive and/or multiplicative scale?

5.2 Data and Methods

5.2.1 Data

Our analysis utilizes recently collected data from a long-standing sociometric study of enteric disease outcomes in western Ecuador. We measure spatial density and social density along with their interactive effects to observe the the association these exposures have with diarrheal and respiratory symptoms.

5.2.2 Study Site

A georeferenced census network of enteric disease risk in Canton Eloy Alfaro region of Ecuador For this analysis, we will utilize data from a long-running observational study of community-level enteric infection risk in northwestern Ecuador. Within Canton Eloy Alfaro in the Esmeraldes Province we selected six study sites in the communities of Borbón, Colon Eloy, Maldonado, Santo Domino and Santa Maria. The communities vary in population size and distance to the commercial center of Borbón and are composed of ethnically diverse (Afro-ecuadorian, Chachi and Mixed-race communities) and indigenous populations. Timbire and Colon Eloy are geographically close to one another (13 minute car drive based on Google Maps) and are about 45 minutes and 36 minutes southeast of Borbón. Maldonado is about 13 minutes east of Borbón. Santo Domingo and Santa Maria are more remote communities 4 hours south and 7 hours southeast of Borbón, respectively. We refer to Borbón as a city center and the other communities as peripheral communities. Borbón has a population of around 7,700 individuals according to the 2010 Ecuadorian census and covers a 192.6 square kilometer area. It is considered a rural area

and has pre-dominantly Afro-Ecuadorian populations (63%), followed by mestizos (21%) and indigenous populations (9%). Timbire is a much smaller town with a total population of 1,037 and 92.9% Afro-Ecudorian population. It is spread out over 22.08 square kilometers. Colon Eloy (del Maria) is similar in population size to Timbire with 1,410 people over 71.72 square kilometers. A majority of the population is Afro-Ecuadorian (93.3%). Maldonado is slighting larger in population size than Timbire and Colon Eloy with 1,703 people across a larger area of land (206.3 square kilometers). A majority of individuals in this village are mestizo (81.8%). Santo Domingo and Santa Maria are smaller communities located in remote locations within the region with less than 1,000 people living in either location. For a map of the study region see [4]. The Environmental Change and Diarrheal Disease (EcoDess) sociometric study includes extensive geo-referenced demographic, household, and social network data for all individuals in the study sites over the age of 12. The study is a part of a larger ongoing study which includes a yearly census of all households in the study areas. Sociometric survey data have been collected since 2003 over four waves of data collection, the most recent of which is utilized in this analysis. In this analysis we expected to see large differences in both the exposure and outcome by study site, especially between the commercial center and the peripheral communities given the much larger population size in Borbón (nearly four times that of some communities).

5.2.3 Data Collection

The data utilized in this analysis were collected from November 30th 2021 until April 1st 2022. The survey was given to every household in designated study communities, with the exception of Borbón. In Borbón only certain central neighborhoods were included in the study sample. The survey lasted on average 35.84 minutes and was conducted by pairs of field staff. The study included eight field staff in total. The staff were trained in sociometric methods and given tablets to collect data. The study utilized the ODK open-source Android software to implement study questions and pre-load information within the tablets. The study tracked various aspects of sociality including 1) social network connectivity (as measured by asking the respondent who they generally share important matters with and who they generally visited within their community) 2) places the respondents visited recently as well as 3) outcomes related to fever and/or diarrhea. Census data was collected earlier in 2021 (as well as prior years) and includes demographic information such as birth date, race, household address, literacy status and household Geographic Position System (GPS) coordinates. Every individual is assigned a studyID as a unique individual identifier and household number as a household level unique identifier. For the purposes of this research, the latitude and longitude of households were projected using SIRGAS 1995 / UTM zone 17S (epsg: 31992) to calculate distances between households accurately in meters. The transfor-

mation projects the round surface of the earth (in longitude and latitude) and flattens each region of the globe for precise measurement.

5.2.4 Sample

Figures 4.1-4.3 illustrate how the final datasets were obtained for mapping and modelling exposures and outcomes. Of the potential 7,005 individual included in a census of the six study communities, 4,346 individuals were interviewed during this survey. Please see the appendix for differences between the census population and the individuals interviewed in the sociometric survey (see Appendix Figures C.1, C.2) Several respondents were removed from the data: a) 243 because they did not complete the survey b) one because they stated that they had not resided in the community for more than three months (and therefore would not be considered within the social network of the community) and c) 45 did not have a recorded GPS coordinate. After exclusions 4,057 individuals remained in the dataset.

For the regression analysis, census data was required. There were several individuals (785) we could not locate in the census, due to changes in census records in the year between survey administration and census capture. Another seven individuals were removed because they had duplicate census records which could not be accurately matched. These individuals were excluded from the regression analysis sample, leaving a total of 3,265 individuals.

We utilized two networks: a network of people whom the respondent stated they shared important matters with (GIM) and those whom the survey respondent generally visited within the village (PV). Of the 3,265 individuals who completed the sociometric survey, 102 and 24 refused to answer the questions related to their GIM and PV networks, respectively. Further, 926 and 1,348 stated that they did not have someone they generally shared important matters with and visited regularly. This left 2,235 and 1,891 who provided names of individuals in their respective GIM and PV networks.

5.2.5 Exposure and Outcome Data

A sociometric survey measures the relationships between actors in a predefined group at a certain point in time. In sociometric surveys every individual is asked about certain topics within which they are to name all the people they identify with on that topic. In the general important matters network, every single person in a village is asked to list the names of people outside their household whom they share very important matters with. For the people visited network, respondents are asked 'How many families do you regularly visit within this community?' and then they are to list those individuals. The nodes of the network then become the connections named and each network is analyzed separately.

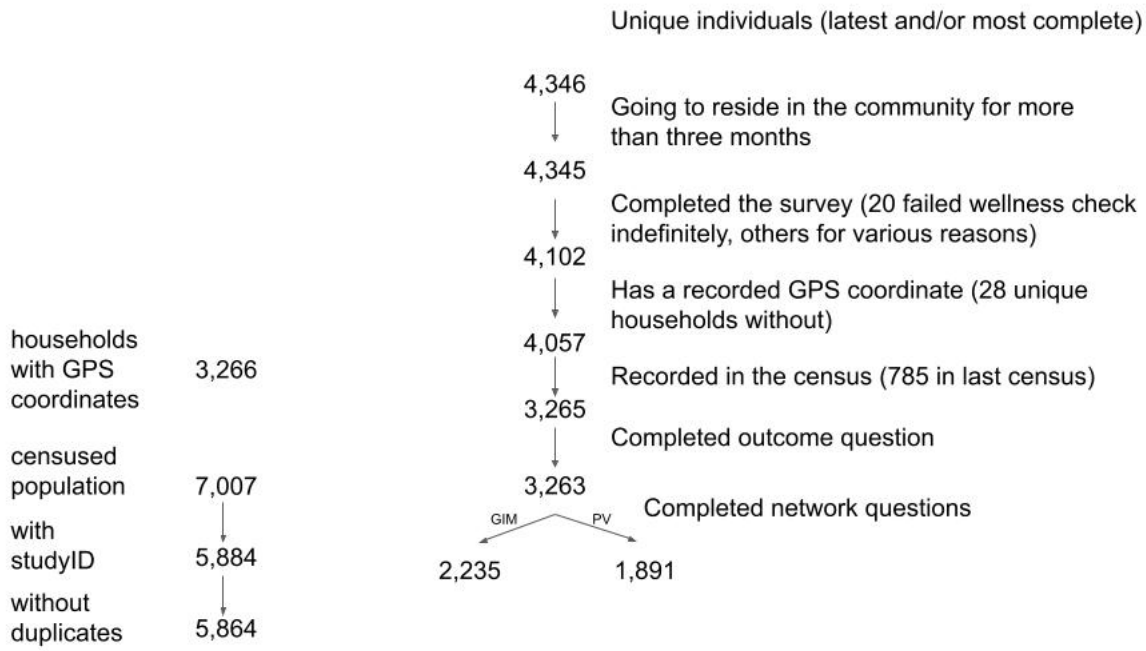

Figure 5.1: Flow chart of survey respondents in 2022 sociometric survey and final data set.

Primary Exposures The two primary exposure variables in our study are the number of contacts of an individual (social density or degree) and the average distance of an individual's household to everyone other household in the community–which we refer to as spatial density. A survey respondent provides the number of people in both of their networks (GIM and PV). Therefore there are two degree values for each survey respondent corresponding to a single network. The average distance of an individual to everyone else in their own community or neighborhood density variable is calculated by generating pairwise distance matrix and multiplying by an adjacency matrix corresponding to community membership. We then sum over each individual row to obtain a vector of values for each survey respondent v_i. The values in this vector are then divided by the number of people within that individual's community to obtain a within community relative density measure. Instead of utilizing an arithmetic mean we utilized a harmonic mean for to calculation to decrease the effect of outliers. This measure is similar to the density measure used in other studies [22]. The measure quantifies the average distance an individual lives away from others in their community. Therefore, a high spatial density measure indicates that the respondent lives in a more dense area, while a low value indicates dispersion.

Outcome At the time of the sociometric survey, questions are also asked about enteric and respiratory disease symptoms. The main outcome variable in this analysis is a two week recall of fever and/or diarrhea. At the time of the sociometric survey, each respondent is asked whether they

have three or more episodes of diarrhea or a fever (generally considered above 100.3F) within the last two weeks. This measure allows us to capture enteric and other illnesses within a period of time that is memorable for the respondent.

Covariates A yearly census is conducted in the six study sites where are all members of a community are asked to share basic demographic information about any individual over the age 12. Many of the identified confounders in our analysis come from the census. The census data collection took place a year prior to sociometric survey administration. The census information is combined with the sociometric survey data to produce the final dataset for analysis. The main known confounders in our model include race (Afro-Ecuadorian, Chachi, Mixed and Other), sex (Female, Male or Other), age (calculated from birth date) and literacy (whether the individual can read and/or write). Race is related to social density given variable levels of connectivity by race. Further, residential segregation by race can be observed within communities. Further, being male, of older age and literate is related to increased social contact. Acute infections are also variable by age, race, sex and literacy in our data set. We generate a variable indicating whether the individual was Afro-Ecuadorian and another variable for Chachi individuals. For sex, a variable was created for Male versus not Male. Literacy was modeled as either being able to read and write versus other (read and not write, write and not read, neither read nor write).

5.2.6 Analysis Methodology

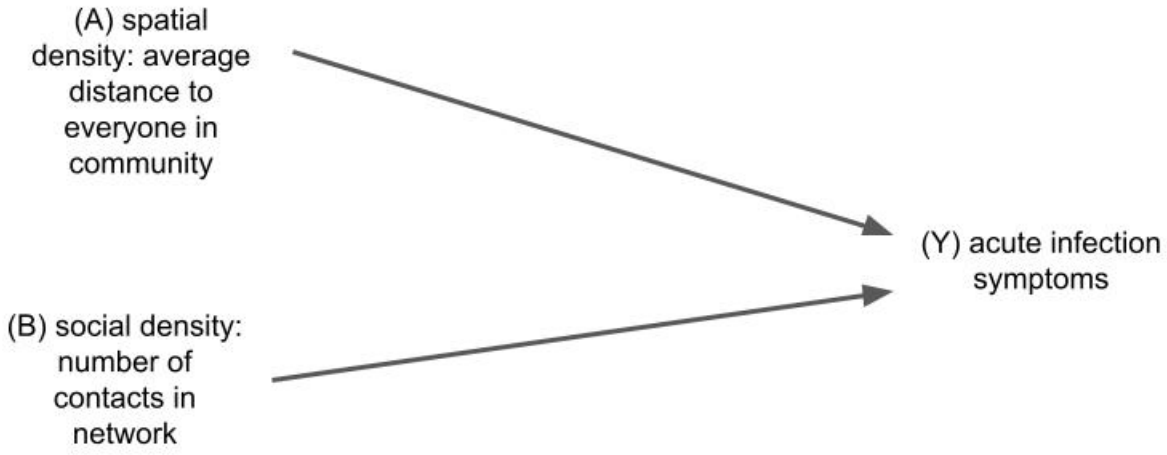

Figure 5.2: Directed acyclic graph of proposed causal relationships between variables

The impact of social and spatial density on acute infection risk We first conduct a descriptive analysis of the differences in risk between the two networks of interest. We also explore differences in both the exposure and outcome by community, especially given that Borbón is a much larger population center compared to the rest of the study sites to determine if stratification by community was needed. It was determined that stratification was needed after this assessment, therefore two equations are presented below for each model. Our regression analysis approach seeks to more directly answer the question: Is there an interaction between spatial and social density? We utilize a Bayesian logistic mixed-effects modeling approach to observe the relationship between social and spatial density and disease risk, given a set of known confounders (see Figure 4.4). Equation 4.5 outlines the model for the five peripheral communities, while 4.6 defines the model for the study site of Borbón. We apply a weakly informative prior to the model with a mean of zero and scale (standard deviation) of 2.5. We transform the outcome into odds ratios for ease of interpretation. We utilized the *rstanarm* package within R.

Testing for multiplicative interaction between social and spatial density Interactive effects can amplify or dampen the effects of individual exposures. Interaction refers to the "effect of one exposure on an outcome to depend in some way on the presence or absence of another exposure" [119]. It is different from effect modification in that both exposures have an effect on the outcome. Both spatial density and social density have an established effect on enteric disease outcomes [22, 23, 118]. Therefore, we seek to understand whether the interaction between social density and spatial generates an additional effect on enteric disease outcomes.

$$\begin{aligned} Y_{ij} = (\beta_0 + \mu_{0j}) + \beta_1 X_1 + \beta_2 X_2 + \\ \beta_3 X_{1*2} + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \\ \beta_7 X_7 + u_j + e_{ij} \\ i = \text{ego}, j = \text{community} \\ X_1 = \text{social density}, X_2 = \text{spatial density} \\ X_{1*2} = \text{social density X spatial density}, X_4 = \text{race} \\ X_5 = \text{sex}, X_6 = \text{literacy}, X_7 = \text{age} \\ u_j = \text{random intercept at community level} \end{aligned} \tag{5.1}$$

$$\begin{aligned} Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_{1*2} + \\ \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + + e_i \\ i = \text{ego}, X_1 = \text{social density}, X_2 = \text{spatial density} \\ X_{1*2} = \text{social density X spatial density} \\ X_4 = \text{race}, X_5 = \text{sex}, X_6 = \text{literacy}, X_7 = \text{age} \end{aligned} \tag{5.2}$$

Interaction can be multiplicative and/or additive. Additive interaction in on the risk difference scale and is sometimes considered to be more relevant from a public health perspective [119] whereas multiplicative interaction looks at risk ratios or relative risk. Multiplicative interaction quantifies the additional effect of combined exposures on the risk ratio scale. It provides insight into the relative risk of disease given an increase in the exposure, but it does not indicate which subgroup might be best to treat. The multiplicative scale is useful from a policy perspective when interventions are less concerned about the baseline level of the exposure. In our analysis, the baseline level of exposure (level of spatial density or social density) may or may not matter, depending on the type of intervention utilized to mitigate disease. Suppose we wanted to intervene on spatial density (where homes are spread out more for example), additive interaction would provide

more insight into whether to intervene on those with high or low social contacts. Multiplicative interaction would not provide this type of information.

In logistic regression models, multiplicative interaction is assessed by adding an additional exposure (a product of two exposures). The exponentiated effect estimate $e^{\Upsilon_{\text{Social density X Spatial density}}}$ therefore equals $OR_{\text{Social density X Spatial density}}/(OR_{\text{Spatial density}}OR_{\text{Social density}})$ where and $e^{\Upsilon_{\text{Spatial density}}} = OR_{\text{Spatial density}}$. Therefore, $e^{\Upsilon_{\text{Social density X Spatial density}}}$ can be referred to as the statistical multiplicative interaction for a logistic model [119]. We can also assess additive interaction with the equation 5.3 to estimate the $RERI_{OR}$ (see Equation 5.3).

$$\begin{aligned} RERI_{OR} &= OR_{\text{Social density X Spatial density}} - OR_{\text{Social density}} - OR_{\text{Spatial density}} + 1 \\ &= e^{\Upsilon_{\text{Social density}}+\Upsilon_{\text{Spatial density}}+\Upsilon_{\text{Social density X Spatial density}}} - e^{\Upsilon_{\text{Social density}}} - e^{\Upsilon_{\text{Spatial density}}} + 1 \end{aligned} \tag{5.3}$$

Overall we utilize various methodologies from spatial epidemiology and social network analysis to understand the role of spatial and social connectivity on infectious disease outcomes. We bring to light new techniques which can be used to visualize cases in social-geographic space to better understand how disease may be spreading and investigate the impact of integrative measures of social and spatial connectivity on disease outcomes.

5.3 Results

Our study included a final sample of 3,265 individuals within the six study sites (see Table 5.1). A majority of our sample, 1,692 (51.8%), came from the community of Borbón. Based on our censused population, 55.7% completed the final sociometric survey. The community-specific response rates ranged from 50.67% in Maldonado to 64.58% in Colon Eloy.

Disease risk varies by community A total of 213 (6.52%) individuals reported fever and/or diarrhea in two weeks prior to survey administration (see Table 5.2). The outcome varied by community ranging from 4.04% of the sample in Timbire to 11.91% of the sample in Colon Eloy. We observed minimal missingness in the outcome variable (0.06%).

Table 5.1: Survey Response Rate within Censused Population

Community Name	Completed Survey	% Community	Total	% Completed
Borbon	1692	51.82	3056	55.37
Colon Eloy	403	12.34	624	64.58
Timbire	297	9.10	541	54.90
Maldonado	418	12.80	825	50.67
Santo Domingo	211	6.46	365	57.81
Santa Maria	244	7.47	453	53.86
Total	3265		5864	55.67

Table 5.2: Fever and/or Diarrhea by Community

Community Name	No	Yes	Missing	% No	% Yes	% Missing
Borbon	1617	75	0	95.57	4.43	0.00
Colon Eloy	354	48	1	87.84	11.91	0.25
Maldonado	392	26	0	93.78	6.22	0.00
Santa Maria	214	29	1	87.70	11.89	0.41
Santo Domingo	188	23	0	89.10	10.90	0.00
Timbire	285	12	0	95.96	4.04	0.00
Total	3050	213	2	93.42	6.52	0.06

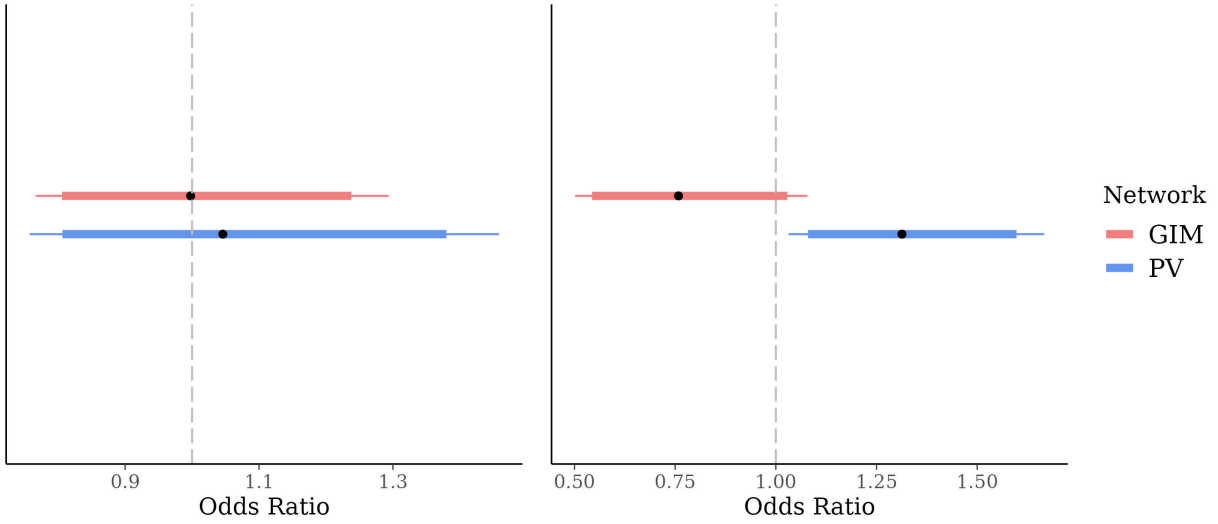

Figure 5.3: (**left**) Borbón (**right**) Peripheral communities. Mean effect estimate and 90% (thick line) and 95% (thin line) credible interval of interaction term (Social density X Spatial density (100m)) in Bayesian (mixed-effects) binomial logit model of fever and/or diarrhea which includes social density, spatial density (100m) and interaction of both along with confouders of both. Model for GIM network is in red and model for PV network is in blue. The dashed vertical line indicates an odds ratio of 1.0 (null).

Our primary exposures of interest follow patterns similar to those identified in previous papers [22, 23, 118, 123, 124]. Increased social density (degree) was protective against disease in the GIM network, odds ratio of 1.09 (95% CrI: 0.78, 1.51), but was risk inducing in the PV network, odds ratio of 0.74 (95% CrI: 0.56, 0.95), in the five peripheral communities. Degree was not a significant predictor against disease in the GIM network but was significant in the PV network at the 95% credible interval level. In terms of spatial density in the peripheral communities, the odds of disease was 1.01 (95% CrI: 1.00, 1.02) in the GIM network and 0.99 (95% CrI: 0.98, 1.00) in the PV network, neither of which were significant at the 95% credible interval level. In Borbón, a one unit increase in social density is associated with an odds of 0.70 (95% CrI: 0.34, 1.32) in the GIM network and 0.92 (95% CrI: 0.42, 1.82) in the PV network. Social density was not significant in either network with Borbón. In terms of spatial density, a one unit increase (100m) was associated with an odds of 1.00 (95% CrI: 1.0, 1.01) in the GIM network and 1.00 (95% CrI: 1.00, 1.01) in the PV network. Spatial density did not have a significant effect on disease risk in Borbón within either network. Given differences of the magnitude and direction of the primary exposures within the two models, our findings confirm that there is effect modification, for our primary exposures and outcome, by population size (Borbón versus more remote communities).

Table 5.3: Bayesian binomial logit model of fever and/or diarrhea of social density, spatial density and interaction with confounders in Borbón. Bold indicate significance at 95% credible interval level.

Network	Parameter	Mean	Lower 95%	Upper 95%
GIM	(Intercept)	0.00	0.00	0.06
	Social Density	0.70	0.35	1.29
	Spatial Density (100m)	1.33	0.79	2.11
	Social Density*Spatial Density	1.05	0.76	1.46
	Afro-Ecuadorian Race	2.63	0.87	10.35
	Chachi Race	0.00	0.00	6.39
	Male	0.66	0.28	1.50
	Literacy	4.31	0.60	80.01
	Age	1.02	0.99	1.04
PV	(Intercept)	0.01	0.00	0.19
	Social Density	0.90	0.39	1.79
	Spatial Density (100m)	1.12	0.58	1.96
	Social Density*Spatial Density	1.05	0.76	1.46
	Afro-Ecuadorian Race	1.47	0.55	4.47
	Chachi Race	0.00	0.00	3.20
	Male	0.79	0.32	1.88
	Literacy	2.14	0.28	47.73
	Age	1.01	0.98	1.04

Social density and spatial density exhibit multiplicative and additive effects on risk We observed non-significant positive multiplicative interaction in the commercial center of (Borbón) ($OR_{\text{Social density X Spatial density}}$ = 1.05 (95% CrI 0.76, 1.46) and 1.05 (95% CrI 0.76, 1.46) in the GIM and PV networks respectively). In the peripheral communities we observed interaction negatively ($OR_{\text{Social density X Spatial density}}$ = 0.76 (0.52, 1.09) and positively 1.32 (1.05, 1.67)) in the GIM and PV networks respectively, but only the PV network effect was significant at the 95% credible interval level. therefore, in Borbón multiplicative interaction of social and spatial density was positive while in the peripheral communities multiplicative interaction was negative. This means living in spatially dense areas of the community combined with having a large number of contacts is risk decreasing in Borbón while it is risk increasing in the peripheral communities. With regard to additive interaction (or the risk difference in illness for one additional meter in contact density) in Borbón, there was positive additive interaction (0.02 (95% CrI: 0.62, -0.94)) in the GIM network and negative additive interaction in the PV network (0.03 (95% CrI: 0.79, -1.29)). With regard to

Table 5.4: Bayesian mixed-effects binomial logit model of fever and/or diarrhea of social density, spatial density and interaction with confounders in peripheral communities. Bold indicate significance at 95% credible interval level.

Network	Parameter	Mean	Lower 95%	Upper 95%
GIM	(Intercept)	0.05	0.01	0.35
	Social Density	1.09	0.77	1.51
	Spatial Density (100m)	1.88	0.71	4.64
	Social Density*Spatial Density	0.76	0.52	1.09
	Afro-Ecuadorian Race	1.14	0.33	5.09
	Chachi Race	4.28	0.88	27.94
	Male	0.90	0.47	1.62
	Literacy	0.87	0.37	2.19
	Age	1.00	0.99	1.02
	Intercept – Colon Eloy	1.54	0.65	4.11
	Intercept – Timbire	0.64	0.19	1.57
	Intercept – Maldonado	0.62	0.19	1.53
	Intercept – Santo Domingo	1.66	0.71	4.74
	Intercept – Santa Maria	0.72	0.20	1.88
	Sigma	2.05	1.01	27.50
PV	(Intercept)	0.18	0.03	1.14
	Social Density	**0.74**	**0.55**	**0.96**
	Spatial Density (100m)	**0.41**	**0.16**	**0.96**
	Social Density*Spatial Density	**1.32**	**1.05**	**1.67**
	Afro-Ecuadorian Race	0.50	0.22	1.22
	Chachi Race	1.63	0.44	6.79
	Male	0.88	0.51	1.49
	Literacy	2.42	0.86	8.13
	Age	1.00	0.98	1.02
	Intercept – Colon Eloy	1.45	0.72	3.33
	Intercept – Timbire	0.71	0.27	1.48
	Intercept – Maldonado	0.75	0.29	1.57
	Intercept – Santo Domingo	1.43	0.68	3.52
	Intercept – Santa Maria	0.71	0.21	1.58
	Sigma	1.59	1.01	10.13

additive interaction in the peripheral communities, there was negative additive interaction in the GIM network (-1.21, (95% CrI: , -0.0347)) and positive additive interaction in the PV network (1.17, (95% CrI: 1.34, 0.75)).

5.4 Discussion

Social density and spatial density predict enteric disease outcomes and display interaction on the multiplicative and additive scales. These variables display effect modification by community size (commercial center versus smaller peripheral communities). Our analysis illustrates that, in relative risk terms, individuals who live in dense neighborhoods with with a high number of social contacts are at highest risk for disease in peripheral communities. In absolute risk terms, our analysis shows that efforts to reduce the effects of spatial density on enteric outcomes will see the most benefit among individuals with fewer social contacts.

Social density and spatial density have a multiplicative and additive effect on enteric disease outcomes The integrative influence of social and environmental factors continues to harness attention within infectious disease literature [125, 126] but few studies provide empirical evidence of how they interact. Social networks and geographic indicators are necessarily an indicator of underlying processes within a community. They do not identify the processes themselves but rather provide clues into what could be influencing disease that is not already accounted for. Many of these processes occur beyond the individual level, making it challenging to quantify and utilize in models of individual-level enteric disease risk. Our study provides some clues into what those underlying processes could be by attempting to integrate two layers of risk: social density and spatial density. We identify significant positive multiplicative interaction in the peripheral communities within a contact network, highlighting the increased risk of those who live in densely populated areas with a high number of social contacts. Further, we identify positive additive interaction in the peripheral communities in the PV network illustrating that for interventions that seek to decreased crowding and population density for the purpose of mitigating enteric disease, applying such interventions among those with fewer social contacts may yield more benefit.

Social and environmental factors interact to influence enteric disease outcomes Our findings shed light on the integrative influence of social and environmental factors for the spread of enteric disease at the individual level. Our analysis illustrates that while social density (degree) and spatial density (neighborhood density) have independent effects on enteric disease outcomes, they also have multiplicative and additive effects when interacted together. Previous studies have demonstrated the independent contribution of both degree and neighborhood density on enteric disease spread [22, 23, 25, 123], but we illustrate the multiplicative and additive influence of both factors beyond their individual contributions to enteric disease risk alone.

5.4.1 Limitations

Our findings represent an important advance in our understanding of the integrative effect of social and environmental processes for enteric disease risk measured at the individual level, but it is important to keep in mind several limitations. First, because we were unable to identify and seek participation from all individuals within the six communities, there is a possibility that our sample does not represent that of the entire population within these communities. An advantage of this study is that all individuals within the rural communities and a sizeable proportion in Borbon have taken the census survey in 2021. Using this demographic information we are able to identify differences between the sampled population and the wider community. Overall, it seems that our sample is older and has more schooling than the censused population, but this largely because the survey was limited to those above the age of 12. We did not observe meaningful differences between the study population and the censused population, based on available demographic information. We did although have to leave out several people within the sample who were censused in the last year, meaning that our sample is limited to those that we were able to identify in the census prior to 2021. The addition of individuals added into the census after 2021 is possible, but will require data matching which has yet to occur. We hope that by the time of publication these individuals will be added to increase the sample size of our study.

Second, although there was extensive oversight and regular meetings with the field team, it is possible that respondents taking the survey could not have responded truthfully or fully about the number and names of people that the share general important matters with and/or visit within their community. The field staff in our study are local members of neighborhoring communities who have been trained in sociometric methods. Further, this is the fourth wave of the sociometric survey beginning in 2003. Given the experience of both the field staff and the study team in conducting this research, it is our hope that such measurement error would be minimized. The study findings also reflect that of two previous studies conducted using the general important matters network [23, 127]. Third, this study was conducted two years after the SARS-CoV-2 pandemic and field staff were required to conduct a wellness check prior to interviewing the respondent. Efforts were made to revisit households that failed the wellness check (reported symptoms of COVID-19), therefore in the end very few respondents remained who continued to fail the wellness check.

Lastly, our data was collected cross-sectionally and therefore reverse causality is possible. For example it is possible that sick individuals reported fewer people in their contact and core-discussion network because they are chronically ill. This extends to distance-based measures as well, where sick individuals could report visiting and sharing important matters with nearby contacts because of their inability to travel. Our survey questions attempt to prevent this by asking who the respondent visits *often* and *generally* shares important matters with (rather than asking about their contact with individuals in the past two weeks). Further, we would expect to see a difference

in the direction of the relationship between contact density and enteric disease in the PV versus the GIM network, but this is not the case. It is also unlikely that acute diarrheal disease would change long term travel patterns or relationships of patients given the nature of the illness being short.

5.4.2 Conclusion

Social and spatial density can be used together to target programming, policy and disease modeling in semi-urban areas Identification of densely populated areas and high contact individuals, can be used identify locusts of risks and to target interventions which seek to mitigate disease. Social contact and geographic proximity are two important, well studied, elements of enteric disease risk [22, 23, 25, 118, 124]. We integrate these measures to illustrate that individuals living in dense areas who also have a high number of social contacts experience increased disease risk, especially in remote communities. We also postulated and confirmed that our primary exposures: social density and spatial density are modified by remoteness and/or community size. Our findings hold true after accounting for other known demographic confounders. Further research is needed to determine how social and spatial density interact in other settings, but based on our analysis, it is important to assess social and spatial density together when assessing environmental and social aspects of individual risk. Further, policies which target crowding and housing density [128, 129] many benefit from taking a contact-based approach to risk-mitigation, where groups of contacts are provided interventions together to improve outcomes. Lastly, disease modelers [90, 130, 131] may consider how social density and spatial density interact when establishing potential parameters to assess the effectiveness of hypothetical interventions. Social and spatial environments interact to create risk and advancing our understanding of this interaction could lead us to improvements in the form of enteric disease morbidity and mortality.

5.4.3 Future Directions

In this analysis, we explored the interactive effects of social and spatial density on acute infection risk within a commercial center and five peripheral communities. Looking forward, refinement of this analysis and several related analyses have been identified to further advance questions related to the integrative effects of social and spatial density on disease risk. To further strengthen this analysis, we will draw upon additional census records which are to be added in the Summer of 2023 by data management coordinators so that we may increase our sample size. Additionally, a of the sample will be needed. Therefore, additional descriptive tables and a re-implementation of the regression model will be added at that time. This analysis brought to light several other research questions, which will materialize into separate but related research papers in the near future. First, a novel measure of contact density was generated after the identification of interactive effects

between social and spatial density. Contact density is a weighted individual-level social-spatial density measure which is an average distance to individuals who are identified as contacts within the GIM and PV networks. Analysis utilizing this measure has advanced but it was determined that the analysis related to contact density warranted a separate or extended version of this paper with tangential research questions to those identified in this analysis and therefore is included in the Appendix (see sections entitled ”Contact Density Methodology”, ”Contact Density Results” and ”Contact Density Discussion”). The analysis related to contact density also included a thorough mapping of the social and spatial networks in each community (see Appendix) which illustrated the impact of contact density on disease risk qualitatively. Lastly, to incorporate a temporal component to questions of social and spatial density on disease risk, the communities of Timbire and Colon Eloy were analyzed across four waves of data collection to determine the relationship between spatial and social connectivity in 2007, 2010, 2013 and 2022. The three related analyses will form a group of complementary papers to be published in the near future.

CHAPTER 6

Conclusion

In this dissertation we utilize three different types of Infectious Disease (ID) data sets: surveillance, census and household survey data to elucidate the mechanisms behind the social processes which drive ID. We drew on frameworks of social disease transmission within traditionally socially mediated infections such as HIV/AIDS and TB and applied them to what have been historically known as "equal opportunity infectors": respiratory illness and enteric disease [77, 132]. By generating and drawing upon a coherent framework of how social and environmental factors lead to disparities in infectious disease outcomes we were able to analyze heterogeneous social, environmental and geographic contexts. We illustrate the importance of diverse and creative approaches to the study of social processes including: the use of policy and behavior-informed periods of transmission to observe mechanisms related to structural racism in the spread of COVID-19 in Michigan over time, the identification of geographic clusters of disability prevalence derived from the census of India and the interplay between increased contact and crowding in the spread of enteric disease in northwestern Ecuador.

Each paper makes a unique contribution towards to the elucidation of particular elements of the frameworks outlined at the beginning of this dissertation. Nonetheless, our papers illustrated the need for further exploration into several more fundamental elements of the framework related to data sources, data collection and levels of analysis. To gain a cohesive understanding of how and why IDs spread among socially vulnerable populations these fundamental elements require attention. Mainly we made observations related to a) the inability of traditional IDs data sources to answer questions of social processes b) the lack of equity-informed data collection approaches within ID surveillance and studies and c) the need for historically informed and cross-population analyses. In this section we explore each of these observations in detail.

6.1 Infectious disease and social determinants research requires a rethinking of traditional data sources.

The study of social processes (such as racial capitalism, social cohesion, colonialization, constructive resilience, development and patriarchy) and their effects on ID requires creativity about data sources [76, 133–135]. Cohort studies (and their derivatives) are the most effective way to study causality within epidemiology. Yet, very few individual-level or household-level cohort studies of disease transmission collect data on social elements, with the exception of STDs/HIV research where contact is socially mediated. Often such data is not only difficult to collect, but the social processes may also have very complicated exposure pathways to the outcome of interest.

Given data constraints, analyses of social factors within ID are limited. Our analysis in Aim 3 which utilizes the EcoDess study provides some insight into social processes over time, but it is a cross-sectional study and therefore causal inference can be limited. The uniqueness of this data set, nonetheless, is the collection of social network characteristics alongside ID outcomes, making it possible to draw inferences about complex social processes like social density, neighborhood density and contact density as we discuss in Aim 3. In the absence of such data, researchers often combine aggregated ID surveillance data with census or other national sample data to infer associations about social processes. In Aim 1 and 2 we utilize census data (from the United States and India respectively) to gain insight into social processes. For our Aim 1 analysis we were ideally positioned to obtain race/ethnicity variables within SARS-CoV-2 surveillance data because of a Federal U.S. government mandate several months into the COVID-19 pandemic after disparities in race/ethnicity grew to astronomical levels within weeks of the virus entering the population [136, 137]. Prior to that point, race/ethnicity data was not routinely collected as a part of ID surveillance throughout the country [138].

Surveillance systems with social data are need for population-level analyses of disparities. Although it is known that diseases like influenza, tuberculosis and HIV/AIDS have clear racial patterning, the collection of race/ethnicity data has never been mandated within surveillance systems at the federal level in the US [139, 140]. The debate is not limited to ID research, there has been resistance to collecting race and ethnicity data within electronic health records and government-based databases of chronic diseases for decades [141]. Nonetheless, it could be argued that the collection of race and ethnicity data as a part of surveillance is even more pressing for ID research given the acute nature of most IDs and the need to act with swiftness in cases of outbreaks and emerging pathogens. Ultimately, there is a need for more policies, in line with the ”Equitable Data Collection and Disclosure on COVID-19 Act” of 2020, which require the collection of race, ethnicity, country of origin and occupation data within ID surveillance systems to begin to answer questions related to social factors in transmission at the population-level [142].

Household transmission studies are well placed to collect data on neighborhood social processes. Community and household level social processes such as residential segregation, occupational exposure and social cohesion impact outcomes can be studied in the context of epidemiological survey data. For example, household-based transmission studies of respiratory illnesses (including various influenza sub-types and RSV) along with hospital-based studies of acquired infections collected across diverse populations with information on income, wealth, race/ethnicity, occupation and social capital could greatly advance questions of infectious disease disparities. Another emerging source of data for ID research, big data–including electronic health records, mobile phone data, social network data and internet searches—could be leveraged (with caution) to better understand how social processes impact ID outcomes [143]. This type of data is vast and requires purposeful contextualization, but it has already shown promise within ID transmission research generally for the past decade and could be extended to include more social exposures [18, 144]. Overall, there is a need to evaluate what questions traditional ID data sources can and cannot answer and re-imagine them with social context in mind, if we are to advance important questions of social equity within ID.

6.2 Studies of ID social processes benefit from collaborative and participatory action research approaches.

Community-based transmission studies are often relegated to the ”global South”. This is with the exception of certain diseases which continue to be endemic in high-income countries such as influenza, HIV/AIDS and hospital acquired infections–which we discuss extensively above. Many studies of ID transmission in the global South are funded by institutions and agencies in the global North. Recent calls for action within the field of global health highlight the need to evaluate the inherent systems of power these types of relationships reinforce [135, 145–147]. Many call for more involvement of local academic voices to set research agendas, evaluate processes and share findings. Researchers in the global South have highlighted the disparity in ID agendas coming from within versus outside their own countries [148].

ID research requires a justice-oriented approach. The role of traditional medical academic centers in global health research is prominent and requires deep inquiry into what is and what is not equitable. A recent publication entitled ”Anti-Racism and Anti-Colonialism Praxis in Global Health—Reflection and Action for Practitioners in US Academic Medical Centers” brings to light three central questions for researchers at these institutions to ask themselves and their colleagues: 1) How do I understand the legacies of racism and colonialism in global health? 2) How do I value different ways of knowing in public health? 3) What are my motivations for global health?

Inquiries into the nature of the relationship between traditional academic centers in the global North and populations in the global South may bring to light assumptions such as "poverty and resource scarcity as the natural immutable reality" of these communities [135]. These types of assumptions greatly inform how social factors are studied in the context of IDs and act as a lens through which such processes are analyzed and reported upon. For example, our own research in northwestern Ecuador was one of the first to demonstrate how community social cohesion, rather than symbols of poverty and strife, were and continue to be a protective factor against ID outcomes [118, 149]. Gaining clarity around the nature of relationships between actors within global health will greatly inform whether and how social factors gain prominence in the study of ID transmission.

Participatory research can serve as a framework for action towards social equity. The study of social equity and social processes within ID transmission in the global South requires the participation of local communities and institutions. The field of participatory action research provides a framework for how communities can equitably contribute to research agendas, processes and the dissemination of questions rooted in novel interventions, new policies and changes to social structures [150]. The approach has been utilized in several studies of disease transmission and mitigation within LMICs and often includes strong collaboration with local research institutions [151–157]. The extension of such studies to include social context and social variables is natural and could further strengthen resulting interventions and policies. One of the challenges of participatory action research studies, for the study of social factors, is their relatively small sample sizes. Yet such challenges could be overcome with the inclusion of social factors in several studies, which are then pooled together to understand community or regional processes. In summary, questions which investigate social processes of ID transmission, anchored in populations of the global South and/or minoritized populations, will produce more meaningful and actionable advances when the means by which such data is attained is also thoughtfully examined and justice oriented.

6.3 The analysis of ID social processes requires historical and cross-population analyses.

In a recent commentary, we draw on the foundational work of Drs. Link, Phelan and Clouston to illustrate how fundamental social causes (such as racial capitalism, socioeconomic inequity and discrimination) will continue to play a prominent role in ID outcomes if they are not addressed, even after intervening on proximal mechanisms such as access to vaccines, testing and pharmaceuticals [32, 75, 158]. To study fundamental causes of ID disparities a wider, more inclusive vision of our sample is needed [46]. In the paper [159], we argue that ID research requires more than a "social lens"–it requires a sociological imagination [159]. Utilizing the framework of revered

sociologist, C.Wright Mills, the paper highlights the need for ID researchers interested in social inequity to "grasp history and biography and relations between the two within society" (p11). Researchers are then free to imagine possibilities for the study of social constructs and the like between societies, nations and points in time. In this way, we remove ourselves from the assumption of natural and inevitable racial and social differences–so that we can imagine a society without these constraints.

Studies of fundamental causes of ID transmission can build upon other disciplinary traditions. Few studies exist which evaluate cross-national and historical changes in social conditions in relation to ID outcomes. Although cross-national studies of IDs exist, they seldom evaluate social conditions and context [160–162]. Descriptive analyses of IDs within certain populations throughout history provide important context on how fundamental social causes impact outcomes [163–165]. Researchers have advanced these questions with the limited data sources available [32, 45], including our own work in Aim 1. Nonetheless, the translation of this research into traditional epidemiological analyses is lacking. Clouston et al. [75] in his paper "A Social History of Disease" provides a framework which elucidates fundamental social causes and how they have a natural ebb and flow throughout the course of a particular diseases. Frameworks such as these are useful in the design of studies and the analyses of surveillance data when evaluating fundamental causes of ID outcomes. The framework provides context and encourages the researcher to determine whether and how the four stages of inequality should considered: 1) natural mortality (when marginalized and non-marginalized populations have equitable outcomes) 2) producing inequality (when unnatural inequalities begin due to fundamental causes) 3) reducing inequality (when both groups experience gains in health, though marginalized populations experience gains more quickly) and 4) reduced disease (when inequities return to minimal levels). Overall, there are large bodies of work within sociology, history and demography which we can draw upon as researchers of ID social processes to inform our own research questions, processes and approaches.

In conclusion, the pursuit of data-driven approaches for the study of social processes of ID requires a creative, justice-oriented and sociologically-informed framework. ID transmission is complex, as is the study of fundamental social causes of disease. A poorly informed framework could result in analyses which not only incorrectly classify exposure outcome relationships but also perpetuate the very inequity they seek to eliminate. After the COVID-19 pandemic and the grave inequities it produced, academic interest in the study of social factors ID transmission is growing. We, collectively as the next generation of researchers, can only hope to develop a coherent framework within which questions of social equity and ID can be pursued. As we traverse this new terrain, we should keep in mind that before all else we must set justice before our eyes and "withhold our selves not therefrom" [166].

APPENDIX A

Timing of Infection was a Key Driver of Race/ethnic Disparities in COVID-19 Mortality During the Pre-vaccine Period

A.1 Additional Figures

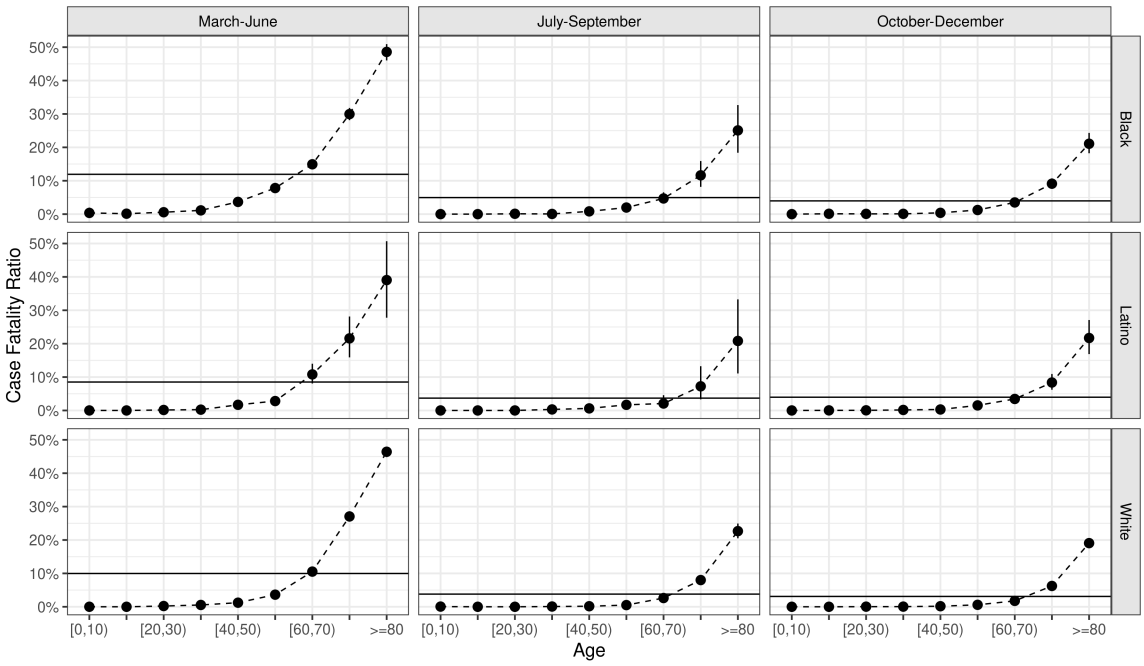

Figure A.1: Age-specific SARS-CoV-2 case fatality rates (CFRs) by race/ethnicity for Michigan residents during three periods of 2020. Each panel represents age-specific CFRs by race/ethnic group (rows) for each of three time periods during the first year of the COVID-19 pandemic (columns). Solid lines in each panel represent the crude (unadjusted) race/ethnic group-specific CFR for all ages. Vertical lines at each point indicate the width of the 95% posterior credible interval (CrI) of the estimated CFR.

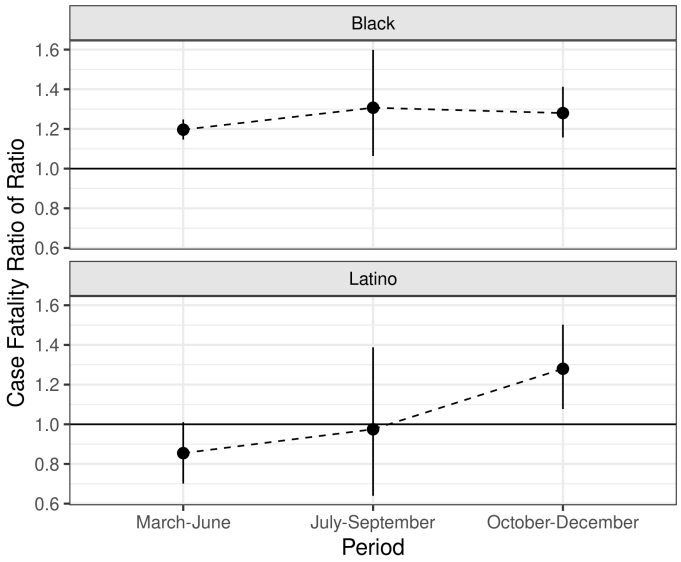

Figure A.2: Modeled race-specific case fatality rate ratio for all ages for Black and Latino residents as compared to White residents of Michigan over three periods in 2020. This ratio indicates how much higher or lower the case fatality of a Black or Latino resident was as compared to White residents during a particular period.

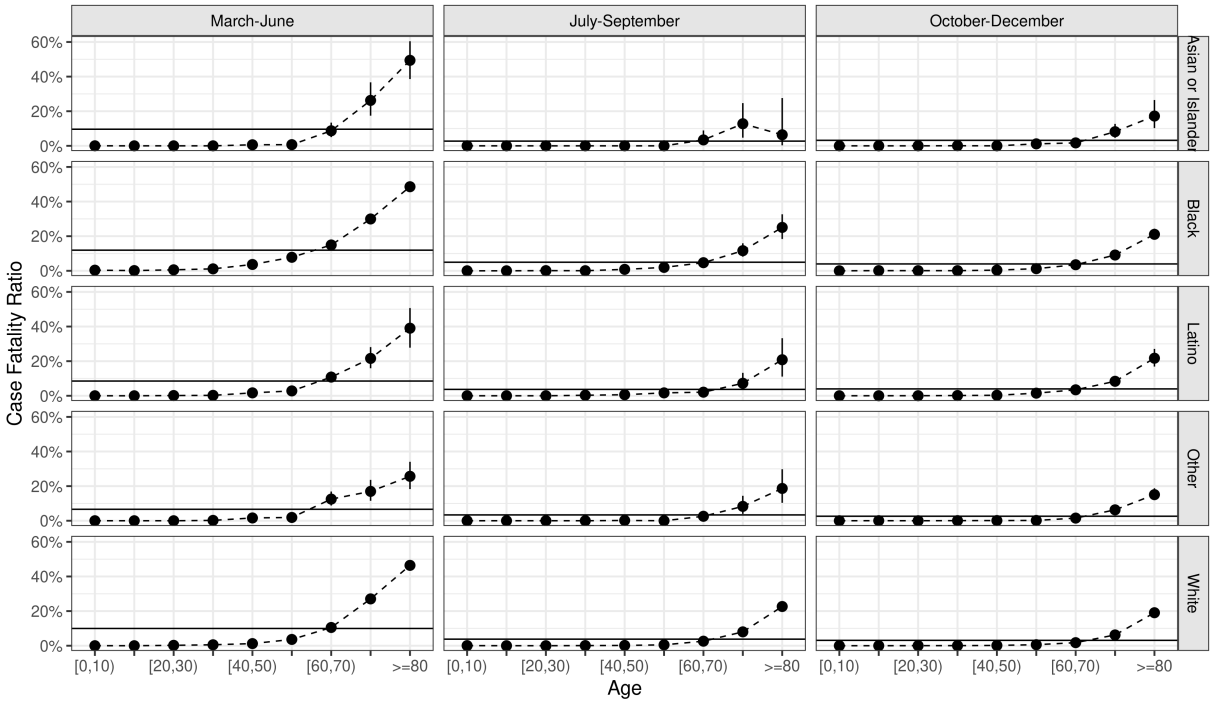

Figure A.3: Modeled age-specific case fatality rate (as a percentage) by race of Michigan residents over three periods in 2020. The solid horizontal line represents the race-specific value for all ages in each time period respectively. The vertical lines around the points indicate credible interval of the case fatality ratio.

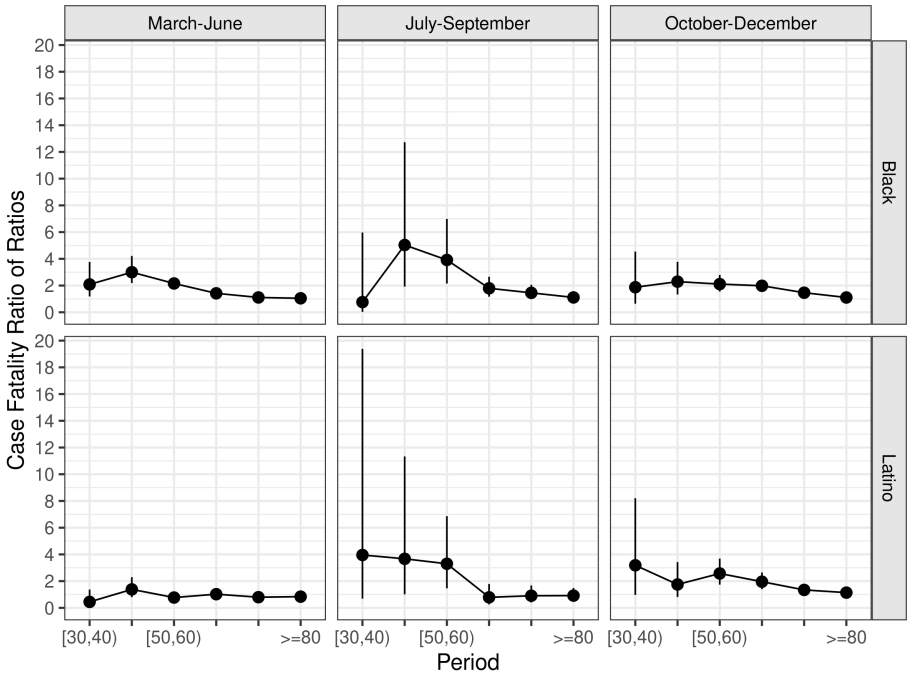

Figure A.4: Modeled race-specific case fatality rate ratio for ages above 30 for Black and Latino residents as compared to White residents of Michigan over three periods in 2020. This ratio indicates how much higher or lower the case fatality of a Black or Latino resident was as compared to White residents during a particular period. The solid vertical lines represent credible interval of the case fatality risk ratio (or case fatality ratio of ratios).

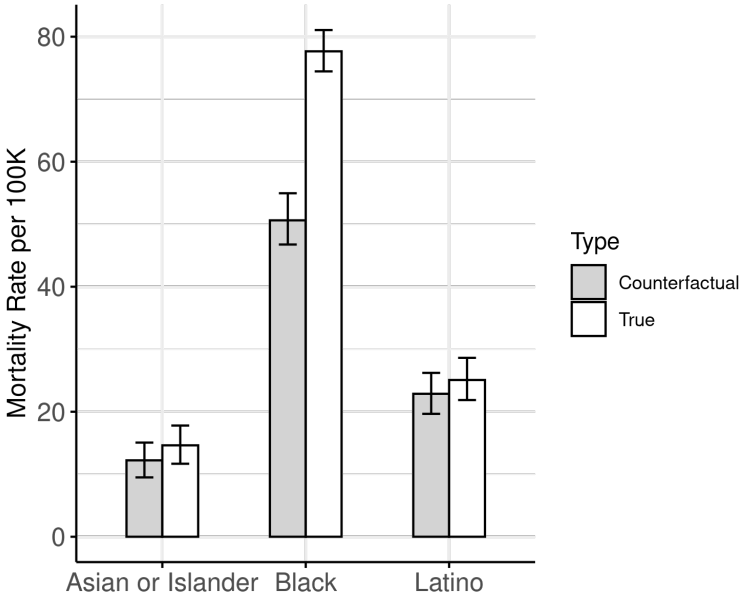

Figure A.5: Counterfactual vs. Observed COVID-19 mortality rates by race/ethnicity in Michigan, 2020. The heigh of each bar illustrates the observed mortality rate (white bars) and an estimate of the group-specific mortality rates that would have resulted if minoritized residents experienced the same pacing of infection as White Michigan residents (gray bars). Vertical brackets represent the width of the 95% posterior credible interval (CrI).

A.2 Sensitivity Analysis

Figure A.6 - A.10 present the results of our sensitivity analysis assessing the robustness of our results to non-differential missingness of race/ethnicity data across race/ethnic groups. Overall, this analysis shows that our results are robust to the possibility of differential missingness of case observations by race/ethnicity.

Figures A.6 & A.10 show that differential risks of missing observations by race/ethnicity overall have small impacts on incidence and incidence rate ratios in the first and second period. These effects are more pronounced in the third period, but their impact on overall mortality rates is limited because of the dramatically decreased case-fatality rates during the final two periods as compared to the first.

Additional sensitivity analysis. Another challenge with surveillance data of large population-wide epidemics occurs if there is significant mortality within the at-risk population, which may result in underestimates of cumulative incidence rates in subsequent periods. We conducted additional sensitivity analyses to determine if the IR, IRR, CFR, or case-fatality rates were robust to the fact that the population denominators in fact changed over time due to COVID-19 mortality within each racial/ethnic group. We found that our indicators were robust to such changes.

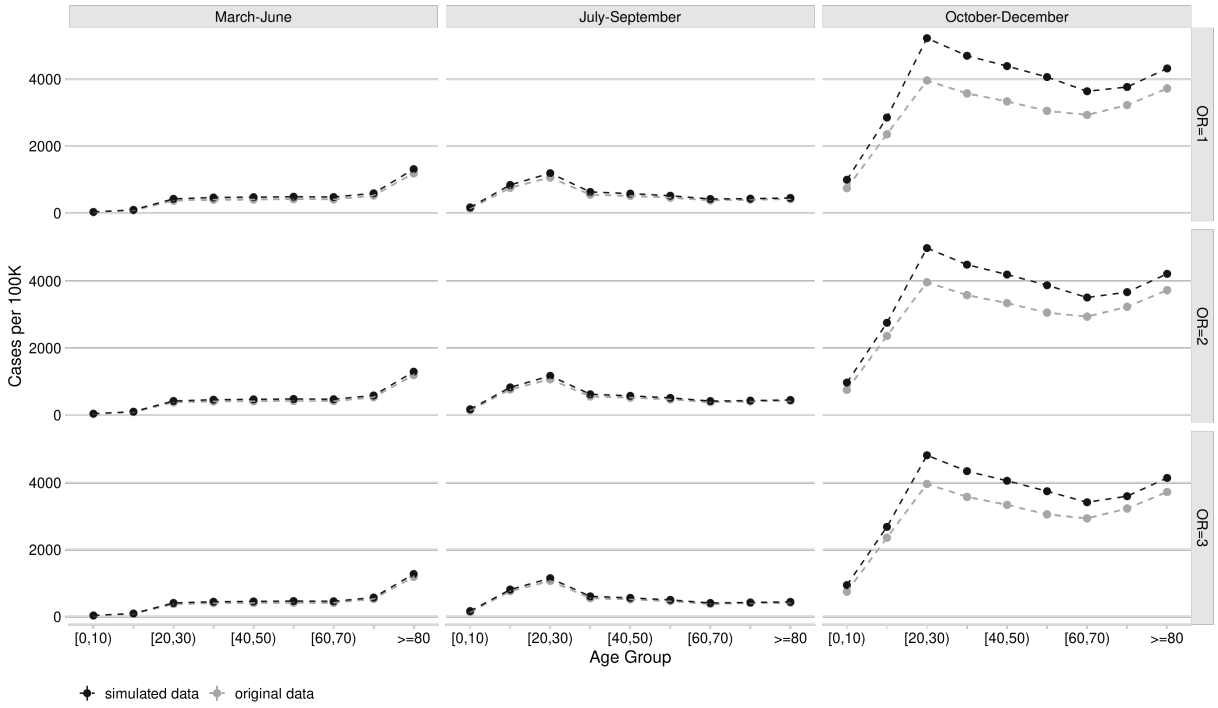

Figure A.6: Modeled age-specific cumulative incidence rate per 100K of COVID-19 infection for White Michigan residents over three periods in 2020. Observed data is shown in grey, and the counterfactual/simulated data is shown in black. The columns progress in terms of the time period. The rows progress in terms of the odds of a Black resident case being missing more than a White resident case. These figures illustrate that as the odds of a missing case being a true Black resident case compared to a White case increase, White residence incidence decreases. The first row illustrates the scenario where the odds of a missing case being White or Black is equal.

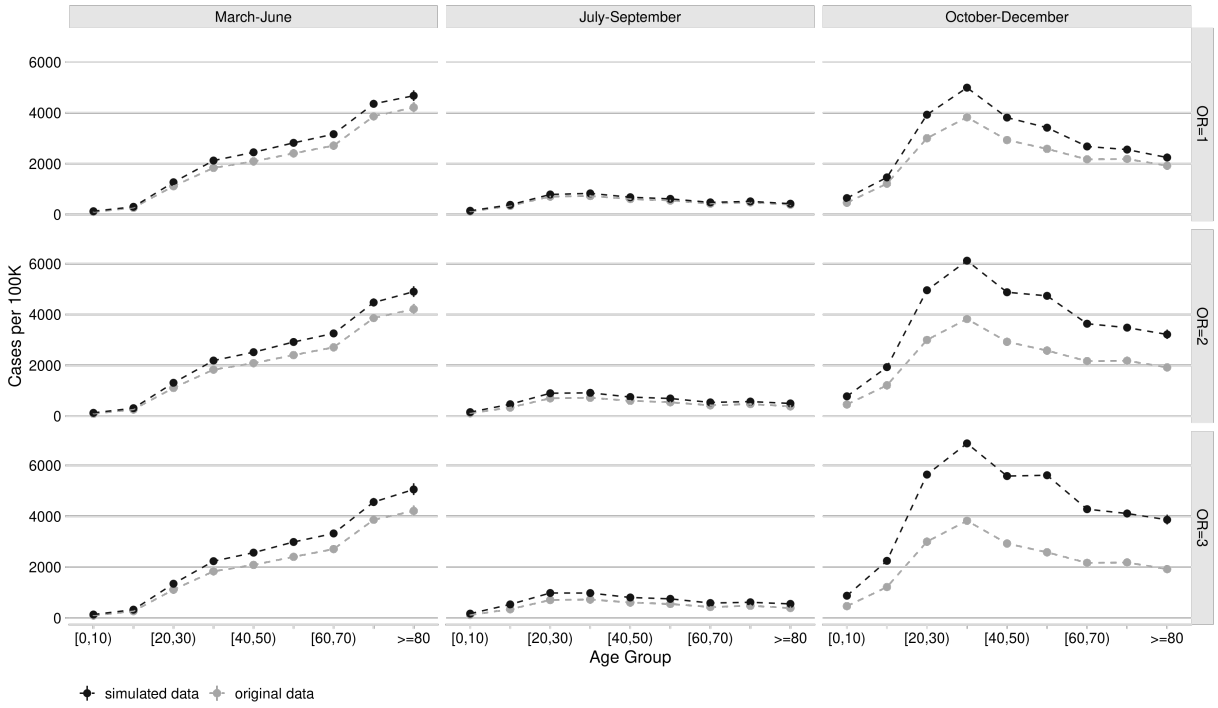

Figure A.7: Age-specific cumulative incidence/100K of COVID-19 infection for Black Michigan residents over three periods in 2020 under three missingness scenarios. The original data (presented in the main body of the paper) is shown in grey, and the counterfactual/simulated data is shown in black. The columns progress in terms of the time period. The rows progress in terms of the odds of a Black resident case being missing more than a White resident case. These figures illustrate that as the odds of a missing case being a true Black resident case compared to a White case increase, Black residence incidence increases. The first row illustrates the scenario where the odds of a missing case being White or Black is equal.

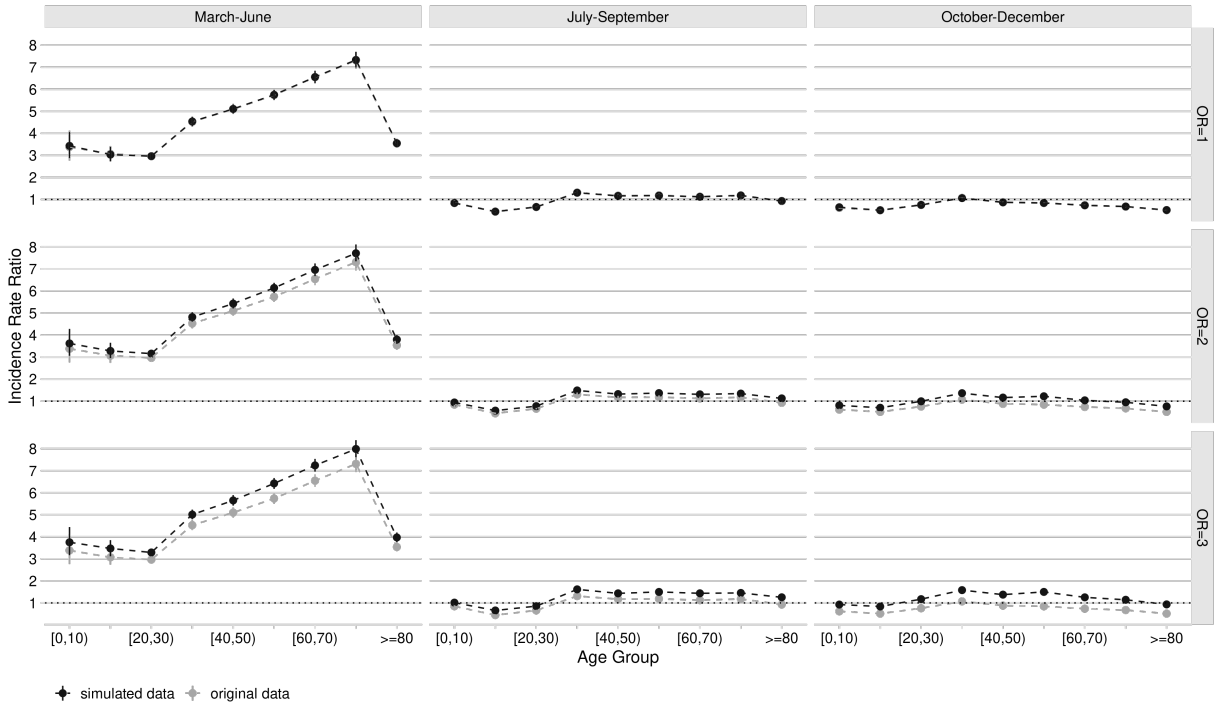

Figure A.8: Age-specific incidence rate ratios for SARS-CoV-2 infection for Black Michigan resident as compared to White residents, over three periods in 2020 under three missingness scenarios. Observed data are shown in grey, and counterfactual/simulated data are shown in black. Each column represents a separate time period. The odds of a Black resident case being missing relative to a White increase with each row, with the top illustrating the scenario in which the odds of a missing case being White or Black is equal. These figures illustrate that as the odds of a missing case being a true Black resident case compared to a White case increase, the Black residence incidence rate ratio increases.

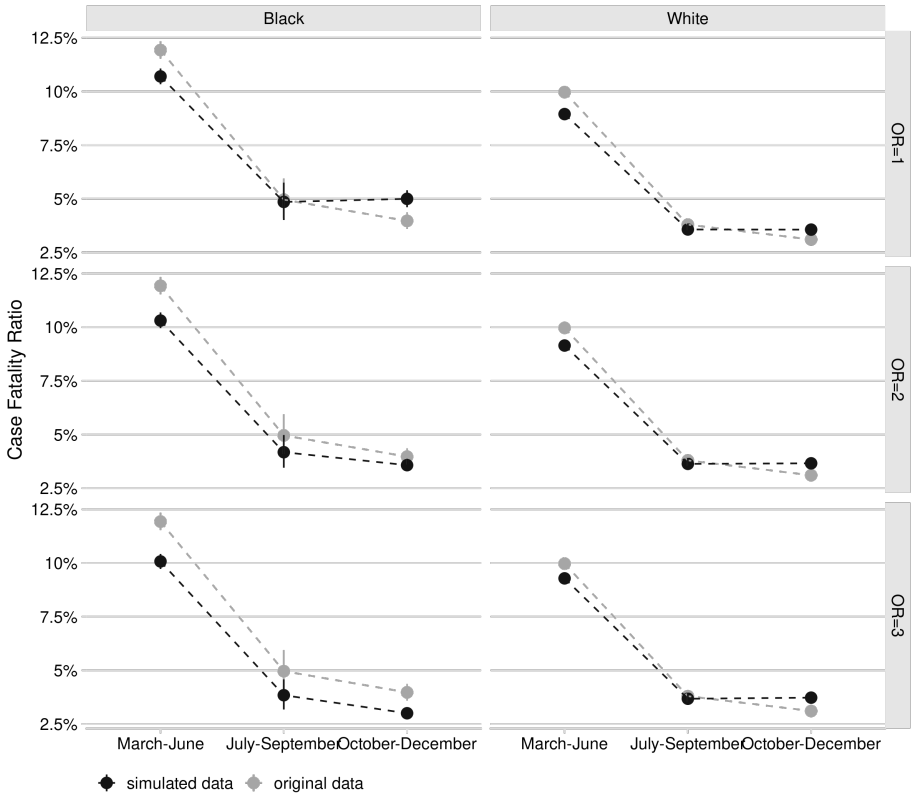

Figure A.9: Modeled race-specific case fatality ratio for all ages for Black and White residents of Michigan over three periods in 2020. This ratio indicates the percentage of cases which resulted in mortality for a given period. The original data (presented in the main body of the paper) is shown in grey, and the counterfactual/simulated data is shown in black. The x-axis progresses in terms of the time period. The rows progress in terms of the odds of a Black resident case being missing more than a White resident case

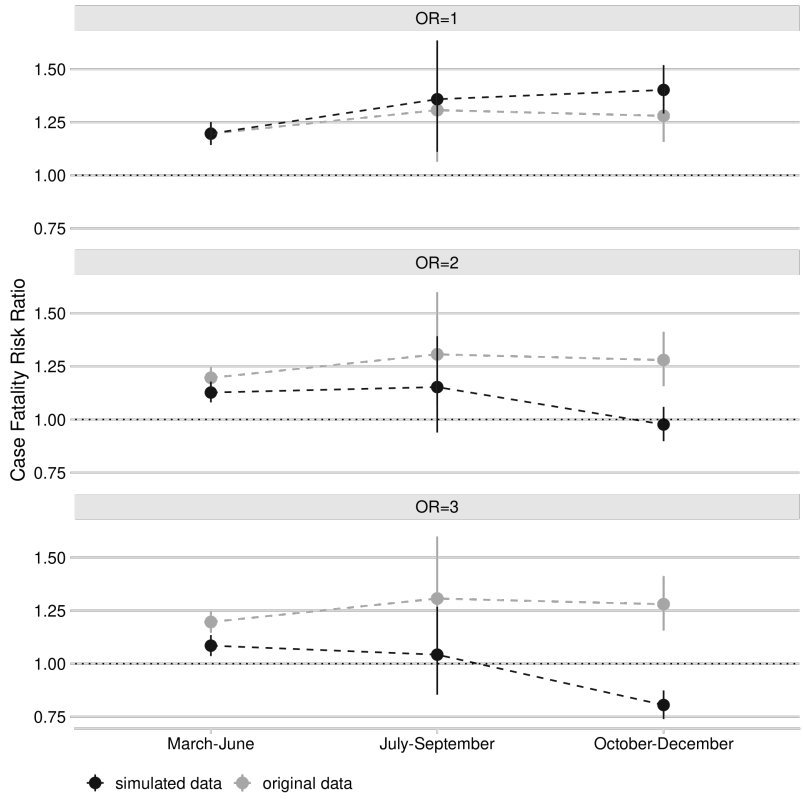

Figure A.10: Modeled race-specific case fatality risk ratio for all ages for Black residents as compared to White residents of Michigan over three periods in 2020. This ratio indicates how much higher or lower the case fatality of a Black resident was as compared to White residents during a particular period. The original data (presented in the main body of the paper) is shown in grey, and the counterfactual/simulated data is shown in black. The x-axis progresses in terms of the time period. The rows progress in terms of the odds of a Black resident case being missing more than a White resident case.

Table A.1: Missing Race/ethnicity Values within Dataset

Period	**Cases**	**Deaths**	**% Cases of Total**	**% of Deaths of Total**
July-September	7639	11	11.87%	1.26%
March-June	9564	153	13.98%	2.39%
October-December	83072	542	22.10%	9.36%

Table A.2: Races within Latino ethnicity within cases of COVID-19 in 2020 in Michigan

Race	Latino	% of Total Latino	Non-Latino	% of Total Non-Latino	Total in Each Race
Mixed	5	0.02	30	0.01	35
Asian	94	0.33	7237	1.52	7331
Black	726	2.53	60584	12.71	61310
Other	12963	45.18	19586	4.11	32549
Unknown	1886	6.57	94790	19.89	96676
White	13016	45.37	294368	61.76	307384
Total	28690	100.00	476595	100.00	505285
% Latino of Total		5.68		94.32	

APPENDIX B

Spatio-temporal Patterns of Disability in India: 2001-2011

B.1 Additional Tables

Type of Disability	P-Value	2001	2011	Change from 2001 to 2011
Total	0.001	0.515	0.551	0.449
Seeing	0.001	0.353	0.515	0.479
Hearing	0.001	0.370	0.679	0.362
Speech	0.001	0.509	0.732	0.563
Movement	0.001	0.721	0.793	0.615
Mental Retardation	0.001	0.717		
Mental Illness	0.001	0.707		
Other	0.001	0.440		
Multiple	0.001	0.657		
Mental	0.001		0.709	0.483

Table B.1: Global Moran's I of Disability in India across two years along with change

Type	Definition
Seeing	A person who cannot see at all (has no perception of light) or has blurred vision even with the help of spectacles will be treated as visually disabled and code 1 will be entered under this question. A person with proper vision only in one eye will also be treated as visually disabled. You may come across a situation where a person may have blurred vision and had no occasion to test whether her/his eyesight would improve by using spectacles. Such persons would be treated as visually disabled.
Speech	A person will be recorded as having speech disability, if she/he is "dumb". Similarly persons whose speech is not understood by a listener of normal comprehension and hearing, she/he will be considered to having speech disability and code 2 will be entered. This question will not be canvassed for children up to three years of age. Persons who stammer but whose speech is comprehensible will not be classified as disabled by speech.
Hearing	A person who cannot hear at all (deaf) or can hear only loud sounds will be considered as having hearing disability and in such cases code 3 be entered. A person who is able to hear, using hearing-aid will not considered as disabled under this category.
Movement	A person who lacks limbs or is unable to use the limbs normally, will be considered having movement disability and code 4 will be entered here. Absence of a part of a limb like a finger or a toe will not be considered as disability. However, absence of all the fingers or toes or a thumb will make a person disabled by movement. If any part of the body is deformed, the person will also be treated as disabled and covered under this category. A person who cannot move herself/himself or without the aid of another person or without the aid of stick, etc., will be treated as disabled under this category. Similarly, a person would be treated as disabled in movement if she/he is unable to move or lift or pick up any small article placed near her/him. A person may not be able to move normally because of problems of joints like arthritis and has to invariable limp while moving, will also be considered to have movement disability.
Mental	A person who lacks comprehension appropriate to her/his age will be considered as mentally disabled. This would not mean that if a person is not able to comprehend her/his studies appropriate to her/his age and is failing to qualify her/his examination is mentally disabled. Mentally retarded and insane persons would be treated as mentally disabled. A mentally disabled person may generally depend on her/his family members for performing daily routine. It should be left to the respondent to report whether the member of the household is mentally disabled and no tests are required to be applied by you to judge the member's disability.

Table B.2: Defintion of Disability in the Indian Census 2001

B.2 Additional Figures

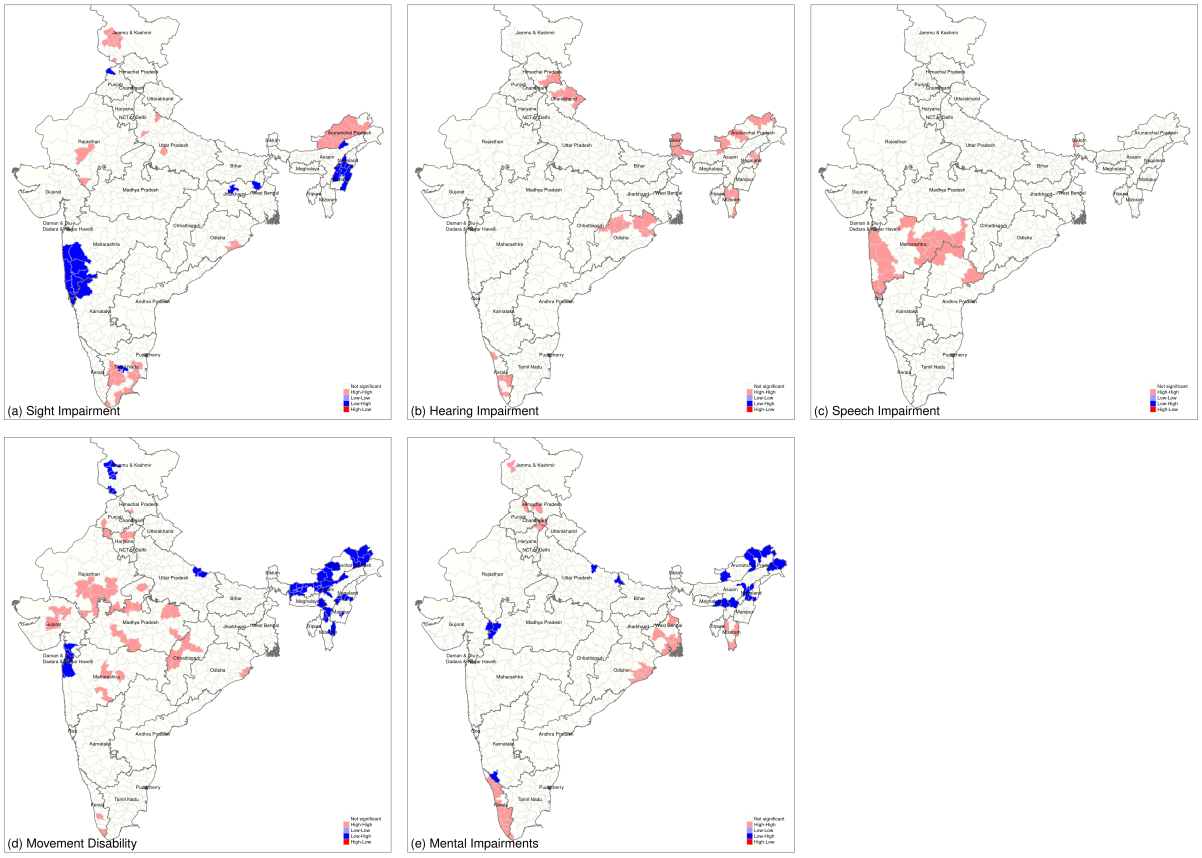

Figure B.1: Clustering of disability prevalence by type in the 2001 Indian Census at the district level as measured by the local indicator of spatial autocorrelation (LISA) with correction for multiple testing bias. Light red indicates high-high clustering, light blue indicates low-low cluster, dark blue indicates low-high clustering and dark red indicators high-low clustering.

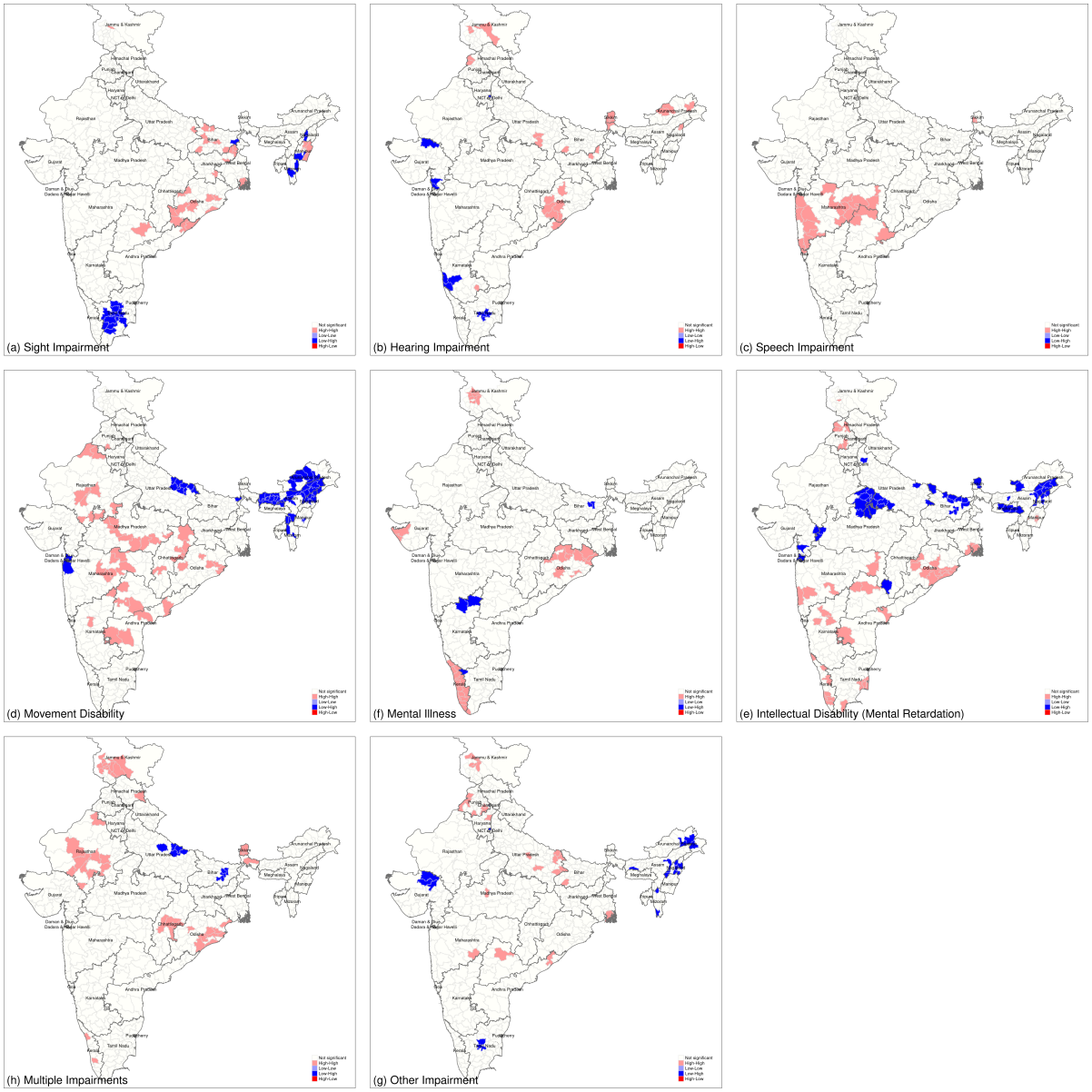

Figure B.2: Clustering of disability prevalence by type in the 2011 Indian Census at the district level as measured by the local indicator of spatial autocorrelation (LISA) with correction for multiple testing bias. Light red indicates high-high clustering, light blue indicates low-low cluster, dark blue indicates low-high clustering and dark red indicators high-low clustering.

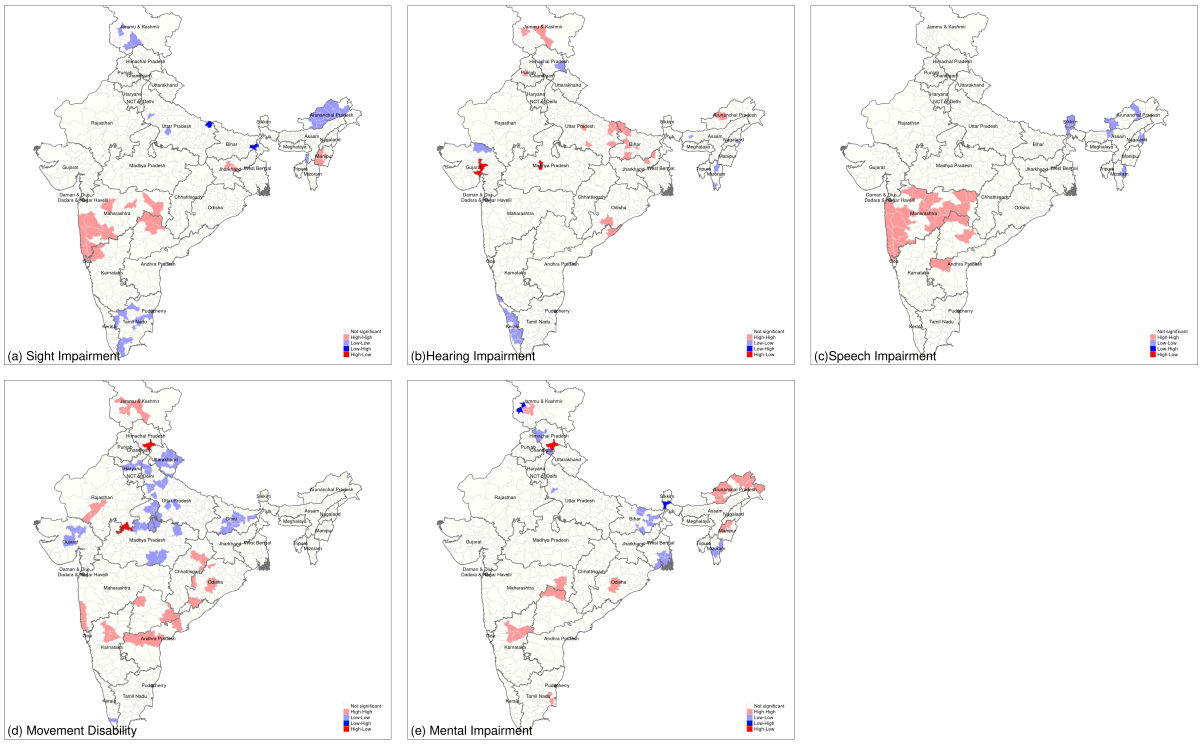

Figure B.3: Clustering of change in disability prevalence by type in the 2001 Indian Census compared to the 2011 Indian Census at the district level as measured by the local indicator of spatial autocorrelation (LISA) with correction for multiple testing bias.

B.3 Additional Methodology

Methodology to Compare 2001 to 2011 Districts Given Boundary Changes The development of the merged data set required many decisions with regard to weighting. For those districts that were not carved out into more than one district, we weighted the population based on the square kilometers (sq km) of the district. We start with those districts that were carved from. We multiplied the 2001 population with the sq km of the 2011 new district divided by the sq km of the 2001 district. In most cases, this weighting resulted in a perfect match (where the sq km from the 2001 district equaled the sum of the newly created district and the 2011 district sq km area). There were although some discrepancies such as for the newly created district Ganderbad from Srinagar and Shupiyan from Pulwama. In those cases, we utilized the values of the square footage reported and attributed these to reporting error. If a district was carved out of more than one district then we used the 2001 population of both districts summed together and then weighted this by the fraction of the sq km of the new 2011 district and the sum of the sq km of the 2001 districts from which it was carved. If two districts were carved of the same district (and that district had another district from which it was carved out of) then we utilized the overlap in the map to determine the square

km contributed from each old district. If one district was carved out from three districts (and those districts were not used to make any other districts) then the difference in the values between original district population and the weighted district population were taken. There were other cases when the sq km of the 2011 district was larger than the sq km of the 2001 district. In the example of Jehanabad in Bihar. Since the difference was 931 and the value was 3968 we assumed that this was a typo and removed the 9. It seems that some states have more accurate sq km data than others. For example, Arunachal Pradesh matched up completely. If there were two districts which were carved out of one district only, then we multiplied the difference in the 2001 population and the weighted population and multiplied by the fraction of the sq km of the new village and the difference between the sq km of the district from which it was carved out from (2001 and 2011). One districts the sq km value was larger in 2011 and there were not an immediate fix possible such as for Kokrajhar in Assam. In this case we assumed that the district sq km did grow and weighted accordingly. There were two new districts in Assam for which it would not possible to calculate the weight, Baksa and Chirang because they were carved out of the same district (and other districts also were carved from to create these districts). Therefore, there were two unknowns in the problem and we could not solve for how many people would have been in this sq km of the district. In this case we utilized a density weight where we summed all the people from the districts from which the new district was carved divided by the sq km of the original districts from which it was carved and obtained person per sq km for each variable which was then multiplied by the sq km of the new district.

APPENDIX C

Do Sociality and Spatiality Interact? Characterizing the Joint Impact of Spatial Proximity and Social Network Connectivity on Acute Infectious Disease Risks and Intervention Efficacy

C.1 Additional Figures for Interaction Analysis

C.1.1 Censused Population Comparison to Sample Population

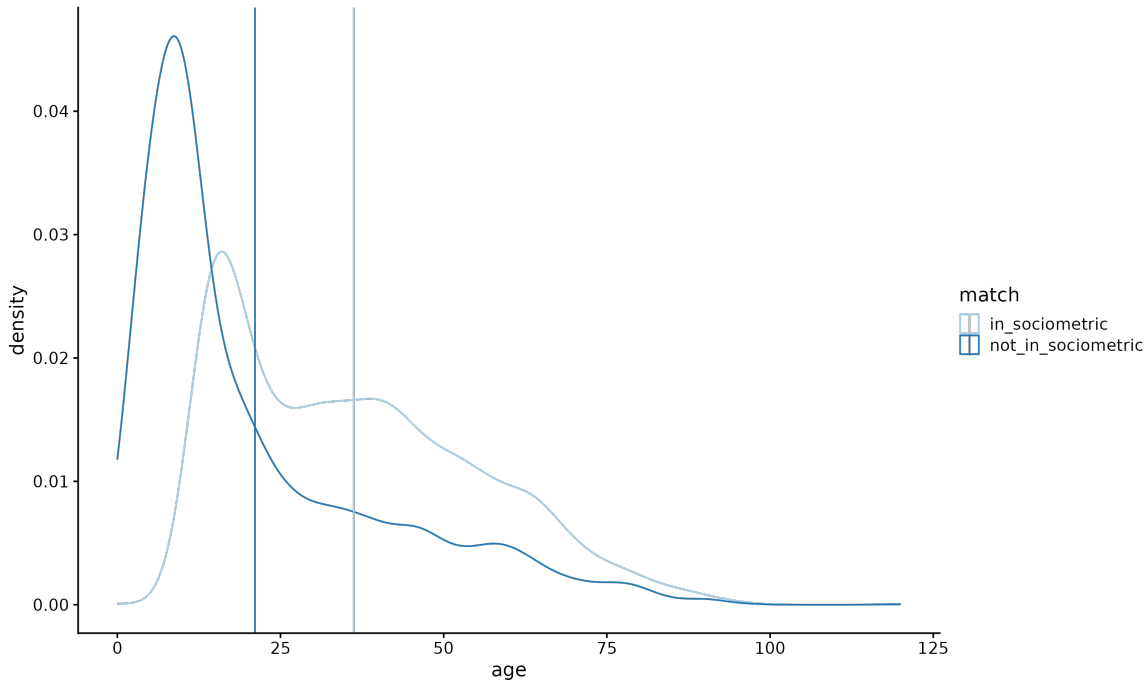

Figure C.1: Differences between the study population and the censused population in all six communities by age.

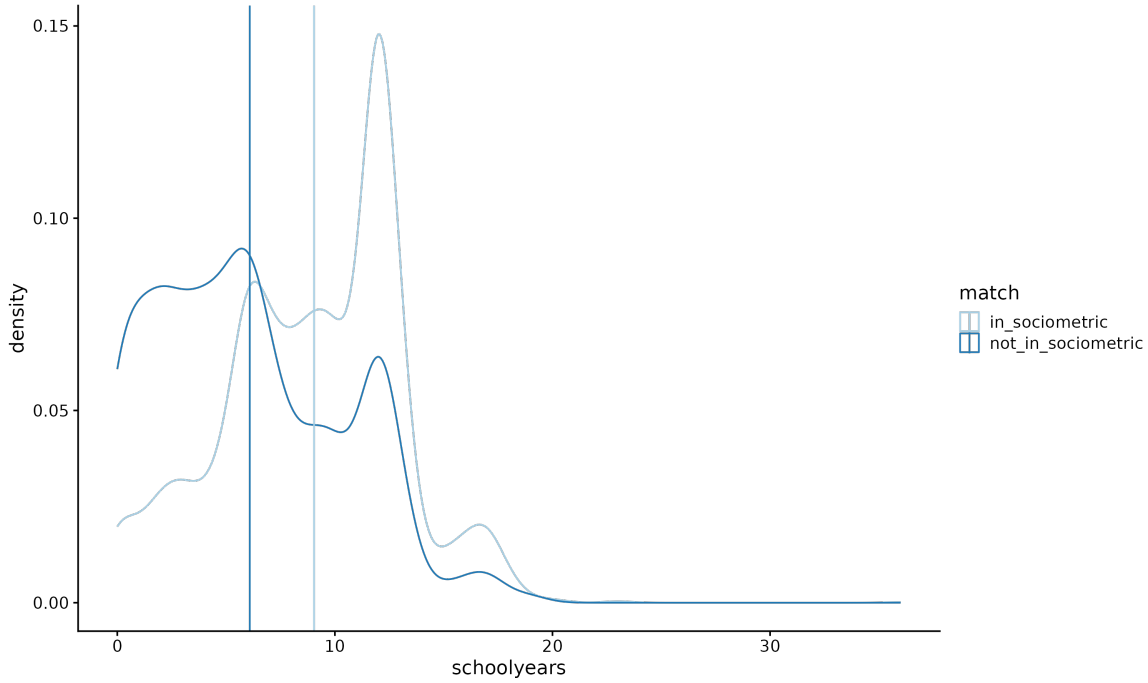

Figure C.2: Differences between the study population and the censused population in all six communities by number of years of schooling.

C.1.2 Histograms of contact density across two types networks

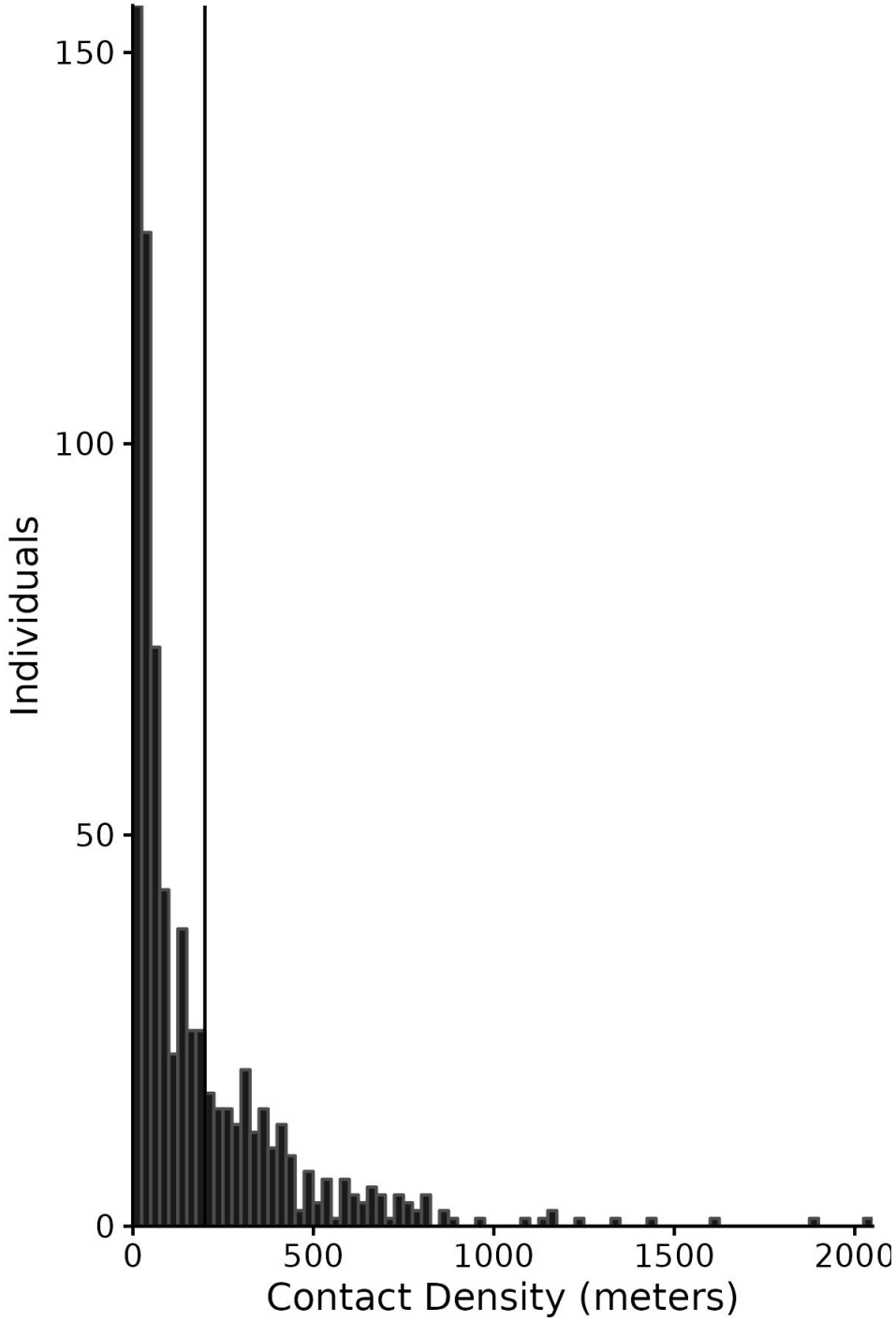

Figure C.3: Histogram of harmonic mean distance (meters) from contacts in the people visited network in Borbon. The vertical line represents the mean (199.4368 meters).

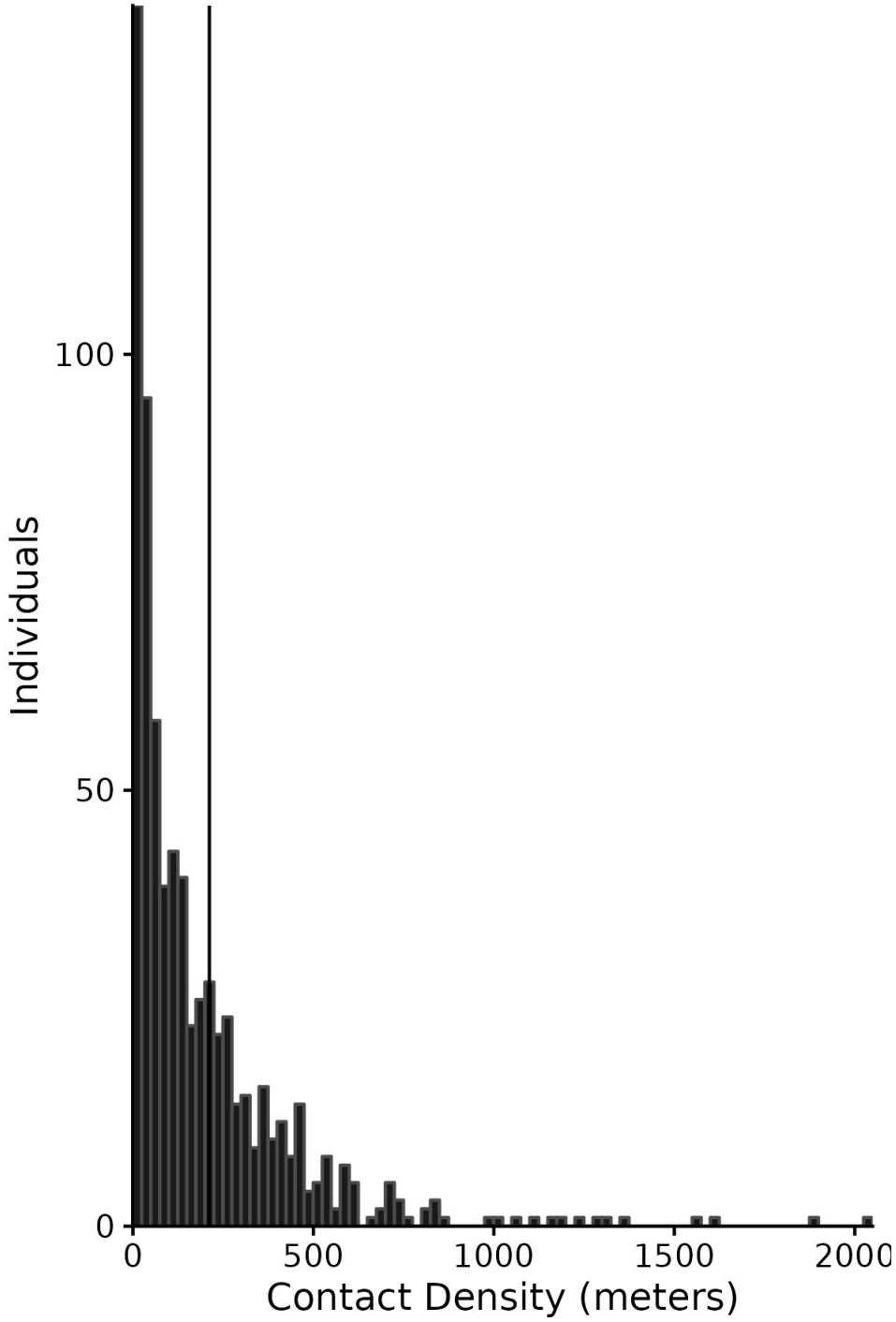

Figure C.4: Histogram of harmonic mean distance (meters) from contacts in the people visited network in Borbon. The vertical line represents the mean (211.8205 meters).

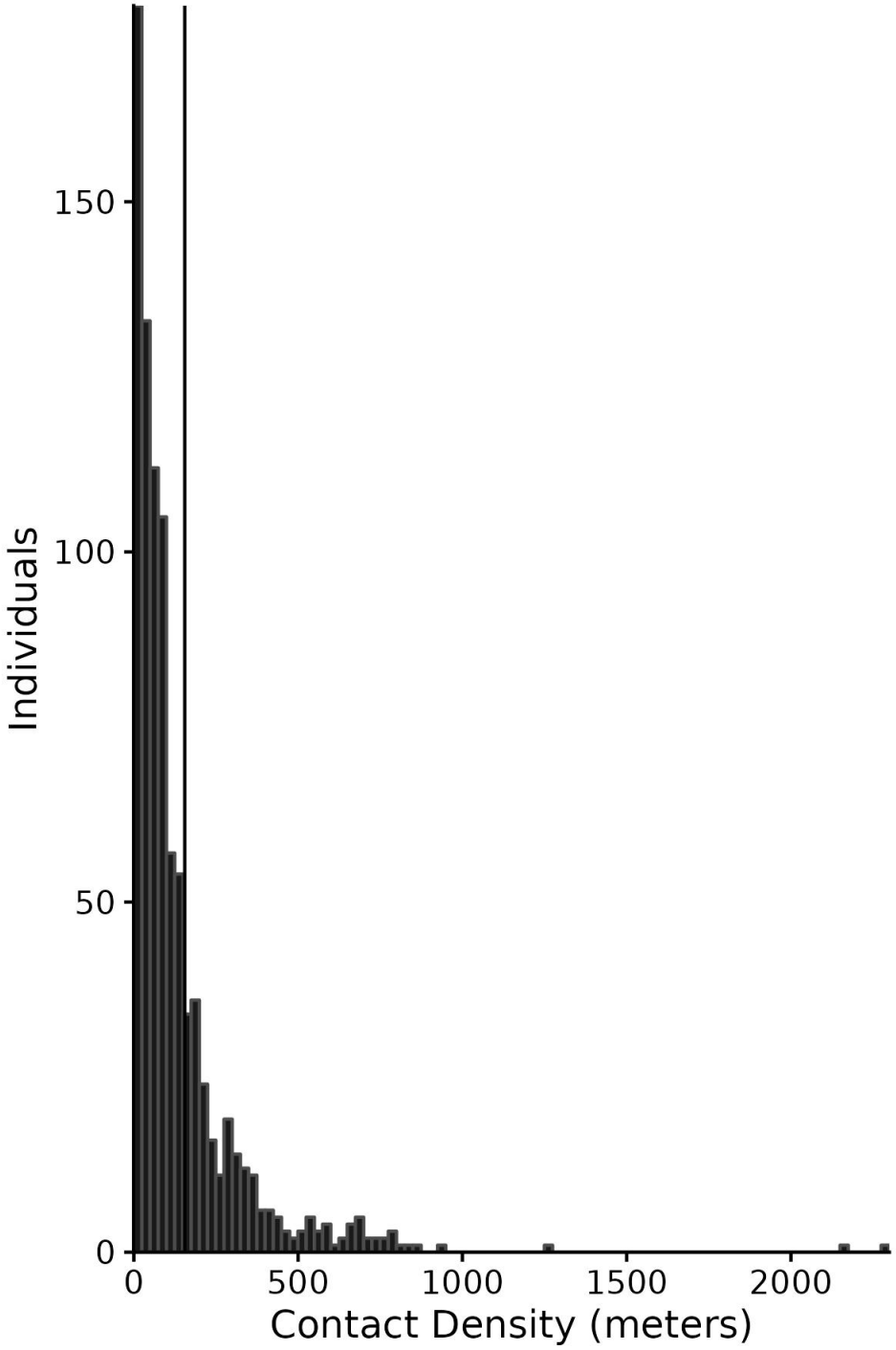

Figure C.5: Histogram of harmonic mean distance (meters) from contacts in the general important matters network in Colon Eloy, Timbire, Maldonado, Santo Domingo and Santa Maria. The vertical line represents the mean (155.0075 meters).

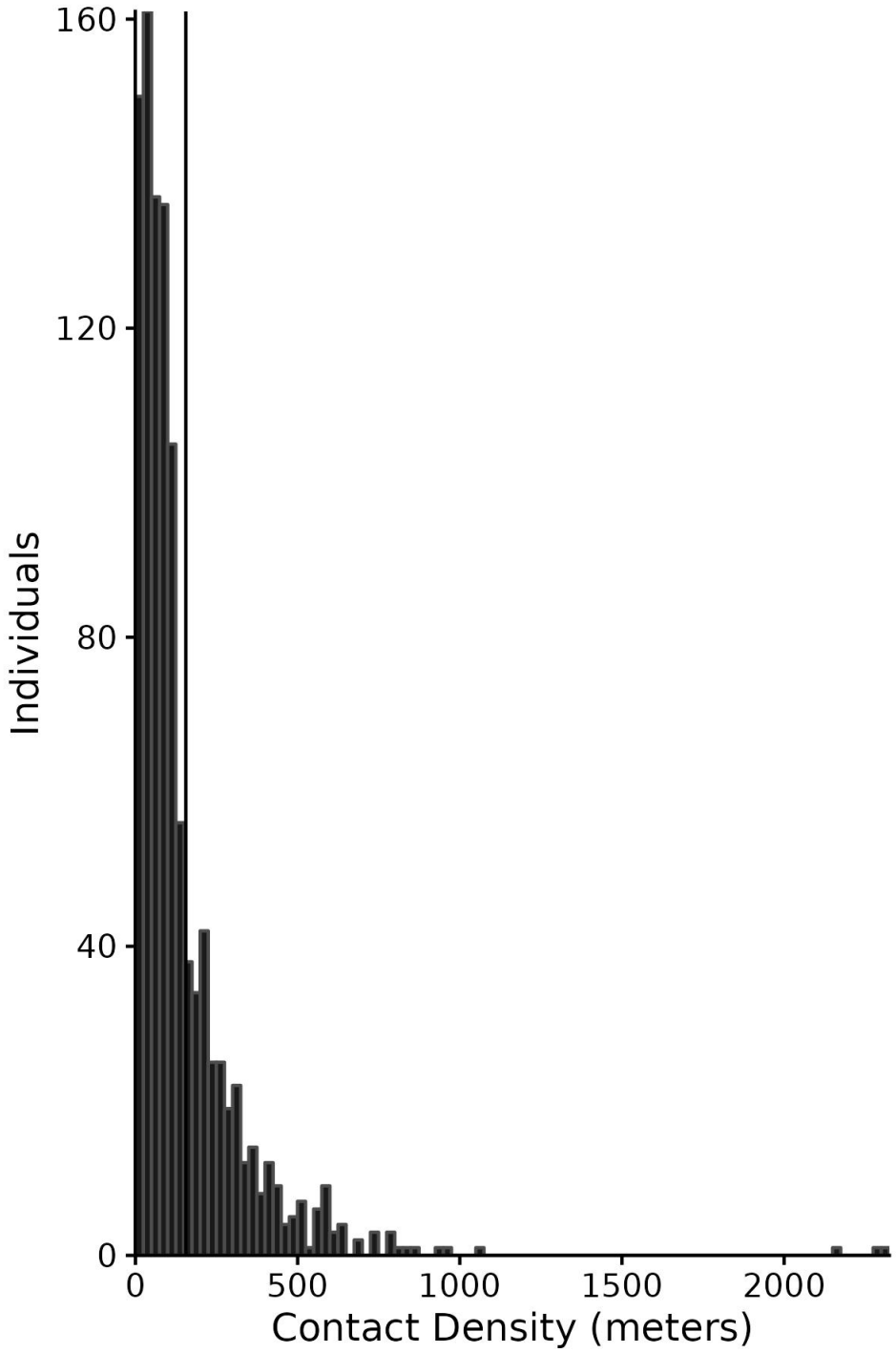

Figure C.6: Histogram of harmonic mean distance (meters) from contacts in the people visted network in Colon Eloy, Timbire, Maldonado, Santo Domingo and Santa Maria. The vertical line represents the mean (155.3287 meters).

C.1.3 Additional Model Results for Interaction Model

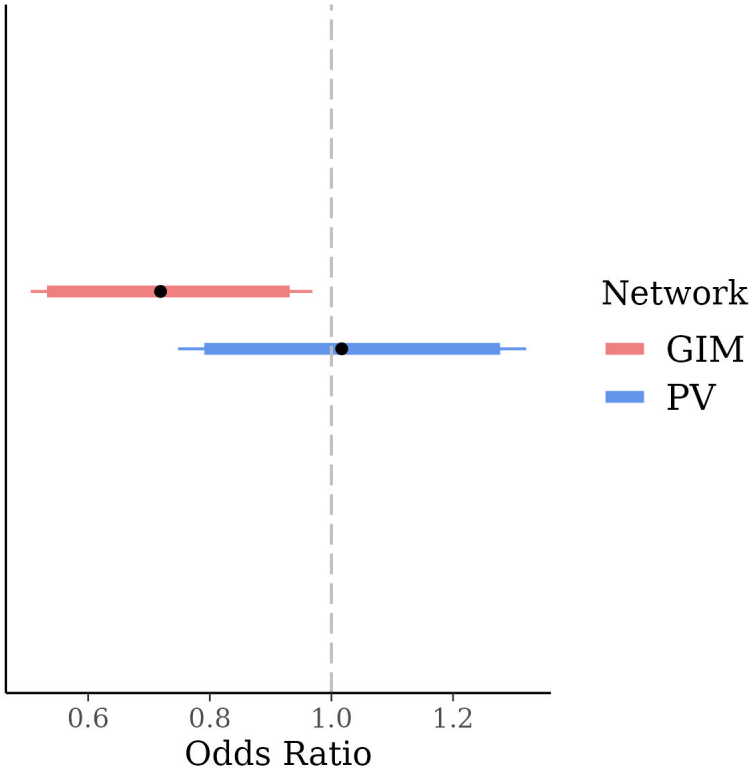

Figure C.7: Odds of fever and/or diarrhea in Borbon per one unit increase in the number of individuals in respondents' network (degree). Results of Bayesian logistic regression with confounders. Credible intervals at the 95% and 90% level.

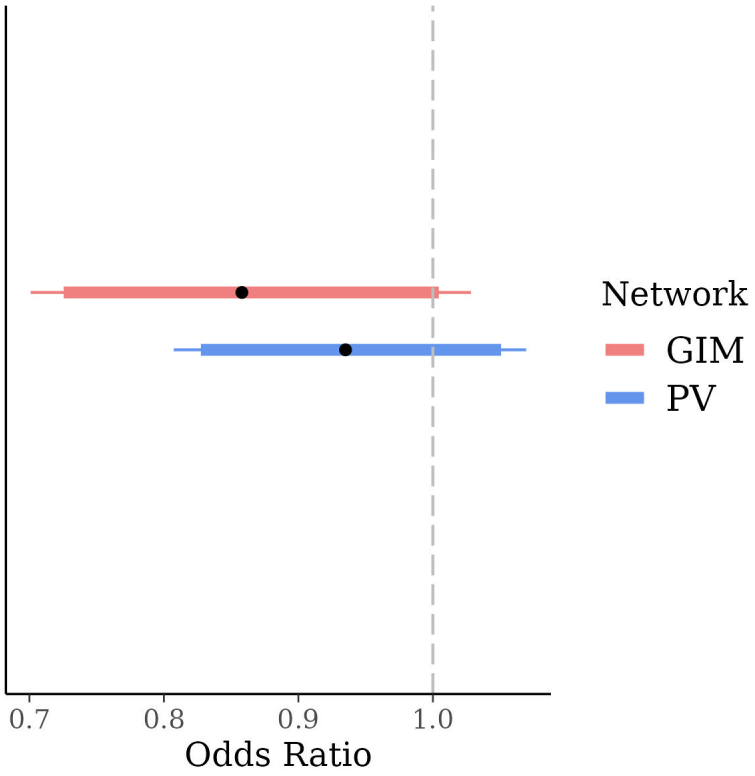

Figure C.8: Odds of fever and/or diarrhea in Colon Eloy, Timbire, Maldonado, Santo Domingo and Santa Maria per one unit increase in the number of individuals in respondents' network (degree). Results of Bayesian mixed effects logistic regression with confounders. Credible intervals at the 95% and 90% level.

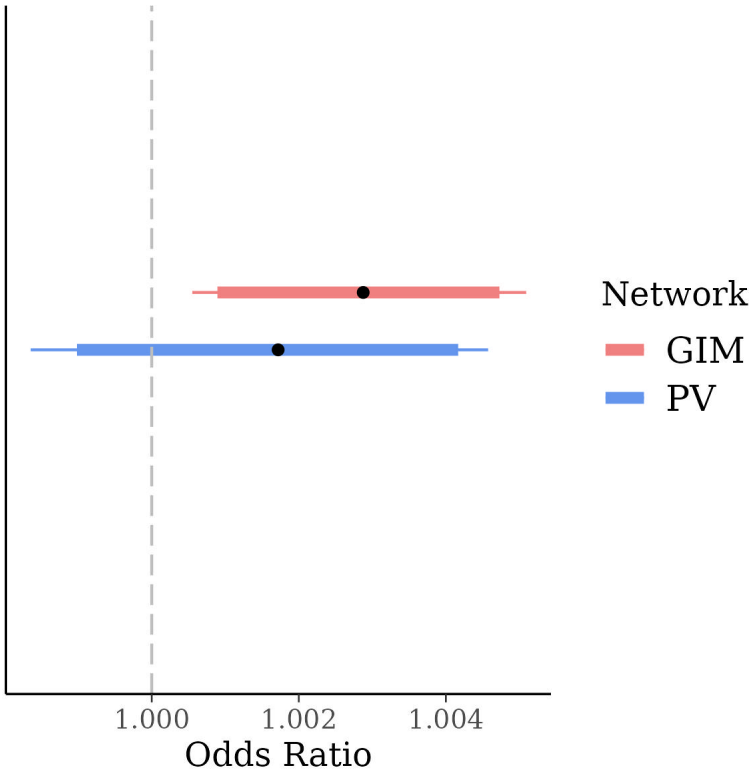

Figure C.9: Odds of fever and/or diarrhea in all communities per one meter increase in distance (harmonic average) from all other individuals in respondents' community. Results of Bayesian mixed effects logistic regression with confounders. Credible intervals at the 95% and 90% level.

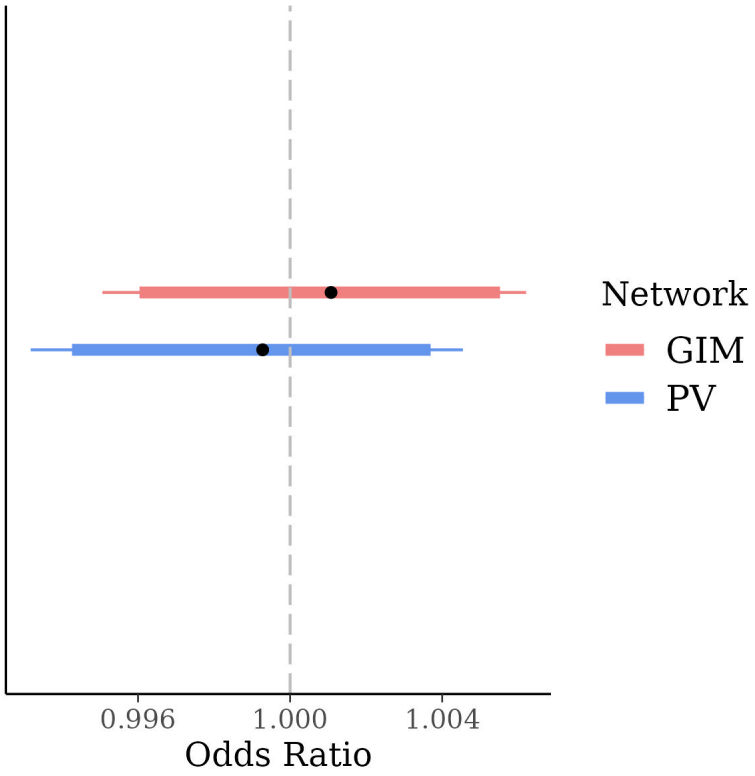

Figure C.10: Odds of fever and/or diarrhea in all communities per one meter increase in distance (harmonic average) from all other individuals in respondents' community. Results of Bayesian mixed effects logistic regression with confounders. Credible intervals at the 95% and 90% level.

C.2 Contact Density Analysis

C.2.1 Contact Density Methodology

In this analysis we explore a novel measure of harmonic mean distance to contacts, which we call contact density. This measure is adapted from a previous approach from Sarkar et al. [167] called spatial degree.

To calculate contact density, we first generate adjacency matrices for each of the two networks in our study, **G** for the general important matters network and **P** for the people visited network. This adjacency matrix includes every ego (survey respondent) within the rows and columns (square matrix). Therefore, the GIM network is represented by a [2,688 X 2,688] matrix and the PV network by a [2,778 X 2,778] matrix with zeroes down the diagonal and ones whenever the individuals represented in the row and/or column state connection (the survey respondent states that the individual is within their network; i.e. an ego-alter pair). This matrix is multipled by a pairwise distance matrix of similar dimensions. The distance matrices for the GIM and PV networks are

denoted by $\mathbf{H}$ and $\mathbf{Q}$ respectively. Therefore we obtain pairwise contact density values with the following equations (see Equation C.1 and C.2)

$$[\mathbf{GH}]_{ij} = g_{i,1}f_{1,j} + g_{i,2}f_{2,j} + ... + g_{i,n}f_{n,j} = \sum_{r=1}^{n} g_{i,r}f_{r,j} \tag{C.1}$$

$$[\mathbf{PQ}]_{ij} = p_{i,1}\,q_{1,j} + p_{i,2}\,q_{2,j} + ... + p_{i,n}\,q_{n,j} = \sum_{r=1}^{n} p_{i,r}\,q_{r,j} \tag{C.2}$$

To obtain the harmonic mean for each individual respondent, we inverse all the values in the matrices described in C.1 and C.2 by dividing one by the result of each ij value within $\mathbf{GF}$ and $\mathbf{PQ}$. We then sum over each individual row (see Equation C.3) to obtain an individual-level measure of contact density.

$$v = (v_i)_n^{i=1} \rightarrow v_i = \sum_{j=1}^{n} x_{ij} \tag{C.3}$$

We therefore obtain a vector v_i where i represents an ego and divide this value by the number of contacts within each ith individuals' network (degree). This measure allows us to quantify the average distance of a respondents contacts. It combines information on who the respondent considers as within their network and where those contacts are located. Therefore a high contact density value indicates that the respondent has social contacts who live in very different physical environments from themselves (or outside of their neighborhood or barrio) and a low contact density value indicates that the respondent has social contacts who share a similar physical environment to themselves.

To explore whether the geographic distance of one's social contacts matters we generate maps to determine the relationship between where in geographic location one's social contacts are and risk of disease. We then implement a Bayesian generalized logistic mixed effects model with a weakly informative prior to determine the association between the harmonic mean distance of one's contacts and disease risk for our multi-community analysis and without mixed effects for the single community analysis.

We first aggregate our social network and illness data to the household level to allow for social network and geospatial mapping. We then generate the social network from the household adjacency matrix (matrix of all households which are connected) with illness status within the vertices (household aggregated illness status of each survey respondent) of the network graph using the *igraph* package of R. Cases of fever and/or diarrhea at the household level are colored in red. We visualized each separate community social network using the Kamada Kawai network layout algorithm which separates large and small clusters of socially connected individuals within the

network. We then visualize the network using household GPS coordinates (mapping the social network to geographic space) of the individual nodes (survey respondents). Case households and those socially connected to case households are visualized by a darker grey edge connection between nodes. We hope to visually capture any patterning–whether those who are ill have social contacts who are far way compared to those that are healthy who have contacts who are nearby. After our exploratory analysis, we hope to better understand how spatial and social factors interact to affect transmission of enteric disease in all six of our study sites.

In the GIM network there were a total of 6,345 pairs of survey respondents (hereafter referred to as egos) and individuals they named as someone whom they share important matters with (hereafter referred to as alters). Please see Appendix Figure C.11 for an illustration of our final sample. Several of these pairs were removed because the alter was not found in our census file at the time of the survey. Of those 2,969 alter-pairs removed, 1,040 alters lived in another community outside of the six study communities and 1,929 resided within our study sites but could not be found in our census, due to changes in the census membership between the time of survey administration and census capture or an a name provided by the respondent which differed from that recorded in our census. Of those 3,376 alters that were found in the census a few did not have GPS coordinates associated with their household leaving, 2,968 pairs. Several of these pairs were removed because the ego did not have appropriate residency status (resided in the community for more than three months), did not complete the survey and/or had missing GPS coordinates did not reside in the same community, leaving 2,862 ego-alter pairs in the final dataset. For the purposes of mapping, fever and/or diarrhea data was required of all nodes (potential individuals in the social network). Therefore, we could only utilize alters who took the sociometric survey, this led to a mapping sample of 2,375 pairs.

In the PV network there were a total of 5,199 pairs of egos and alters. Please see Appendix Figure C.11 for an illustration of our final sample. Several of these pairs did not have an alter with a valid studyID, leaving 3,397 pairs in the dataset. Several alters also were not located in the census due to changes in membership, leaving 3,063 pairs. Further, eight of these pairs had an alter without a GPS coordinate associated with their study ID and 55 of these pairs were removed because the ego: 1) did not live in the community for more than three months, 2) had an incomplete survey or 3) missing GPS coordinates. This led to a final dataset of 2,990 ego-alter pairs, of which 2,577 were used for mapping purposes because both the ego and alter took the sociometric survey allowing us to record outcome data for both.

The regression models are detailed in the equations below. The first equation is for model of peripheral communities and the second equation is for the model of Borbon.

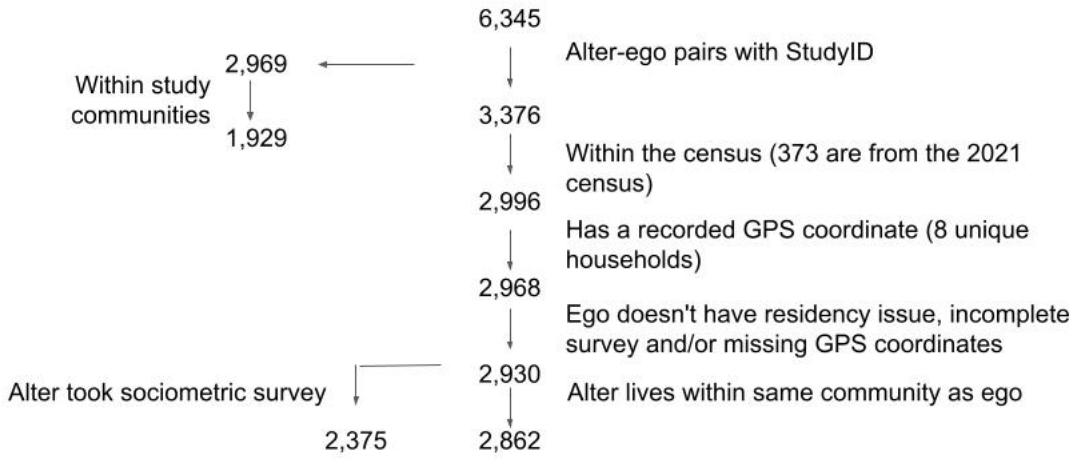

Figure C.11: Flow chart of individual ego-alter pairs (edges) in 2022 sociometric survey for the general important matters network and final data set.

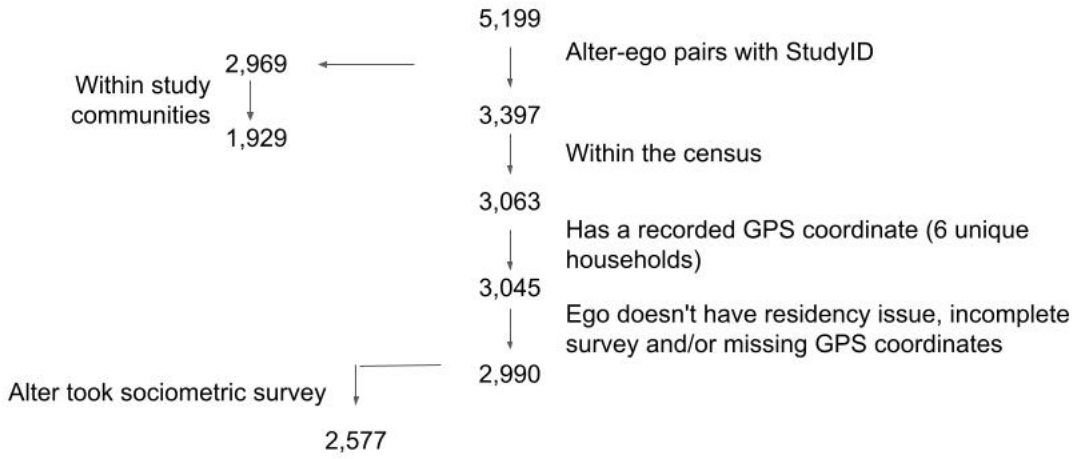

Figure C.12: Flow chart of individual ego-alter pairs (edges) in 2022 sociometric survey for the people visited network and final data set.

$$
\begin{aligned}
Y_{ij} = (\beta_0 + \mu_{0j}) + \beta_1 X_1 + \beta_2 X_2 + \\
\beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + u_j + e_{ij} \\
i = \text{ego}, j = \text{community} \\
X_1 = \text{contact density}, X_2 = \text{race} \\
X_3 = \text{sex}, X_4 = \text{literacy}, X_5 = \text{age} \\
u_j = \text{random intercept of community}
\end{aligned}
\tag{C.4}
$$

$$
\begin{aligned}
Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \\
+ \beta_4 X_4 + \beta_5 X_5 + e_i \\
i = \text{ego}, X_1 = \text{contact density}, X_2 = \text{race} \\
X_3 = \text{sex}, X_4 = \text{literacy}, X_5 = \text{age}
\end{aligned}
\tag{C.5}
$$

C.2.2 Contact Density Results

Contact density differs by illness status The average contact density within the GIM and PV network in Borbòn were 199.44 and 211.82 meters, respectively (see histograms in Figures C.3 and C.4). The average contact density within the GIM and PV network in the more remote communities were 155.01 and 155.33 meters (see histograms in Figures C.5 and C.6).

In Borbòn, the mean contact density of an individual who was not ill in the PV network was higher compared to those who were ill (242.8 meters (SD 248.1m) vs. 189.5 meters (SD 245.0m)) but this difference was not statistically significant at the p=0.05 level (see Table C.1 and C.3). In the more remote communities, the mean contact density of an individual who was not ill in the PV network was *lower* compared to the those who were will (149.8 meters (SD 194.3.9m) vs. 98.3 meters (SD 93.0m) respectively) and this difference was significant at the p=0.05 level (see Table C.2 and C.4). The histogram of the ill and not ill group within the PV networks for Borbòn and the more remote communities illustrates the same (see Figure C.14 and C.16).

In Borbòn, the mean contact density of an individual who was not ill in the GIM network was higher compared to those who were ill (245.6 meters (SD 231.1m) vs. 173.3 meters (SD 240.7m) but this difference was not statistically significant at the p=0.05 level (see Table C.1 and C.3). In the more remote communities, the mean contact density of an individual who was not ill in the GIM network was lower compared to the those who were will ((138.4 meters (SD 190.9m) vs. 110.9 meters (SD 141.5m) respectively) but this difference was also not significant at the p=0.05 level (see Table C.2 and C.3). The histogram of the ill and not ill group within the GIM networks

for Borbòn and the more remote communities illustrates the same (see Figure C.13 and C.15).

Table C.1: Summary of contact density (meters) in General Important Matters (GIM) and People Visited networks in Borbon.

Network	Illness	Mean	Median	SD	Min	Max
GIM	No	173.3	74.0	240.7	1.0	2033.1
	Yes	245.6	177.6	231.1	1.0	687.9
PV	No	189.5	108.2	245.0	1.0	2033.1
	Yes	242.8	123.9	248.1	7.6	975.4

Table C.2: Summary of contact density (meters) in General Important Matters (GIM) and People Visited networks in five remote communities.

Network	Illness	Mean	Median	SD	Min	Max
GIM	No	138.4	78.8	190.9	1.0	2280.8
	Yes	110.9	64.7	141.5	1.0	687.2
PV	No	149.8	93.1	194.3	1.0	2312.9
	Yes	98.3	74.4	93.0	2.0	479.4

Table C.3: T-test of differences in contact density (meters) from contacts in ill versus not ill groups in General Important Matters (GIM) and People Visited networks in Borbon.

Network	Total Mean	No Ill Mean	Ill Mean	T-statistic	P-value
GIM	-72.37	173.27	245.64	-1.70	0.10
PV	-53.21	189.54	242.75	-1.20	0.24

Table C.4: T-test of differences in contact density (meters) from contacts in ill versus not ill groups in General Important Matters (GIM) and People Visited networks in five rural communities.

Network	Total Mean	No Ill Mean	Ill Mean	T-statistic	P-value
GIM	27.48	138.41	110.93	1.48	0.14
PV	51.48	149.79	98.31	4.46	0.00

Geo-referenced network structure of communities varies in commercial center versus adjacent communities Overall, the Kamada-Kawai visualization illustrated that the structure of social networks in Borbon compared to the rest of the communities are different. The social network

structure of Borbòn is characterized by unconnected groups of individual and less social density. The GIM network for the largest community, Borbòn, is illustrated in Figure C.30 of the Appendix. A large periphery of edges (connections between two individuals referred to as nodes) can be seen (creating a semi-circle around the central network). This periphery contains many dual and tri-node groups of edges. This pattern is expressed in the PV network as well but with larger groups of edges in the periphery (see Figure C.23). The cases of fever and/or diarrhea in the GIM network of Borbòn seem to be concentrated more in the center of the network (where individuals with many connections lie) compared to the PV network.

The Borbòn networks present quite differently from the adjacent communities of Timbire, Colon Eloy, Maldonado, Santo Domingo and Santa Maria (see Figure C.43 below for the PV network of Santa Maria and Appendix Figures C.26, C.30, C.34, C.38 and C.42 for the GIM networks and C.27, C.31, C.35 and C.39 for the PV networks) where there is more social density. In the adjacent communities, the pattern of peripheral edges (circle around the central large network block) is less prevalent, with nearly every node connected to the main central network.

The georeferenced visualizations illustrate the relationship between our exposure of interest, contact density, and the outcome, cases of fever and/or diarrhea. We identify that there is more dual clustering (socially and spatially) in the adjacent communities compared to Borbòn. In Borbòn, socially connected nodes who are ill (dark grey lines) are dispersed throughout the community (see Figure C.24 below for the GIM network). In adjacent communities, for example in Santa Maria (see Figure C.45 for the PV network below and Appendix Figure C.44 for the GIM network) the geo-reference GIM and PV networks are more concentrated in groups of dark lines connecting cases in closer proximity. The visualization of social network maps both spatially and aspatially allowed us to determine the need to stratify the models with Borbòn separated from the remaining adjacent communities to account for effect modification [168], this was confirmed by model results.

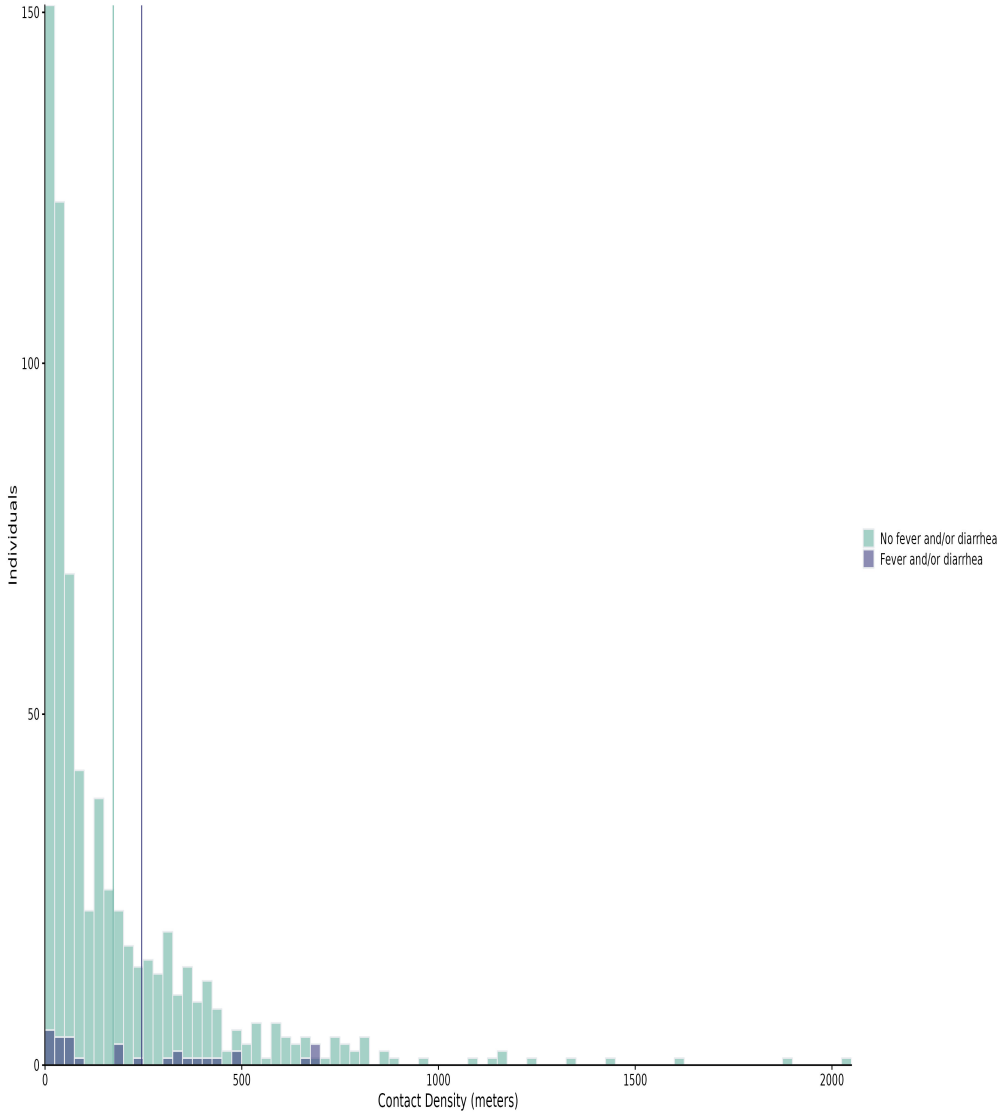

Figure C.13: Histogram of differences in harmonic mean distance (meters) from contacts in group with and without fever and/or diarrhea in the general important matters network in Borbon. The vertical line represents the mean within the group.

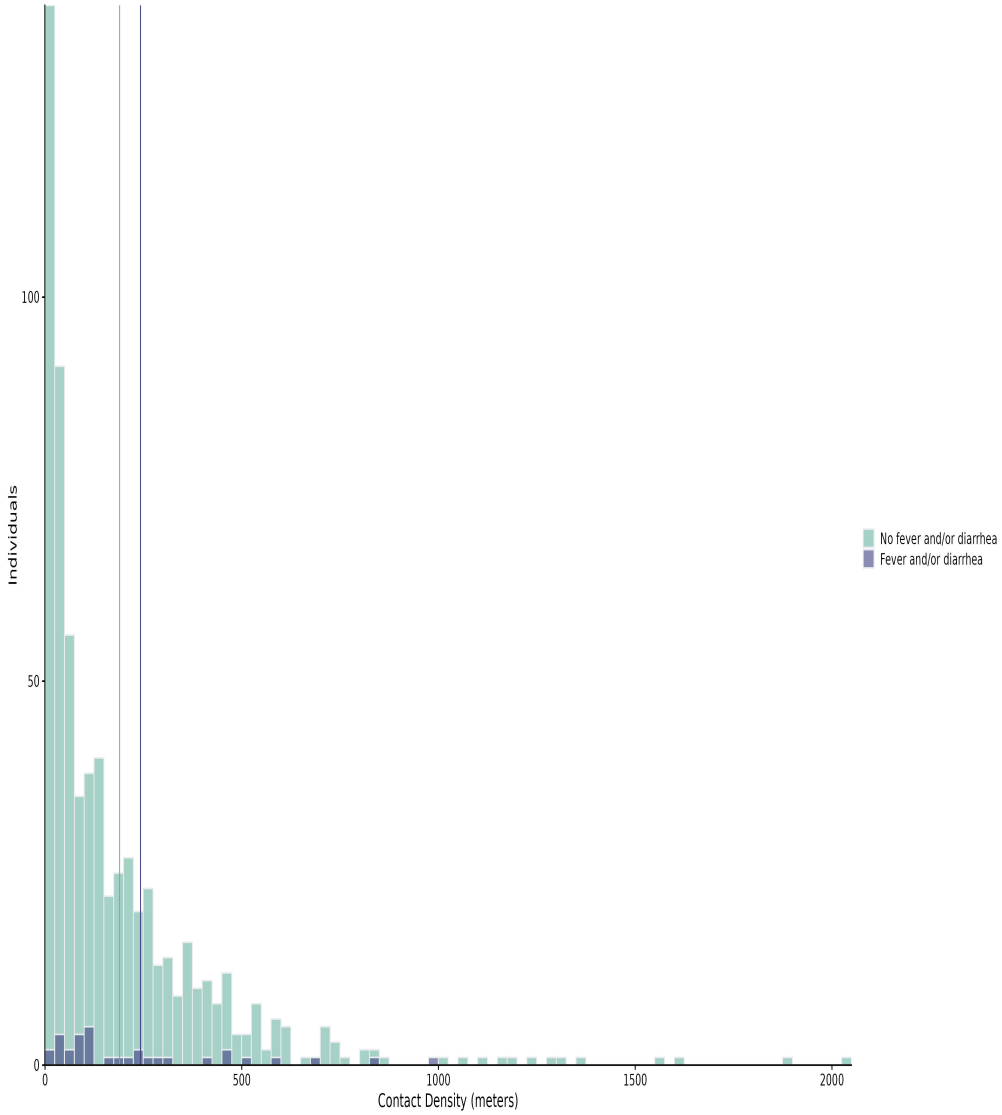

Figure C.14: Histogram of differences in harmonic mean distance (meters) from contacts in group with and without fever and/or diarrhea in the people visited network in Borbon. The vertical line represents the mean within the group.

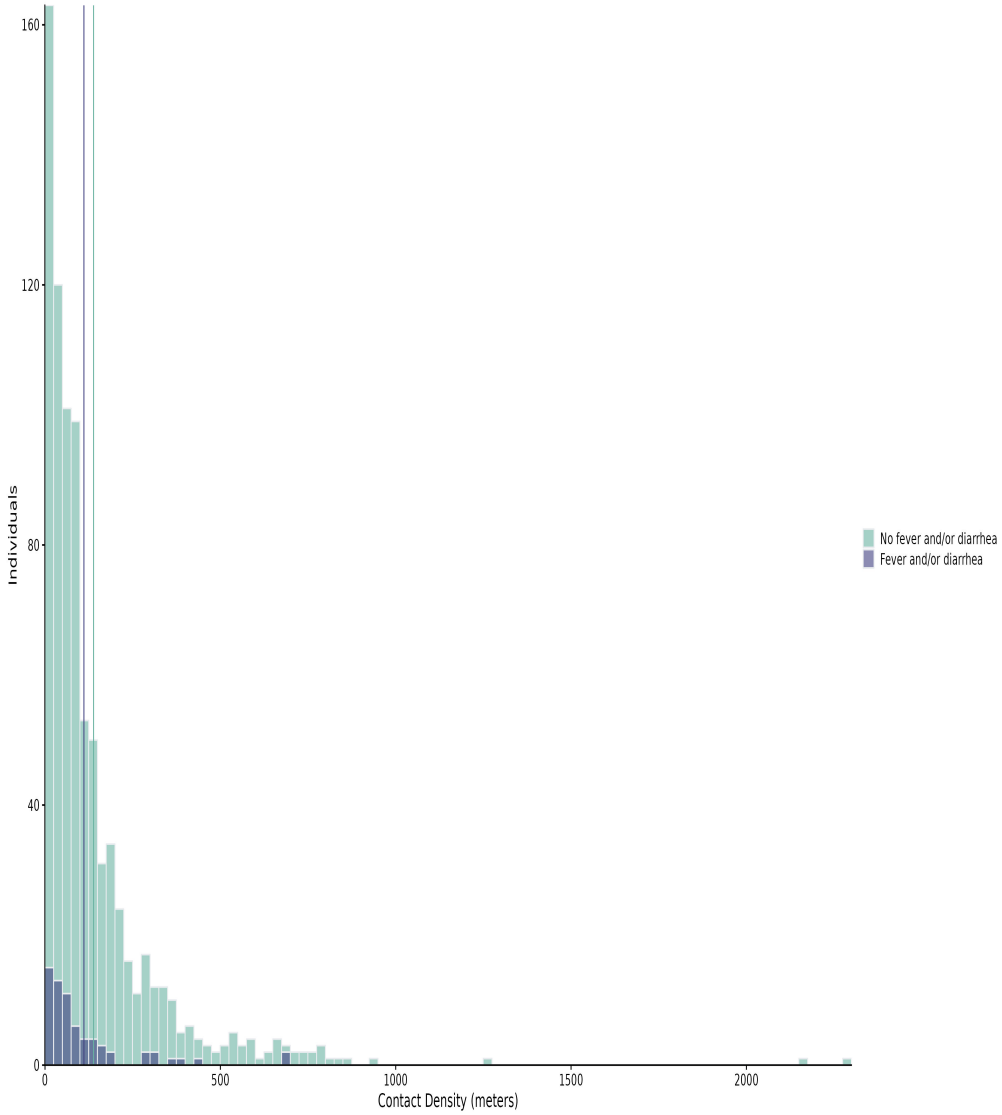

Figure C.15: Histogram of differences in harmonic mean distance (meters) from contacts in group with and without fever and/or diarrhea in the general important matters network in Colon Eloy, Timbire, Maldonado, Santo Domingo and Santa Maria. Timeline The vertical line represents the mean within the group.

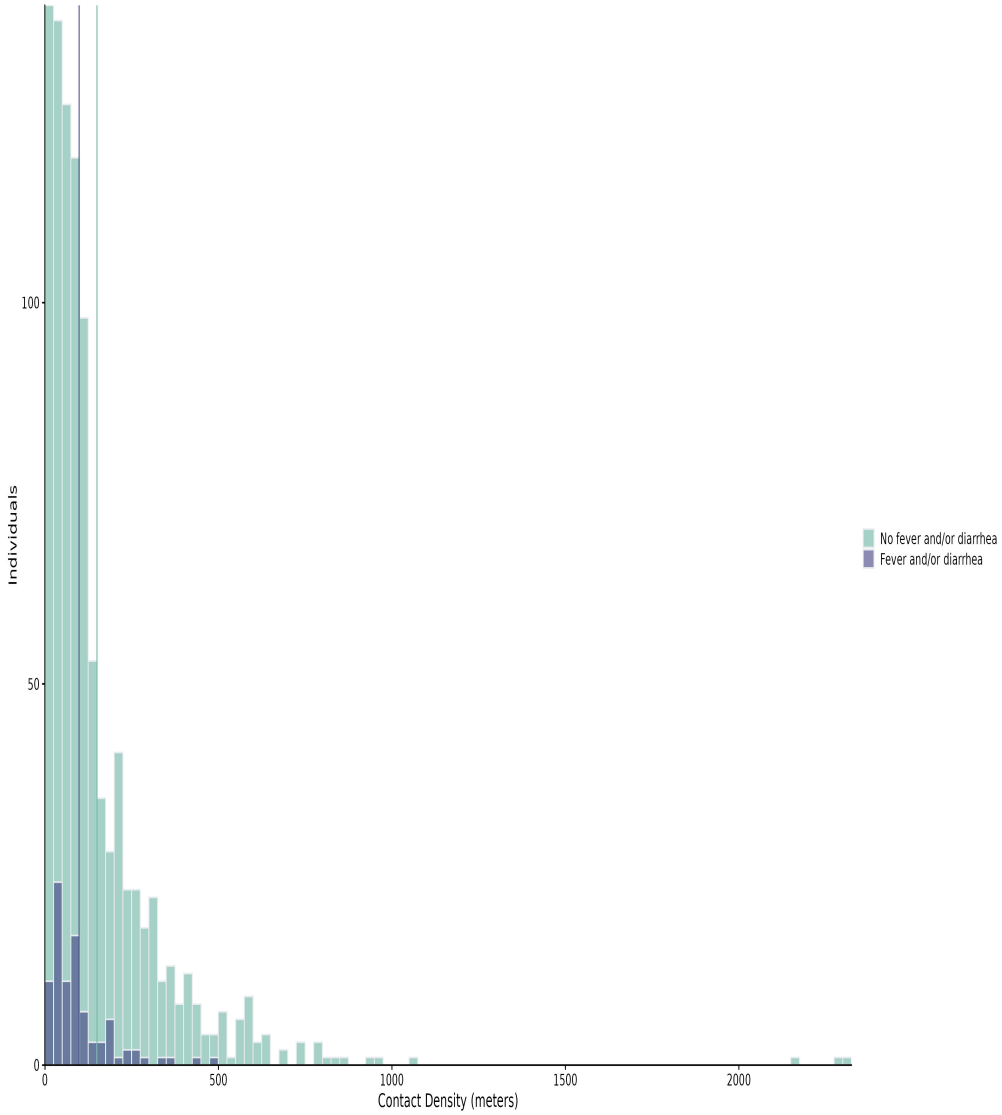

Figure C.16: Histogram of differences in harmonic mean distance (meters)from contacts in group with and without fever and/or diarrhea in the people visited network in Colon Eloy, Timbire, Maldonado, Santo Domingo and Santa Maria. The vertical line represents the mean within the group.

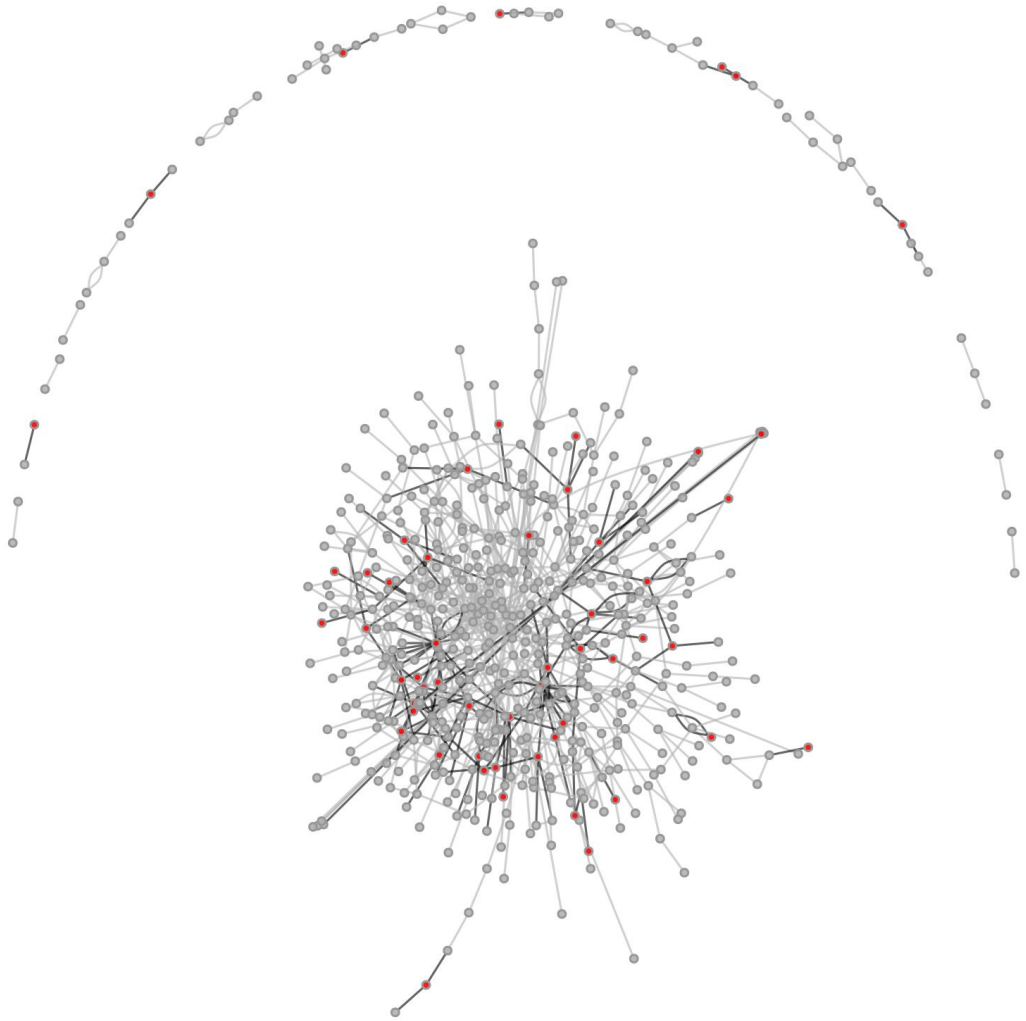

Figure C.17: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Borbòn. Mapping in social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

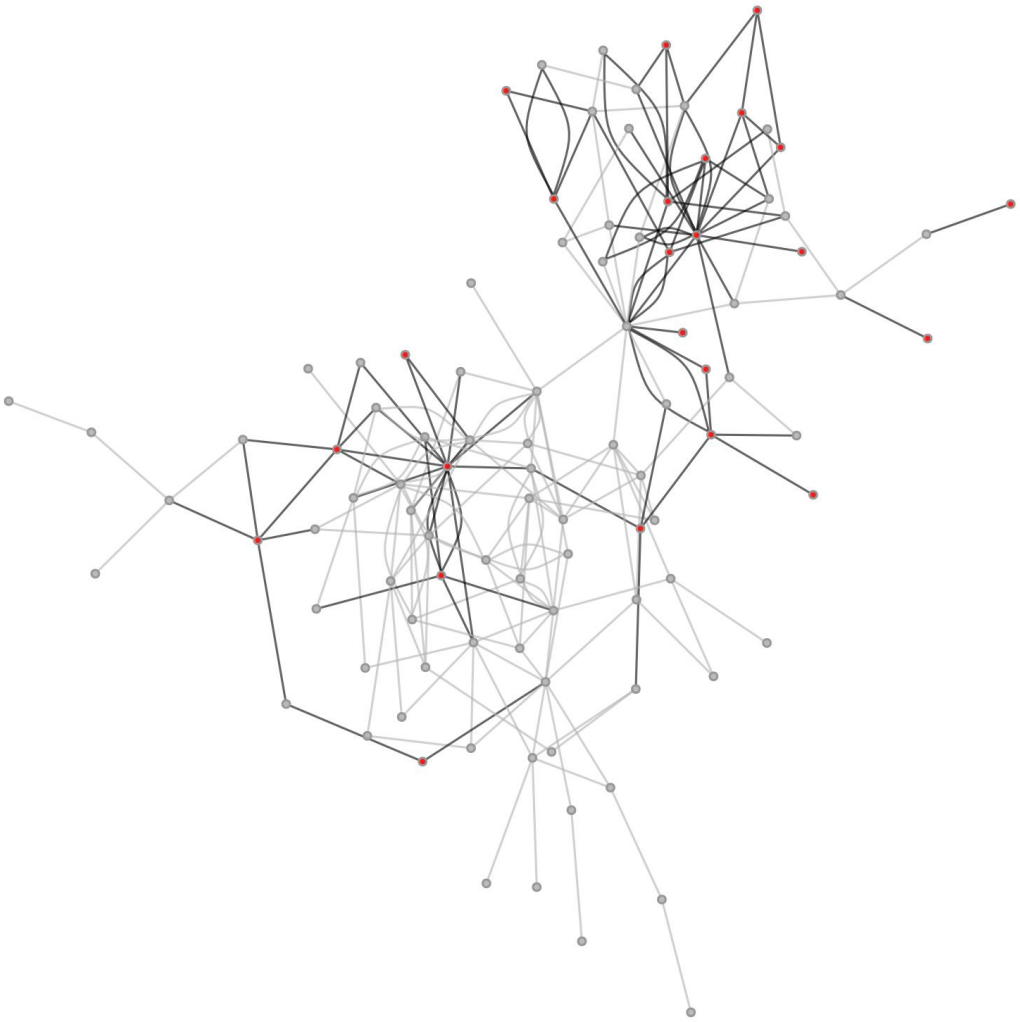

Figure C.18: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santa Maria. Mapping in social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

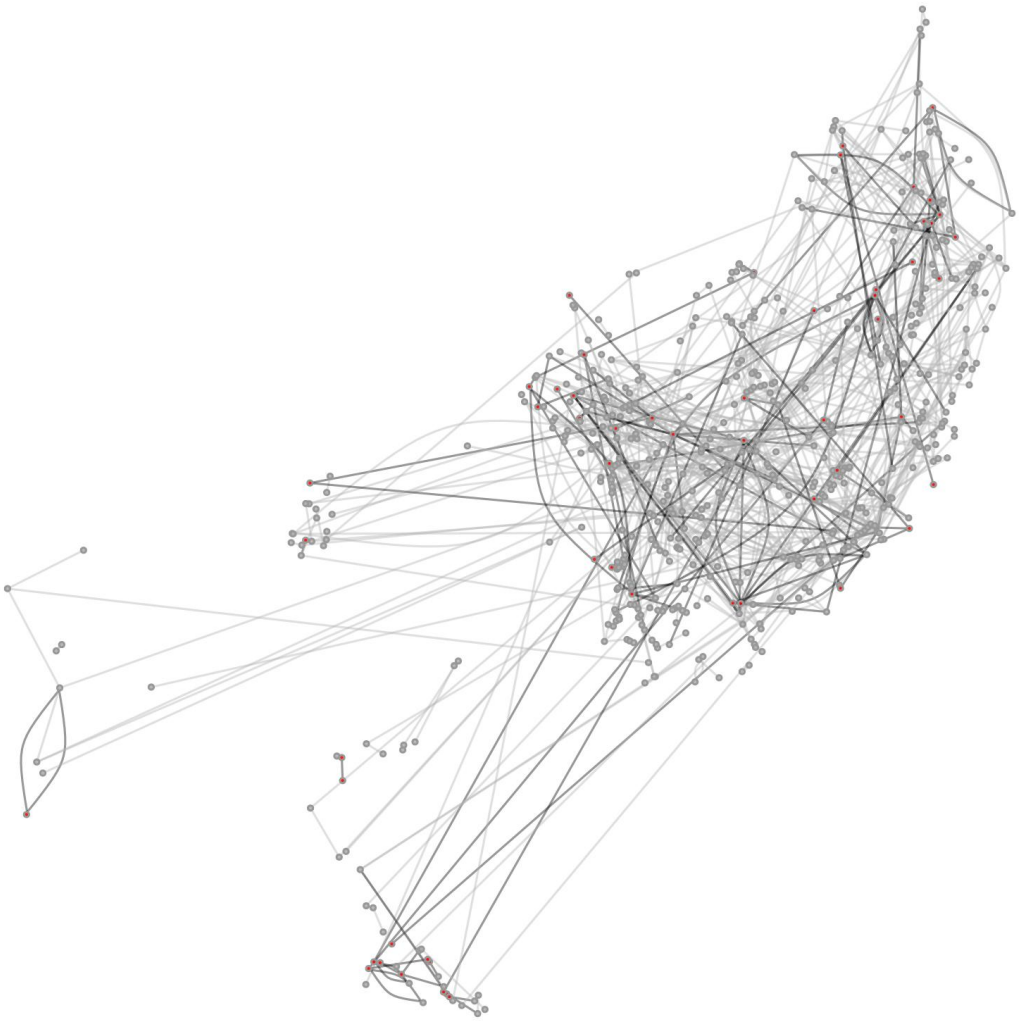

Figure C.19: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Borbòn. Mapping in social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

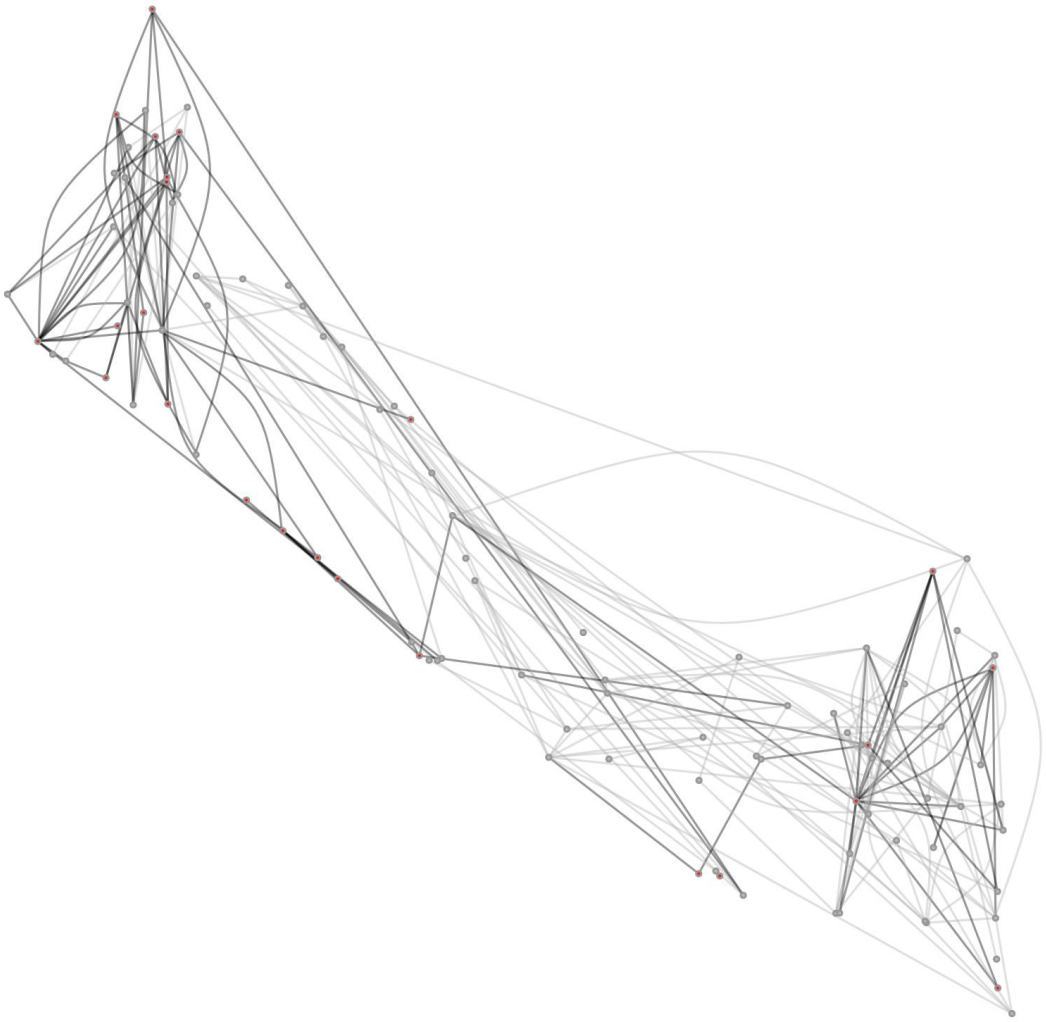

Figure C.20: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santa Maria. Mapping in GPS coordinates of homes and social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

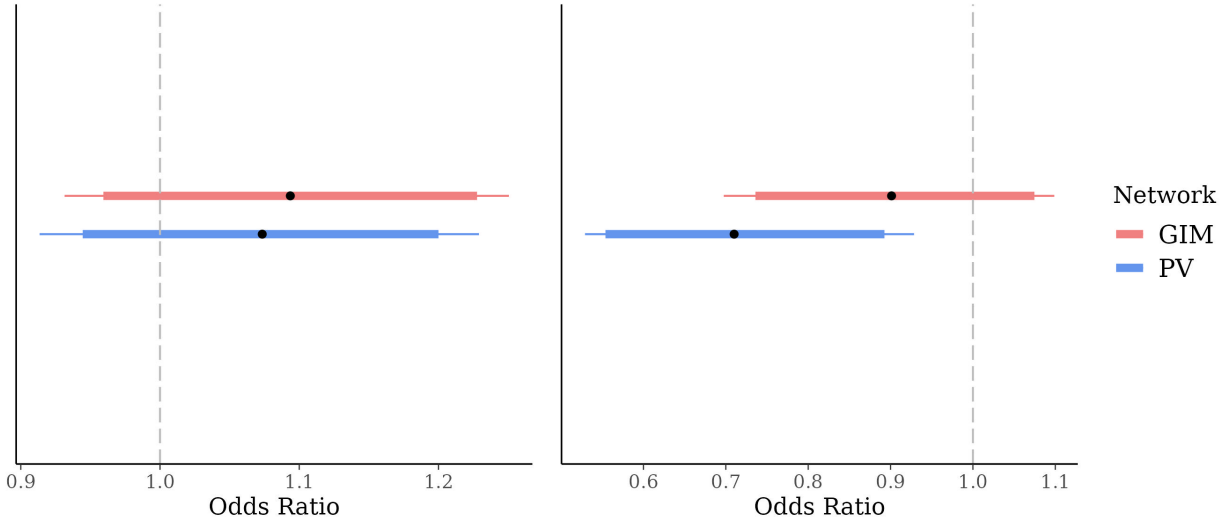

Figure C.21: (**left**) Borbón (**right**) Peripheral communities. Mean effect estimate and 90% (thick line) and 95% (thin line) credible interval of contact density (per 100m) in Bayesian (mixed-effects) binomial logit model of fever and/or diarrhea which includes confounders of exposure. Model for GIM network is in red and model for PV network is in blue. The dashed vertical line indicates an odds ratio of 1.0 (null).

Contact density plays a role in disease risk and does so differently in commercial center vs. peripheral communities Similar to our descriptive findings, our model analyses suggests that contact density is a significant risk factor for illness after accounting for known confounders in the PV network within peripheral communities (see Figure C.21 right and Table C.6). We report the results of our model as the mean of parameter distributions along with the 90% and 95% posterior credible intervals. In the peripheral communities the relationship was negative in both the GIM and PV networks, while in the more urban area of Borbòn, the relationship was positive (See Figure C.21 left and Table C.5). Therefore, having contacts who are further away (on average) in the peripheral communities, decreases risk while in Borbòn having contacts further away (on average) increases the risk of disease.

C.2.3 Contact Density Discussion

Contact density is modified by factors related to community structure The geographic location of one's social contacts becomes increasing relevant in particular contexts, for example in remote communities where shared environmental risk factors (like well water use, shared latrines and small scale husbandry) continue to be a strong driver of infectious disease outcomes. Diseases that are spread from person-to-person and via environmental intermediaries (a prime example being enteric diseases) induce a shared geographic risk space between contacts. Further within(LMICs,

Table C.5: Bayesian binomial logit model of fever and/or diarrhea of contact density with confounders in Borbon.

Network	Parameter	Mean	Lower 95%	Upper 95%
GIM	(Intercept)	0.00	0.00	0.06
	Contact Density (100m)	1.09	0.93	1.25
	Afro-Ecuadorian Race	2.07	0.76	7.10
	Chachi Race	0.00	0.00	6.37
	Male	0.60	0.24	1.37
	Literacy	3.43	0.48	62.61
	Age	1.01	0.99	1.03
PV	(Intercept)	0.01	0.00	0.16
	Contact Density (100m)	1.07	0.91	1.23
	Afro-Ecuadorian Race	1.40	0.55	4.29
	Chachi Race	0.00	0.00	4.23
	Male	0.76	0.32	1.74
	Literacy	1.87	0.24	35.23
	Age	1.01	0.99	1.04

especially within settings with lower population densities, distance remains a relevant measure of changing environmental context. We propose a metric we denote as contact density, which is the harmonic mean of the spatial distance of an individual to all their social network contacts. This measure allows us to capture the joint social and spatial nature of infection risk, i.e. that both proximity and social interaction are important contributors to exposure and infection risk. This approach allows us to ask whether the geographic distribution of an individual's social contacts impacts their specific risk. To what extent are these spatial relationships reflective of person-to-person contact rather than exposures in the shared environment? We explore integrated social and spatial case networks in a location with endemic diarrheal disease. We generate maps of enteric disease cases that compare Euclidean versus social distance. These maps allow us to visualize risk differences between networks where social contacts are geographically dense versus those where social contacts are geographically dispersed. We then generate an integrated social and spatial connectivity metric we call contact density (the harmonic mean distance of one's contacts) and evaluate its influence on enteric disease incidence. Does the physical proximity of your contacts matter in terms of disease risk? We hypothesize that spatio-social connectivity (being physically close to those that you socially interact with the most) influences enteric disease incidence at the individual level, even after accounting for measures (similar to previous studies) of physical and social connectivity independently along with demographic confounders. Our analysis illustrates that given the variable geo-social network structures of remote compared to more densely population communities, contact density is risk inducing within networks in remote areas and protective

Table C.6: Bayesian mixed-effects binomial logit model of fever and/or diarrhea of contact density with confounders in peripheral communities.

Network	Parameter	Mean	Lower 95%	Upper 95%
GIM	(Intercept)	0.08	0.01	0.45
	Contact Density (100m)	0.90	0.70	1.10
	Afro-Ecuadorian Race	1.05	0.32	4.53
	Chachi Race	3.40	0.73	20.61
	Male	0.91	0.50	1.61
	Literacy	0.80	0.34	2.05
	Age	1.00	0.98	1.02
	Intercept – Colon Eloy	1.48	0.72	3.81
	Intercept – Timbire	0.70	0.24	1.53
	Intercept – Maldonado	0.68	0.23	1.48
	Intercept – Santo Domingo	1.48	0.70	3.86
	Intercept – Santa Maria	0.75	0.22	1.75
	Sigma	1.72	1.01	12.44
PV	(Intercept)	0.11	0.02	0.52
	Contact Density (100m)	0.71	0.53	0.93
	Afro-Ecuadorian Race	0.46	0.20	1.09
	Chachi Race	1.47	0.42	5.75
	Male	0.91	0.54	1.54
	Literacy	2.38	0.84	8.18
	Age	1.00	0.98	1.01
	Intercept – Colon Eloy	1.45	0.81	3.18
	Intercept – Timbire	0.78	0.33	1.42
	Intercept – Maldonado	0.85	0.39	1.61
	Intercept – Santo Domingo	1.22	0.65	2.70
	Intercept – Santa Maria	0.77	0.27	1.49
	Sigma	1.43	1.00	5.24

in more densely populated areas. In remote areas as contact density (average distance of contacts) decreases, enteric disease incidence decreases. We postulate that in remote communities contact density highlights an underlying process of disease clustering in geo-social space. In other words, living in a densely populated part of the community and having many social contacts are both risk inducing but experiencing these both together (by having one's social contacts within close proximity) is significantly more risk inducing than each risk factor alone in peripheral communities. We hypothesize that the direction of contact density's effect is negative in peripheral communities because environmental factors relation to share latrine use, local husbandry and shared well water may play more a role in these communities compared to commercial centers like Borbon. Therefore, living in densely populated areas within the community and interacting with neighbors regularly (by visiting them) in those same areas increases one's risk more so than one of the aforementioned risk factors. The direction of this relationship is observed in the contact (PV and core-discussion network alike GIM, reinforcing the findings.

C.2.4 Contact Density: Additional social network maps across two types of networks

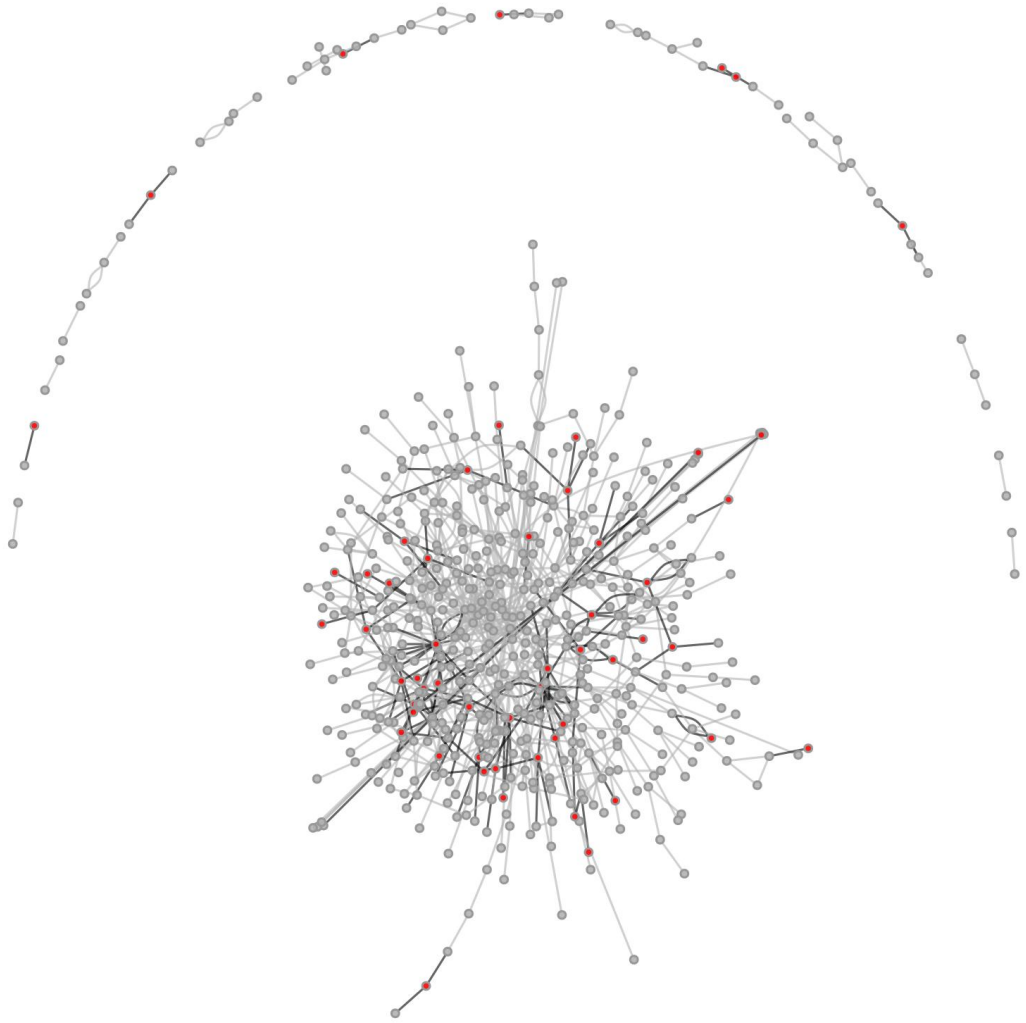

Figure C.22: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Borbon. Mapping in social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

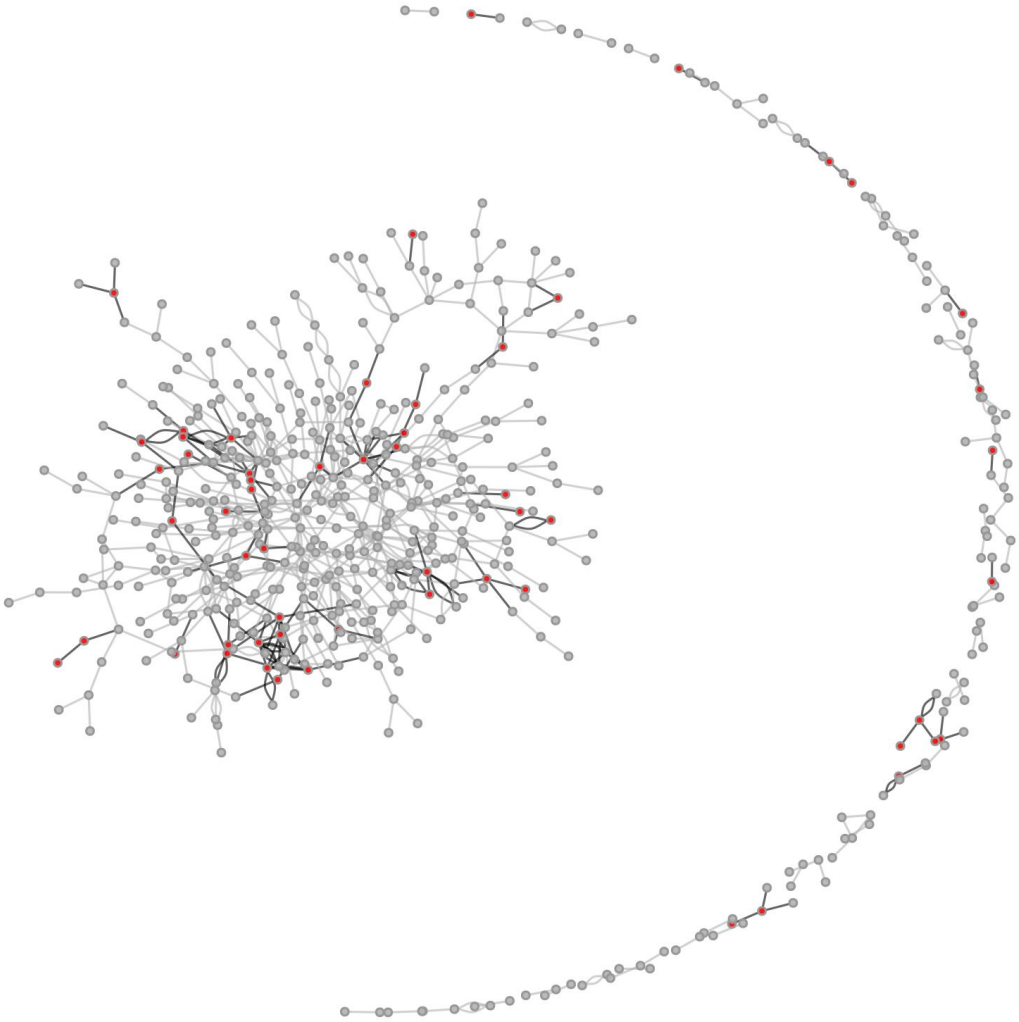

Figure C.23: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Borbon. Mapping in social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

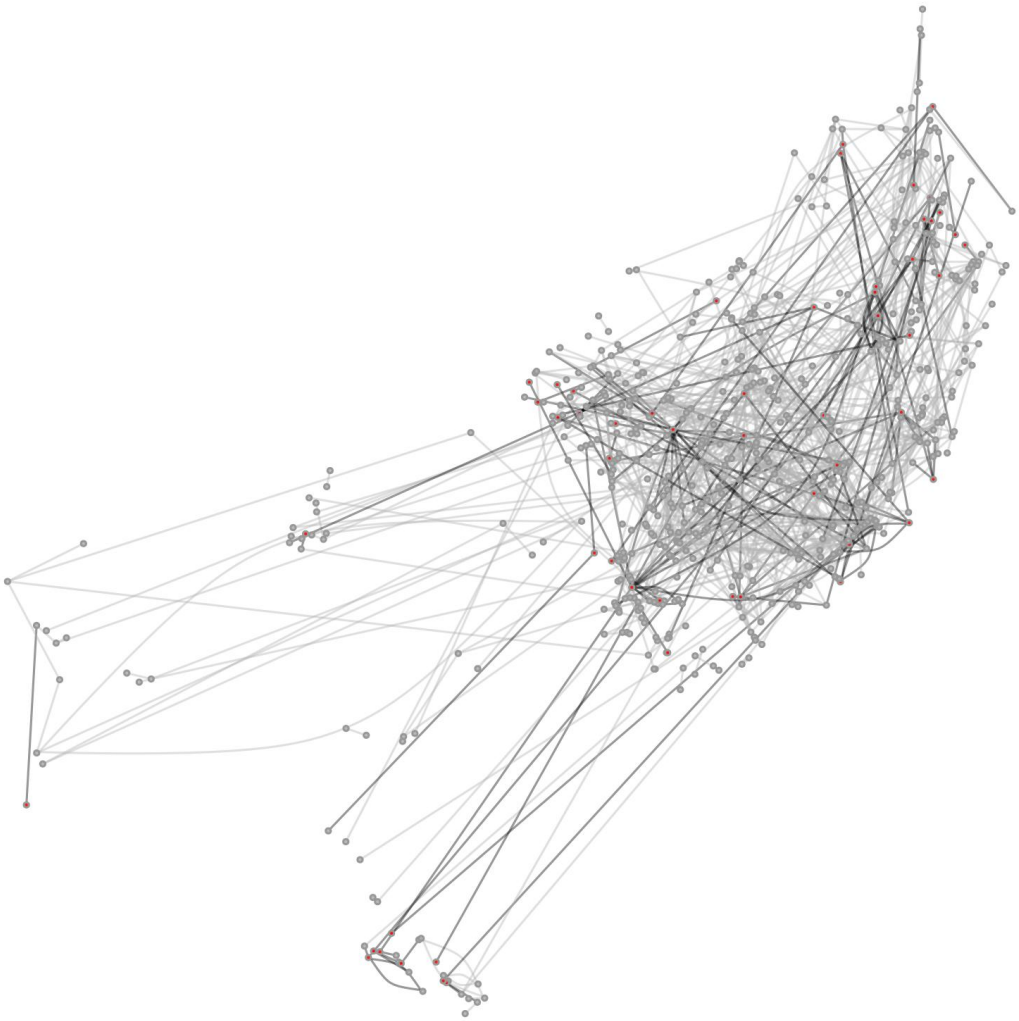

Figure C.24: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Borbon. Mapping in GPS coordinates of homes and social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

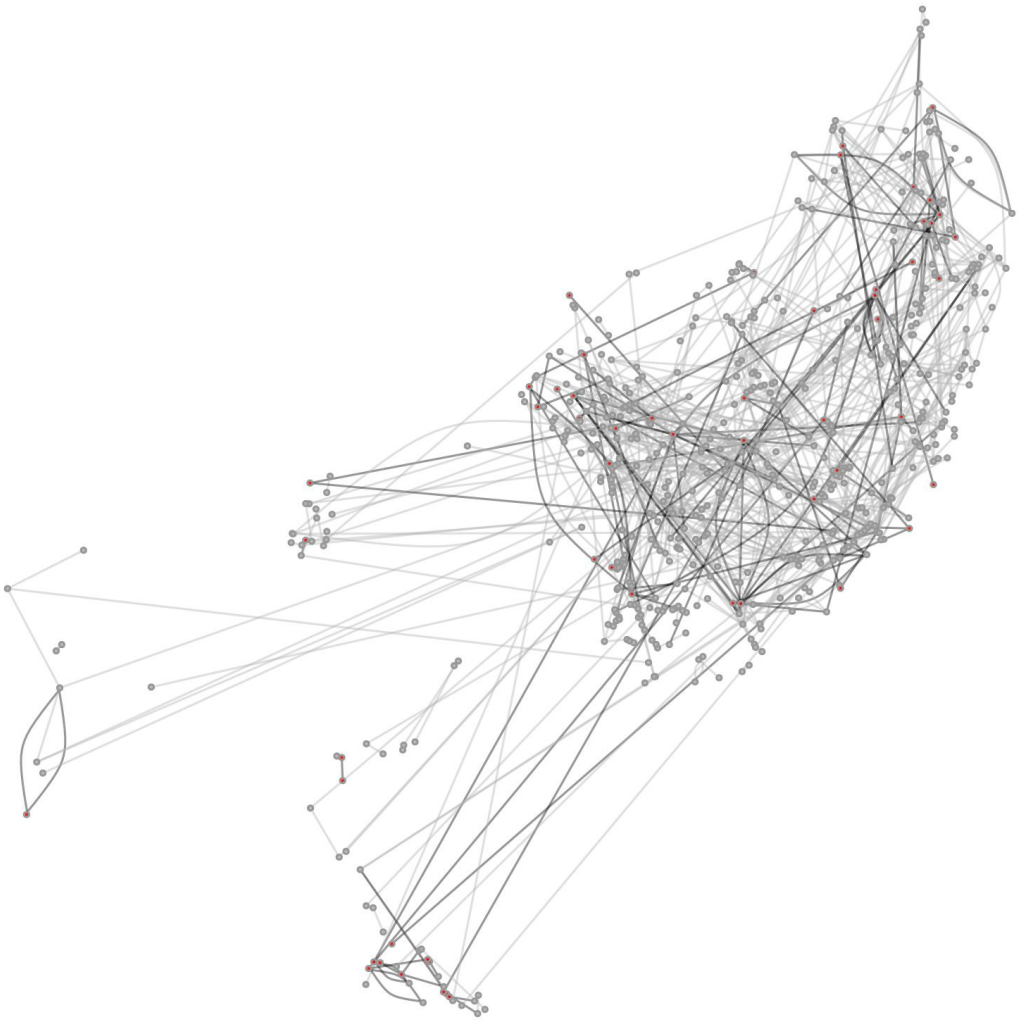

Figure C.25: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Borbon. Mapping in GPS coordinates of homes and social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

Figure C.26: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Timbire. Mapping in social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

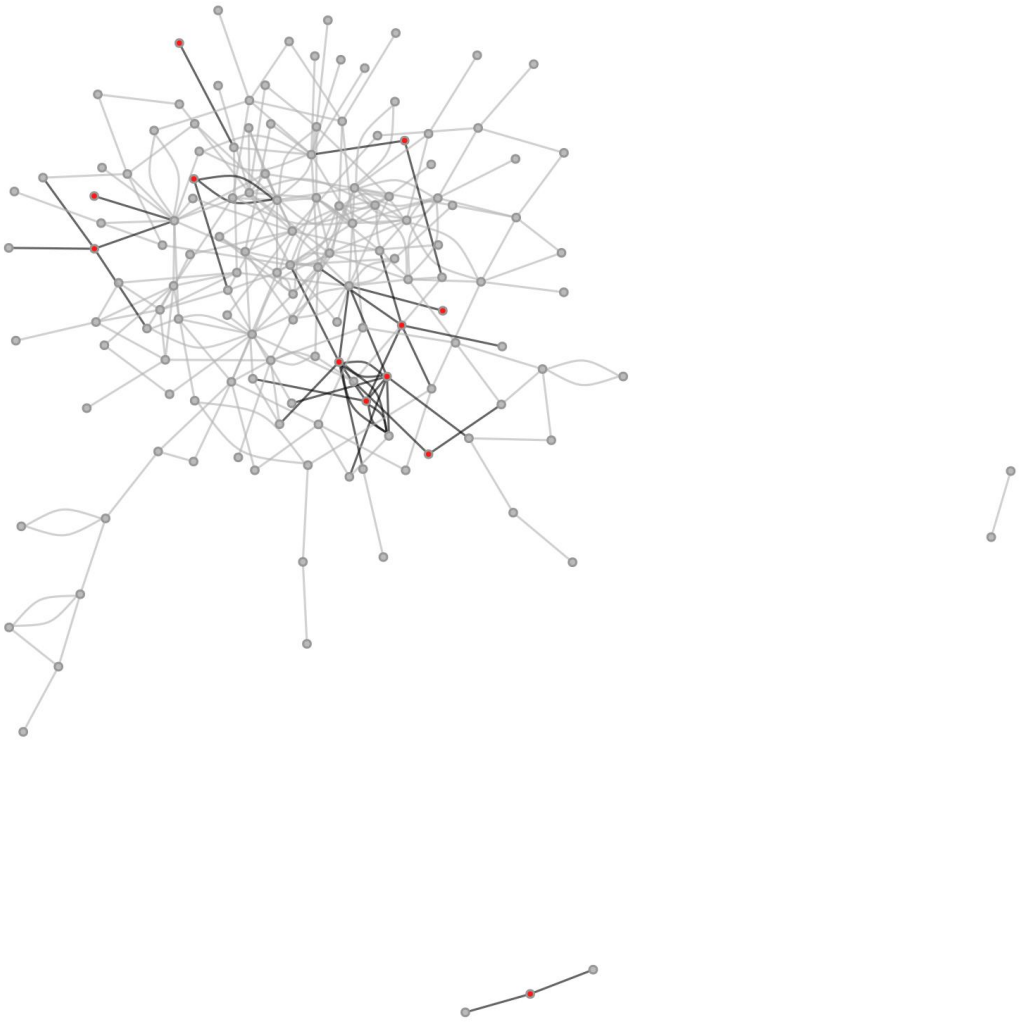

Figure C.27: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Timbire. Mapping in social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

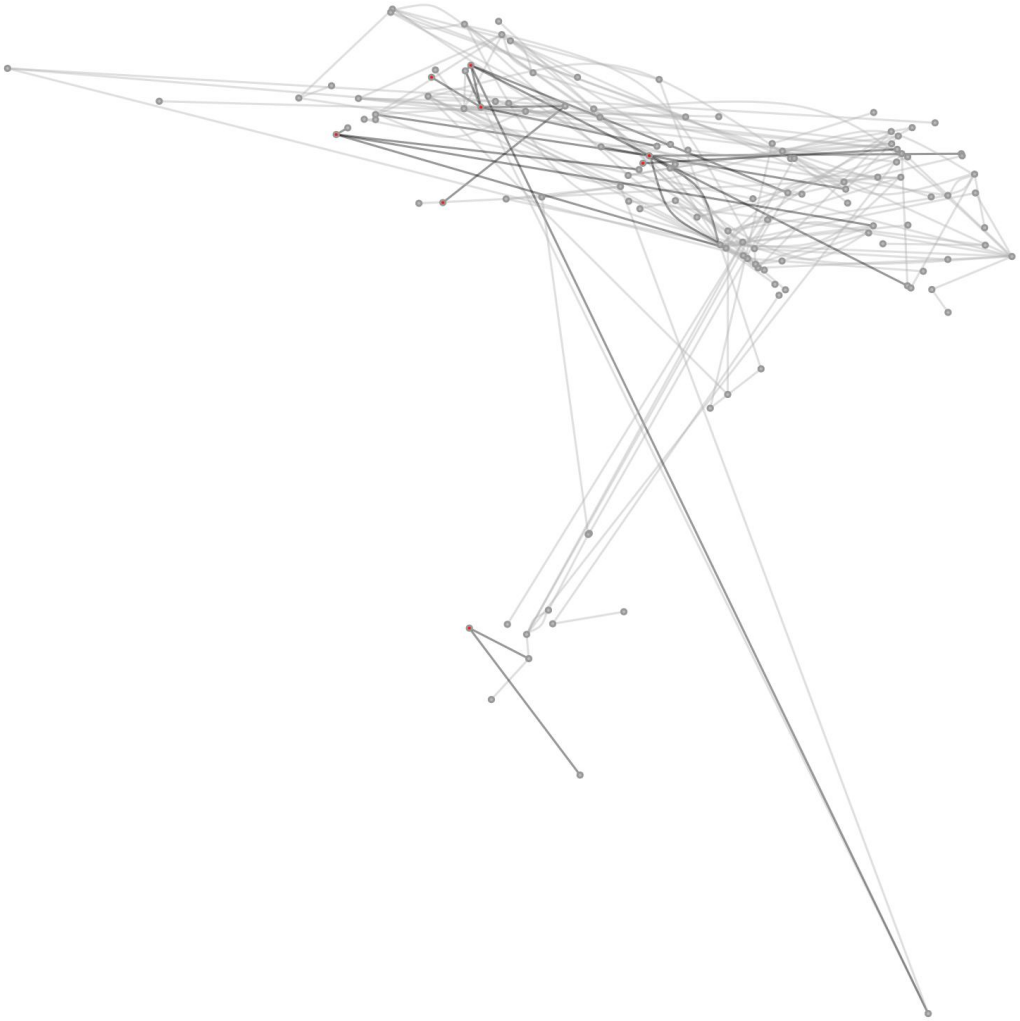

Figure C.28: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Timbire. Mapping in GPS coordinates of homes and social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

Figure C.29: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Timbire. Mapping in GPS coordinates of homes and social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

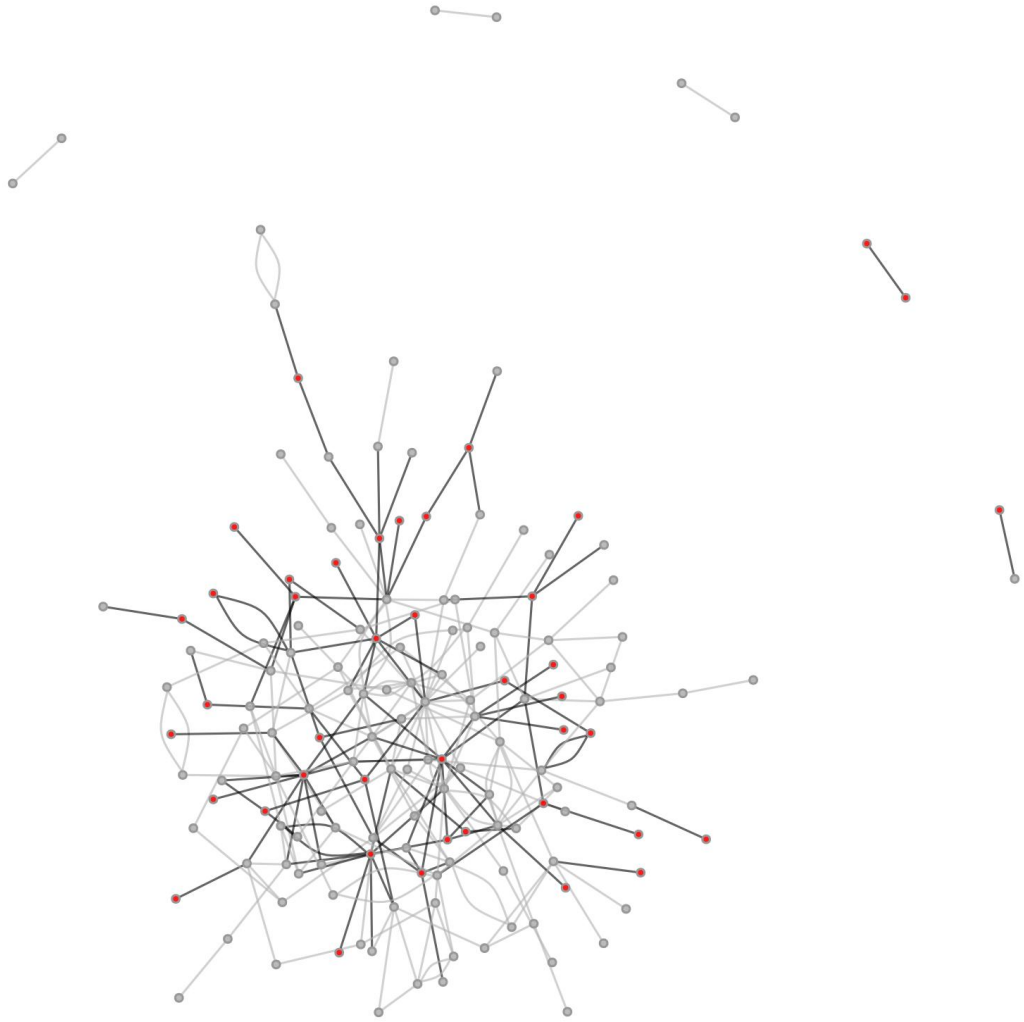

Figure C.30: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Colon Eloy. Mapping in social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

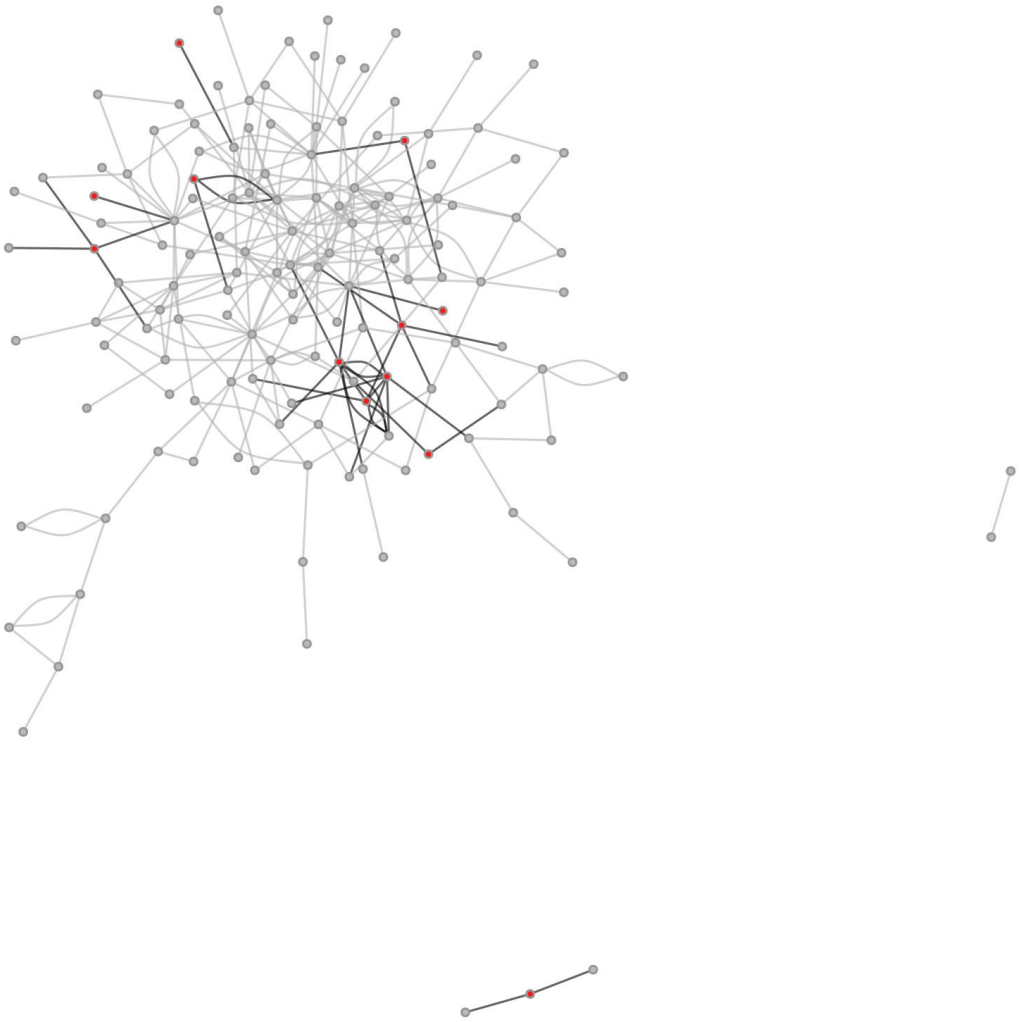

Figure C.31: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Timbire. Mapping in social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

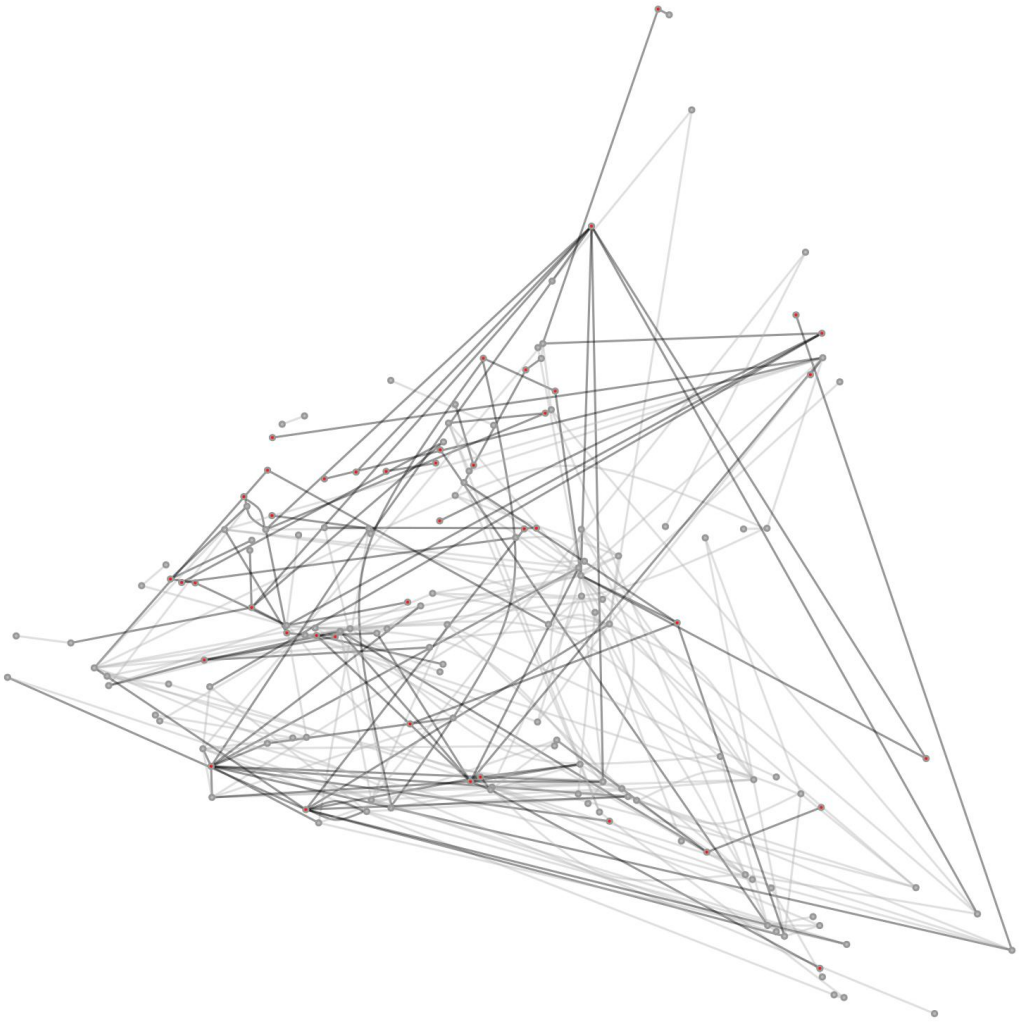

Figure C.32: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Colon Eloy. Mapping in GPS coordinates of homes and social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

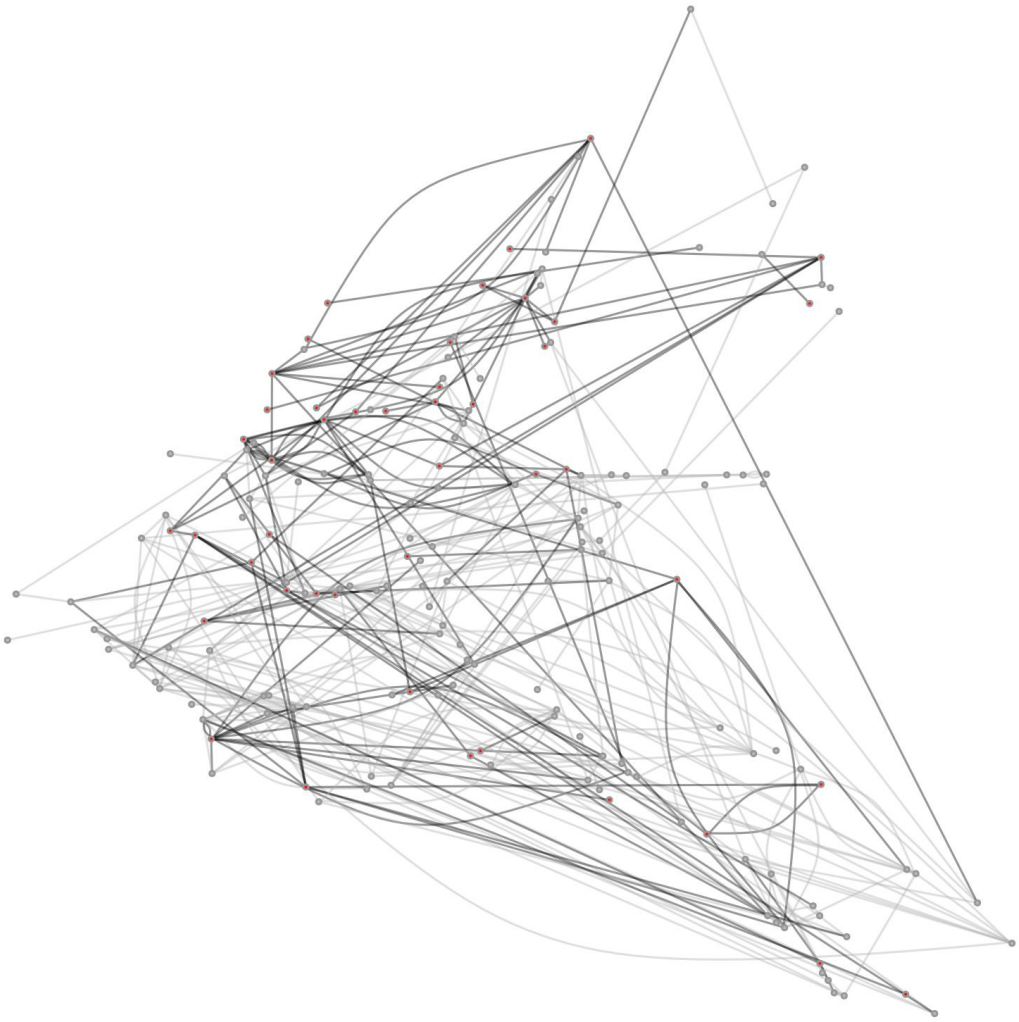

Figure C.33: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Colon Eloy. Mapping in GPS coordinates of homes and social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

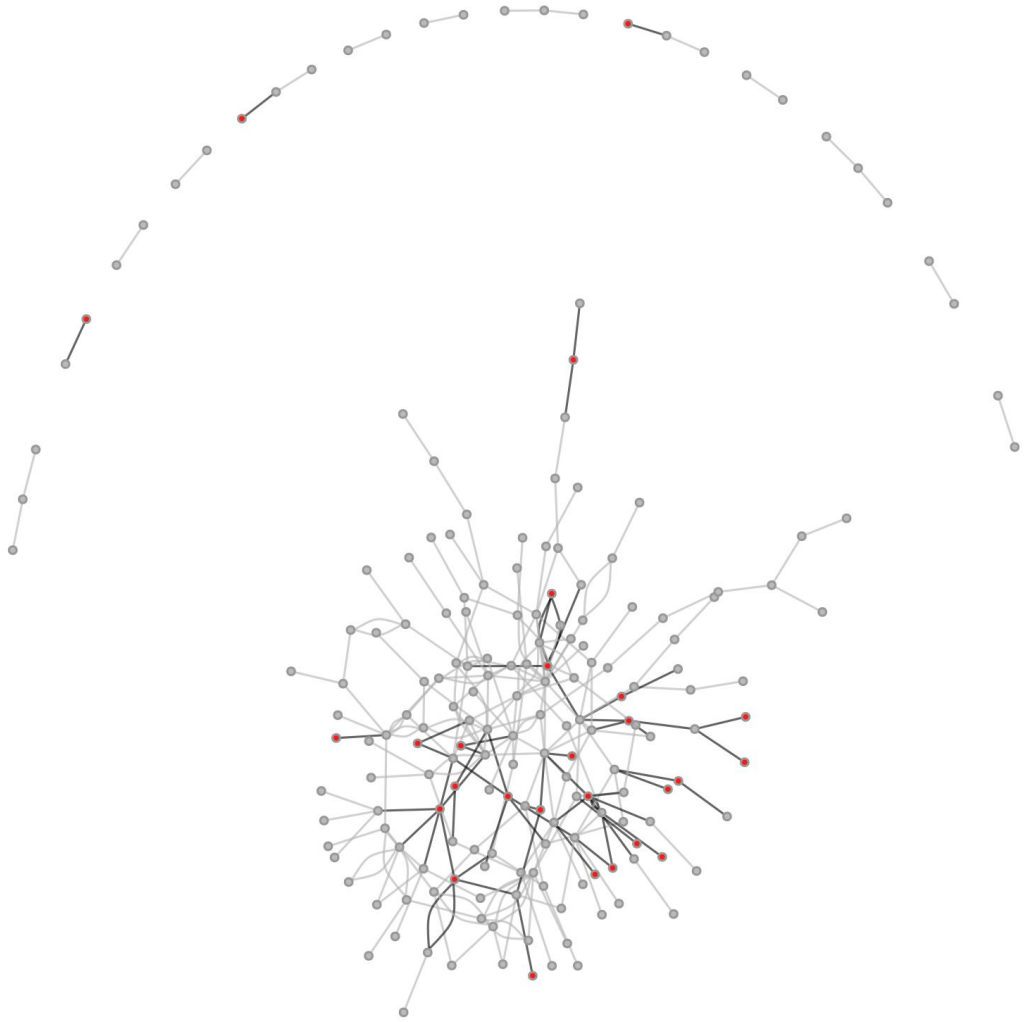

Figure C.34: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Maldonado. Mapping in social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

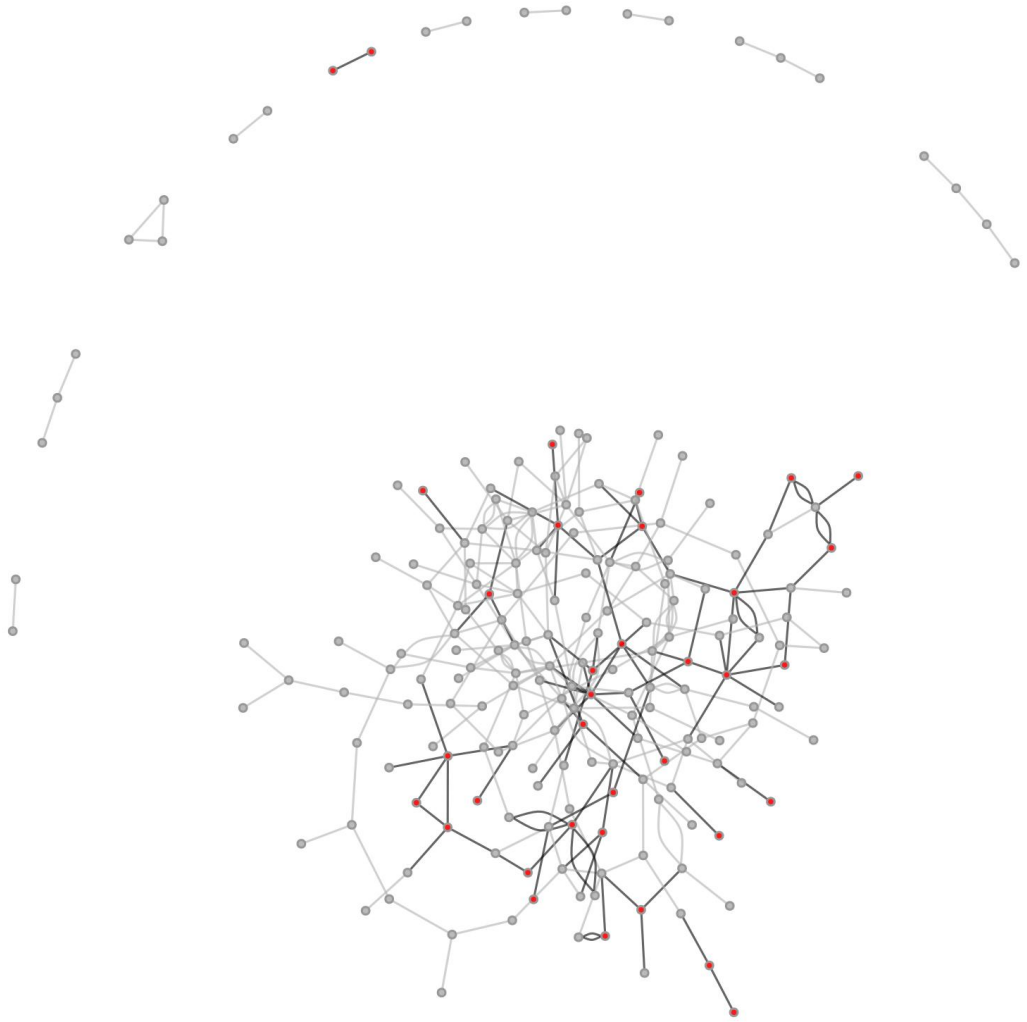

Figure C.35: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Maldonado. Mapping in social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

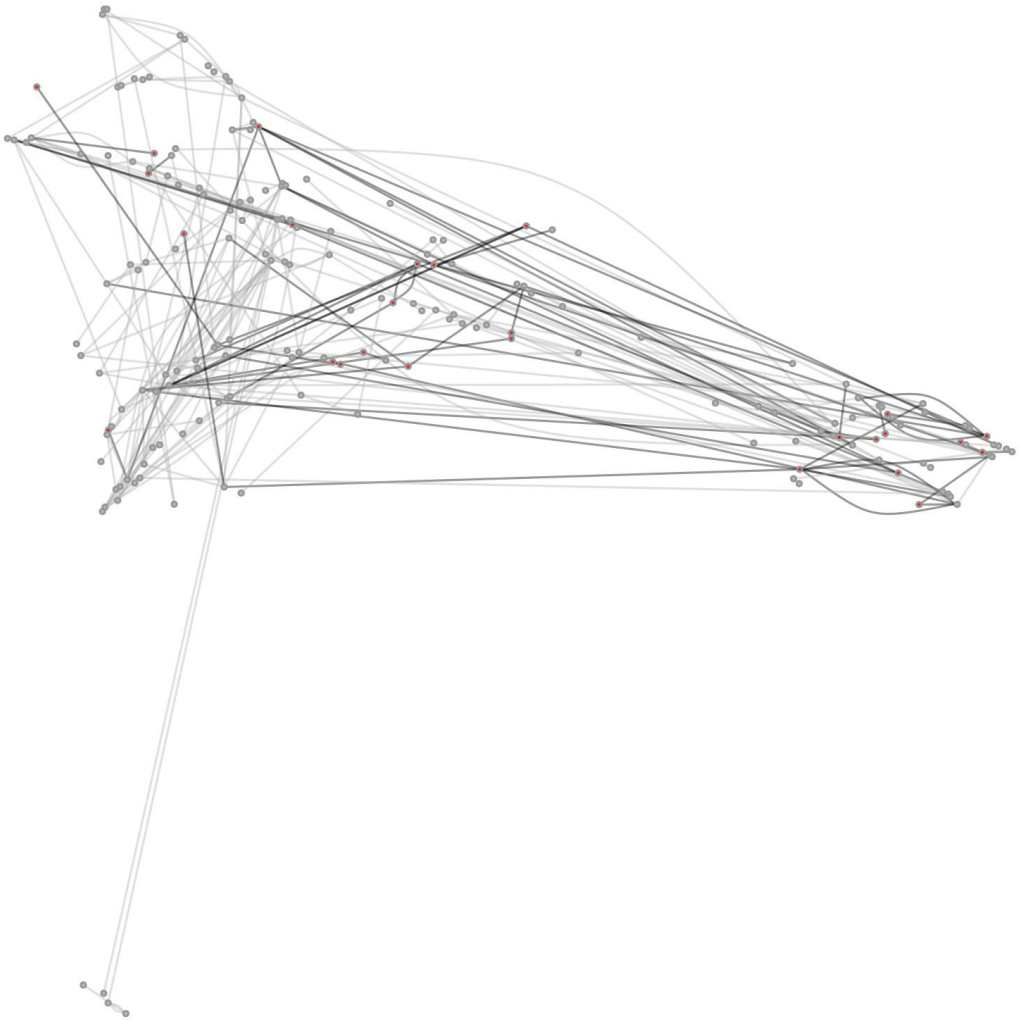

Figure C.36: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Maldonado. Mapping in GPS coordinates of homes and social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

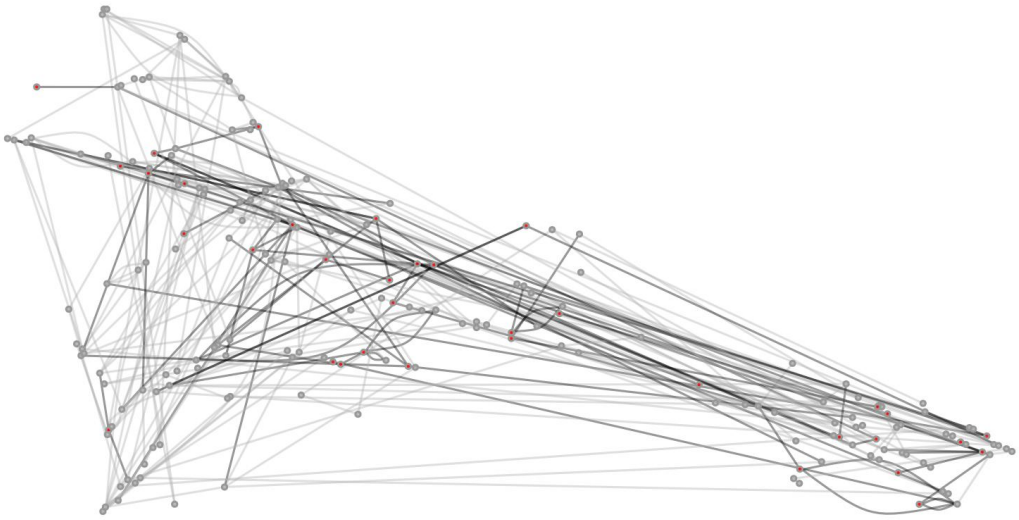

Figure C.37: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Maldonado. Mapping in GPS coordinates of homes and social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

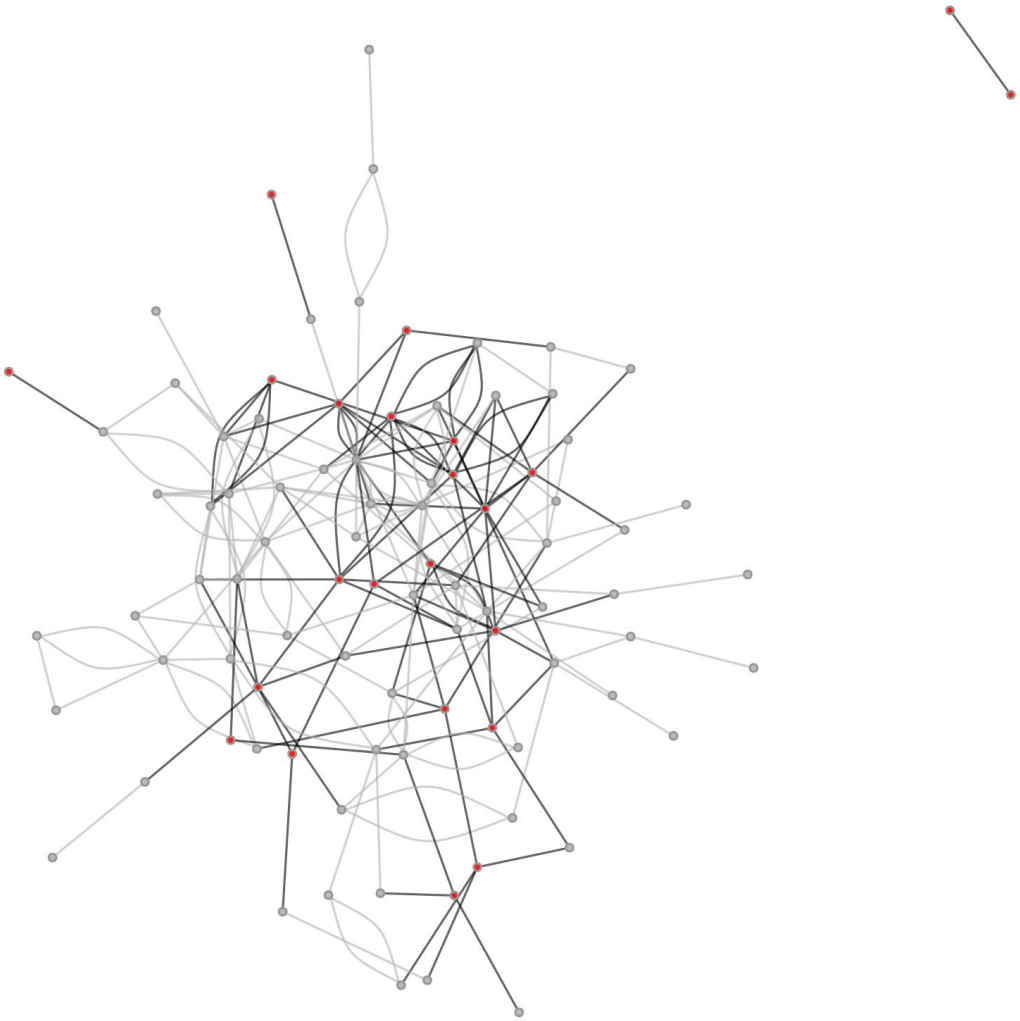

Figure C.38: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santo Domingo. Mapping in social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

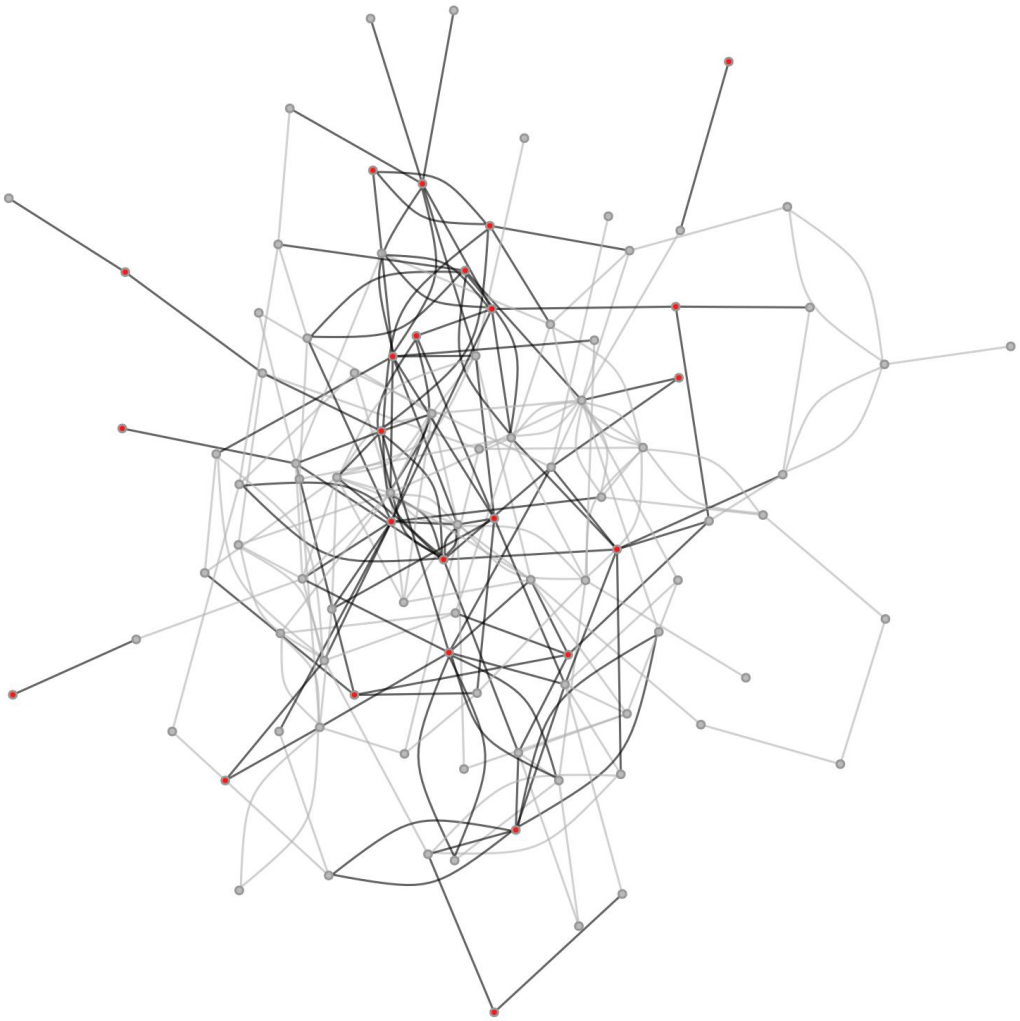

Figure C.39: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santo Domingo. Mapping in social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

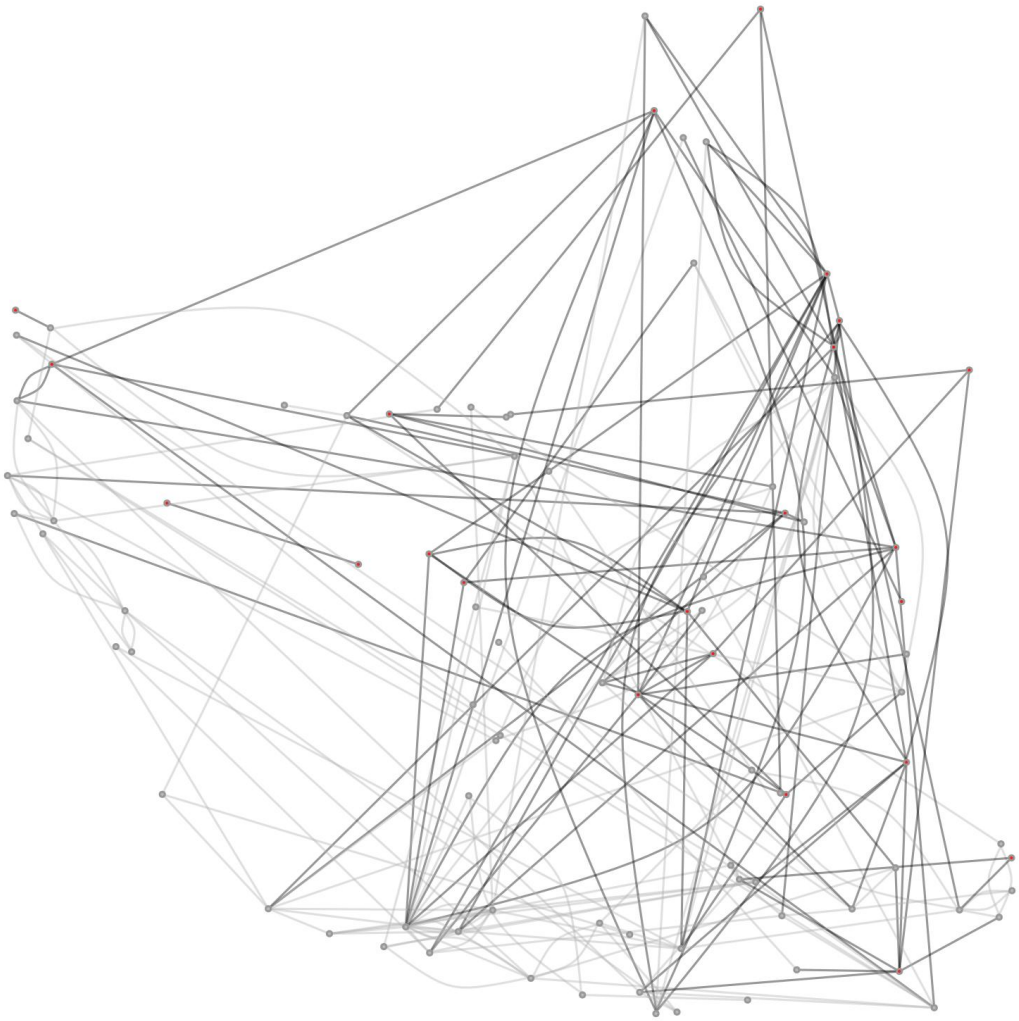

Figure C.40: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santo Domino. Mapping in GPS coordinates of homes and social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

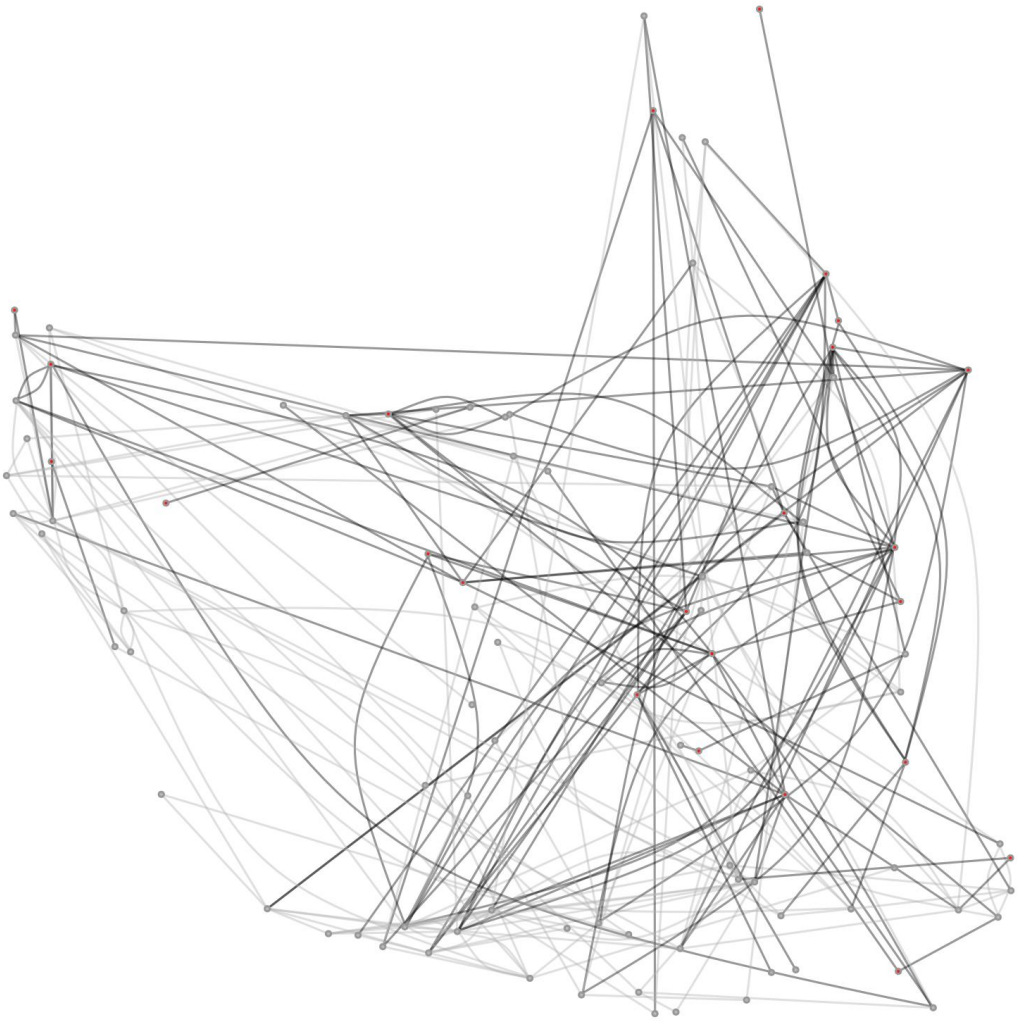

Figure C.41: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santo Domino. Mapping in GPS coordinates of homes and social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

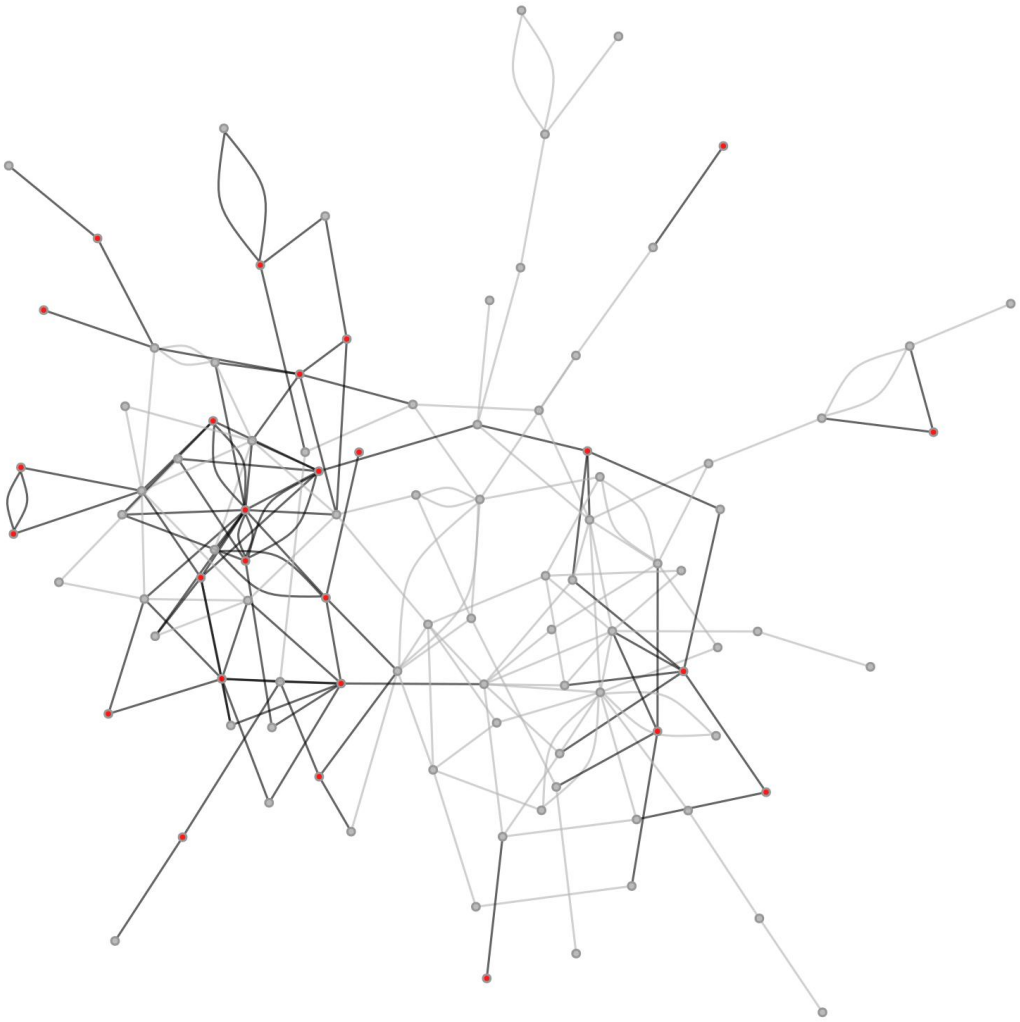

Figure C.42: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santa Maria. Mapping in social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

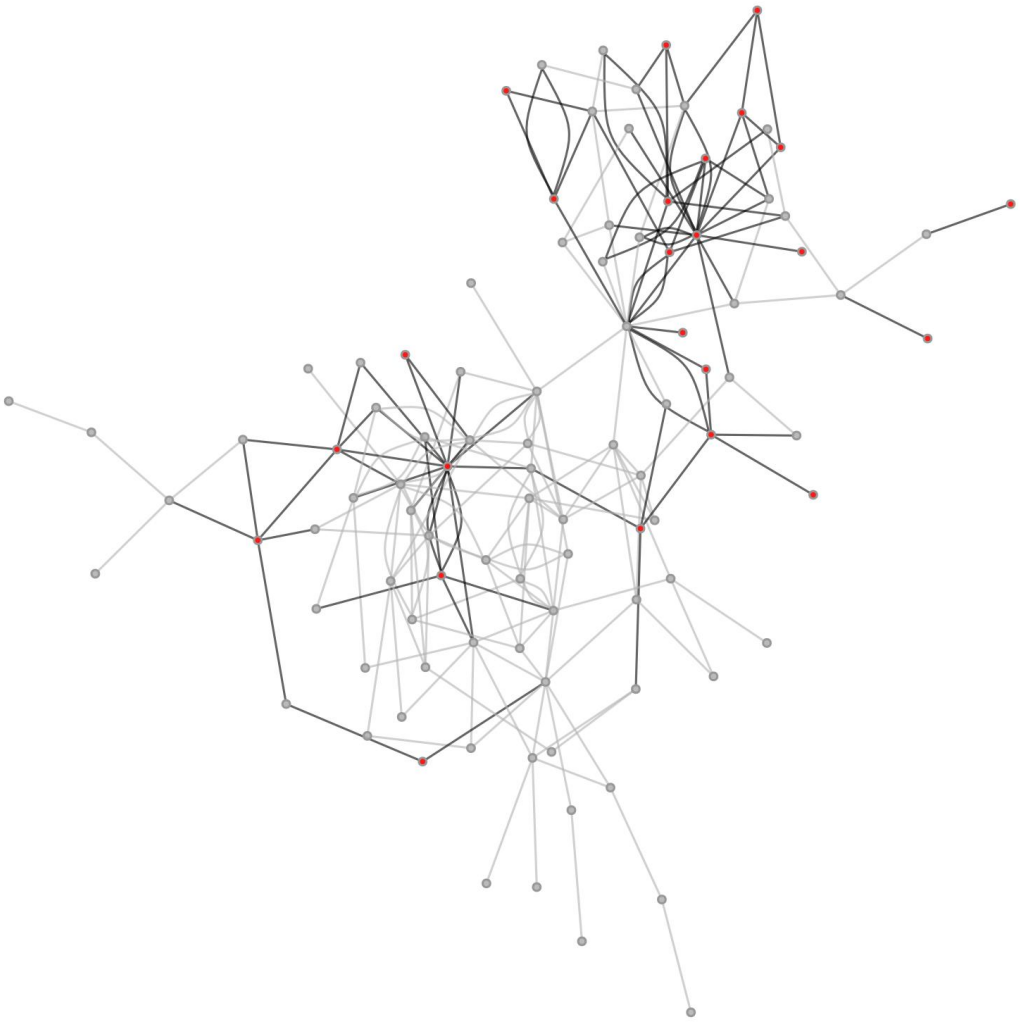

Figure C.43: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santa Maria. Mapping in social network (people visted). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

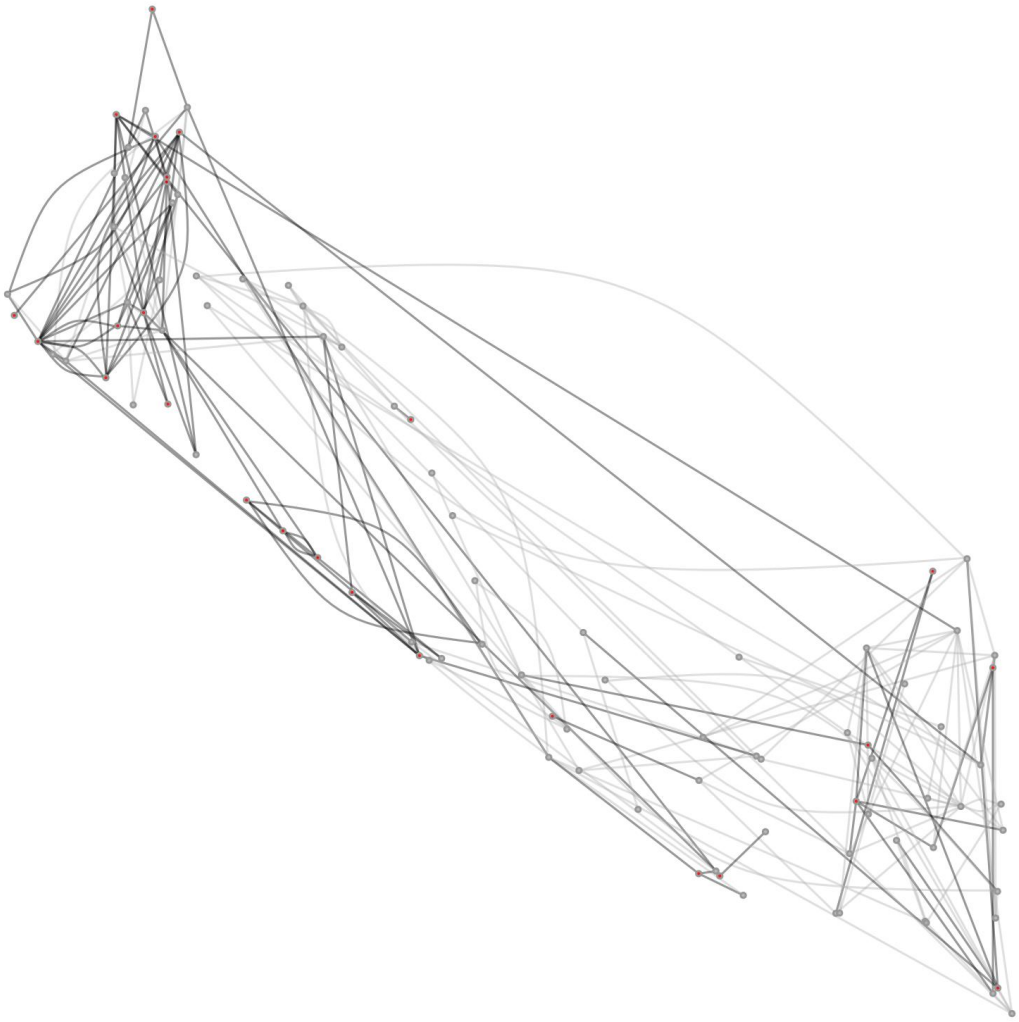

Figure C.44: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santa Maria. Mapping in GPS coordinates of homes and social network (general important matters). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

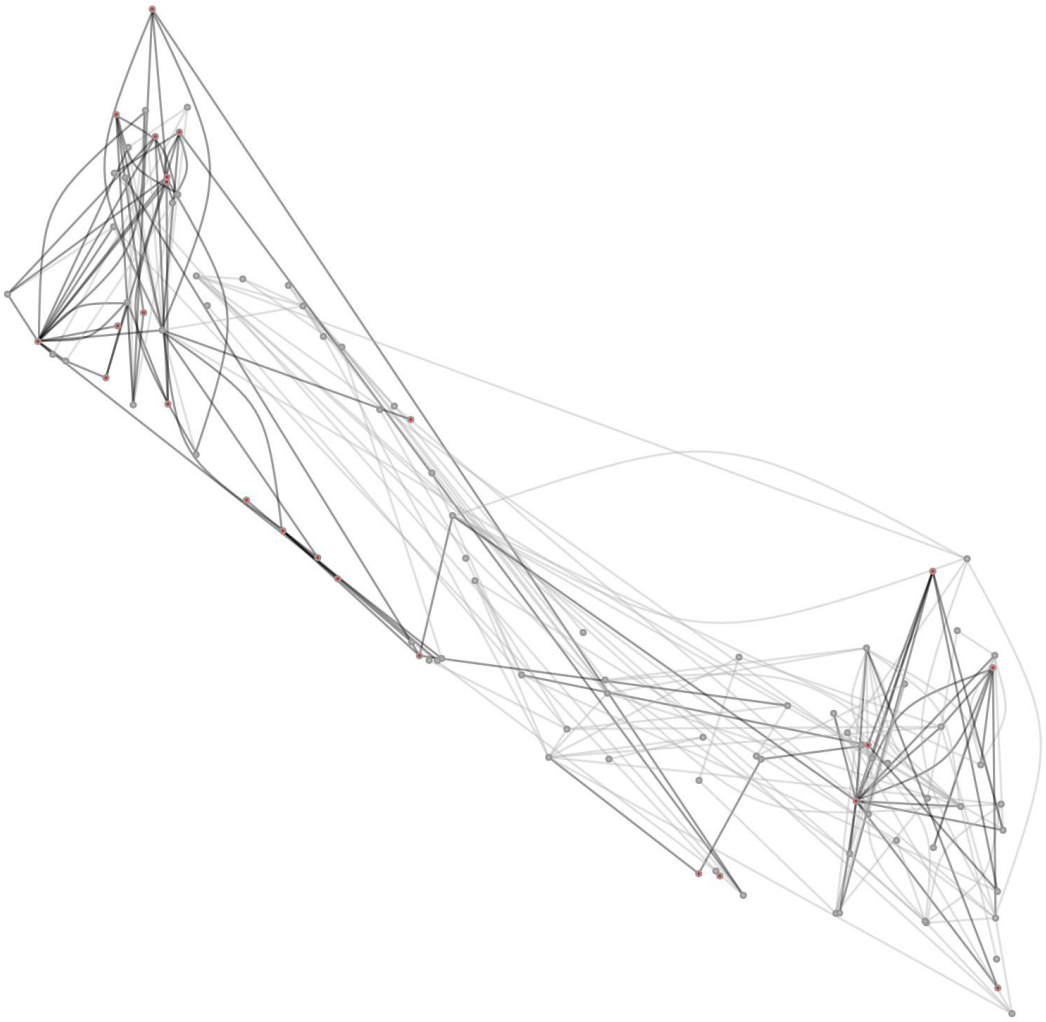

Figure C.45: Cases of diarrhea and/or fever at the household-level in the two weeks prior to sociometric survey administration in winter of 2022 in Santa Maria. Mapping in GPS coordinates of homes and social network (people visited). Red node is a case of diarrhea and/or fever and dark grey edges are connections to those cases.

BIBLIOGRAPHY

[1] World Health Organization. A Conceptual Framework for Action on the Social Determinants of Health. URL https://www.who.int/publications/i/item/9789241500852.

[2] D Pedrazzoli, D Boccia, P J Dodd, K Lönnroth, D W Dowdy, A Siroka, M E Kimerling, R G White, and R M G J Houben. Modelling the social and structural determinants of tuberculosis: Opportunities and challenges. 21(9):957–964. ISSN 1815-7920. doi: 10.5588/ijtld.16.0906. URL https://europepmc.org/articles/PMC5566999.

[3] Eisenberg, Marisa. Michigan COVID-19 Modeling Dashboard. URL https://epimath.github.io/covid-19-modeling/#about.

[4] Joseph N.S. Eisenberg, William Cevallos, Karina Ponce, Karen Levy, Sarah J. Bates, James C. Scott, Alan Hubbard, Nadia Vieira, Pablo Endara, Mauricio Espinel, Gabriel Trueba, Lee W. Riley, and James Trostle. Environmental change and infectious disease: How new roads affect the transmission of diarrheal pathogens in rural Ecuador. 103(51): 19460–19465. doi: 10.1073/pnas.0609431104.

[5] Global Burden of Disease Collaborative Network. Estimates of global, regional, and national morbidity, mortality, and aetiologies of diarrhoeal diseases: A systematic analysis for the Global Burden of Disease Study 2015. 17(9):909–948. ISSN 1473-3099. doi: 10.1016/S1473-3099(17)30276-1. URL https://www.sciencedirect.com/science/article/pii/S1473309917302761.

[6] Ashley M. Fox. The Social Determinants of HIV Serostatus in Sub-Saharan Africa: An Inverse Relationship between Poverty and HIV? 125:16–24. ISSN 0033-3549. doi: 10.1177/00333549101250S405. URL https://doi.org/10.1177/00333549101250S405.

[7] Angelica Geter, Madeline Y. Sutton, and Donna Hubbard McCree. Social and structural determinants of HIV treatment and care among black women living with HIV infection: A systematic review: 2005–2016. 30(4):409–416. ISSN 0954-0121. doi: 10.1080/09540121.2018.1426827. URL https://doi.org/10.1080/09540121.2018.1426827.

[8] Joel Negin, Clive Aspin, Thomas Gadsden, and Charlotte Reading. HIV Among Indigenous peoples: A Review of the Literature on HIV-Related Behaviour Since the Beginning of the Epidemic. 19(9):1720–1734. ISSN 1573-3254. doi: 10.1007/s10461-015-1023-0. URL https://doi.org/10.1007/s10461-015-1023-0.

[9] William L. Jeffries and Kirk D. Henny. From Epidemiology to Action: The Case for Addressing Social Determinants of Health to End HIV in the Southern United States. 23(3):340–346. ISSN 1573-3254. doi: 10.1007/s10461-019-02687-2. URL https://doi.org/10.1007/s10461-019-02687-2.

[10] Samuel Roberts. *Infectious Fear: Politics, Disease, and the Health Effects of Segregation*. Univ of North Carolina Press, . ISBN 978-0-8078-3259-2.

[11] R. Duarte, K. Lönnroth, C. Carvalho, F. Lima, A. C. C. Carvalho, M. Muñoz Torrico, and R. Centis. Tuberculosis, social determinants and co-morbidities (including HIV). 24(2): 115–119. ISSN 2531-0437. doi: 10.1016/j.rppnen.2017.11.003. URL https://www.sciencedirect.com/science/article/pii/S2173511517301641.

[12] Grace A. Noppert, Ryan E. Malosh, Elizabeth B. Moran, Shama D. Ahuja, and Jon Zelner. Contemporary Social Disparities in TB Infection and Disease in the USA: A Review. 5(4): 442–449, . doi: 10.1007/s40471-018-0171-y.

[13] Madhumita Dobe, Ram Narayan Mandal, and Ayan Jha. Social Determinants of Good Hand-Washing Practice (GHP) Among Adolescents in a Rural Indian Community. 36 (2):172–177, April/June 2013. ISSN 0160-6379. doi: 10.1097/FCH.0b013e318282ac42. URL http://journals.lww.com/familyandcommunityhealth/fulltext/2013/04000/social_determinants_of_good_hand_washing_practice.10.aspx.

[14] Jason Corburn and Chantal Hildebrand. Slum Sanitation and the Social Determinants of Women's Health in Nairobi, Kenya. 2015:e209505. ISSN 1687-9805. doi: 10.1155/2015/209505. URL http://www.hindawi.com/journals/jeph/2015/209505/.

[15] Rajib Dasgupta. *Urbanising Cholera: The Social Determinants of Its Re-Emergence*. Number 2 in New Perspectives in South Asian History. Orient BlackSwan. ISBN 978-81-250-4660-8.

[16] Eleanor E. Friedman, Hazel D. Dean, and Wayne A. Duffus. Incorporation of Social Determinants of Health in the Peer-Reviewed Literature: A Systematic Review of Articles Authored by the National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention. 133(4):392–412. ISSN 0033-3549. doi: 10.1177/0033354918774788. URL https://doi.org/10.1177/0033354918774788.

[17] Michael Marmot. Health equity in England: The Marmot review 10 years on. 368:m693. ISSN 1756-1833. doi: 10.1136/bmj.m693. URL http://www.bmj.com/content/368/bmj.m693.

[18] Caroline Buckee, Abdisalan Noor, and Lisa Sattenspiel. Thinking clearly about social aspects of infectious disease transmission. 595(7866):205–213. doi: 10.1038/s41586-021-03694-x. URL https://doi.org/10.1038/s41586-021-03694-x.

[19] Jamie Bedson, Laura A. Skrip, Danielle Pedi, Sharon Abramowitz, Simone Carter, Mohamed F. Jalloh, Sebastian Funk, Nina Gobat, Tamara Giles-Vernick, Gerardo Chowell, prefix=de useprefix=true family=Almeida, given=João Rangel, Rania Elessawi, Samuel V. Scarpino, Ross A. Hammond, Sylvie Briand, Joshua M. Epstein, Laurent Hébert-Dufresne, and Benjamin M. Althouse. A review and agenda for integrated disease models including social and behavioural factors. 5(7):834–846. ISSN 2397-3374. doi: 10.1038/s41562-021-01136-2. URL http://www.nature.com/articles/s41562-021-01136-2.

[20] Jon Zelner, Nina B. Masters, Ramya Naraharisetti, Sanyu Mojola, and Merlin Chowkwanyun. There are no equal opportunity infectors: Epidemiological modelers must rethink our approach to inequality in infection risk. . URL http://arxiv.org/abs/2109.00580.

[21] Johns Hopkins Coronavirus Resource Center. Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). URL https://coronavirus.jhu.edu/map.html.

[22] Sarah J. Bates, James Trostle, William T. Cevallos, Alan Hubbard, and Joseph N.S. Eisenberg. Relating diarrheal disease to social networks and the geographic configuration of communities in rural Ecuador. 166(9):1088–1095. doi: 10.1093/aje/kwm184.

[23] Jonathan L. Zelner, James Trostle, Jason E. Goldstick, William Cevallos, James S. House, and Joseph N.S. Eisenberg. Social connectedness and disease transmission: Social organization, cohesion, village context, and infection risk in rural Ecuador. 102(12):2233–2239, . doi: 10.2105/AJPH.2012.300795.

[24] Carolina Perez-Heydrich, Jill M. Furgurson, Sophia Giebultowicz, Jennifer J. Winston, Mohammad Yunus, Peter Kim Streatfield, and Michael Emch. Social and spatial processes associated with childhood diarrheal disease in Matlab, Bangladesh. 19(1):45–52, . doi: 10.1016/j.healthplace.2012.10.002. URL http://dx.doi.org/10.1016/j.healthplace.2012.10.002.

[25] Michael Emch, Elisabeth D. Root, Sophia Giebultowicz, Mohammad Ali, Carolina Perez-Heydrich, and Mohammad Yunus. Integration of Spatial and Social Network Analysis in Disease Transmission Studies. 102(5):1004–1015, . doi: 10.1080/00045608.2012.671129.

[26] M. Chang, S. L. Groseclose, A. A. Zaidi, and C. R. Braden. An ecological analysis of sociodemographic factors associated with the incidence of salmonellosis, shigellosis, and E. coli O157:H7 infections in US counties. 137(6):810–820. ISSN 0950-2688. doi: 10.1017/S0950268808001477.

[27] Rebecca Jackson, Daniel Smith, Farzaneh Tabnak, and Duc Vugia. Disparities of Shigellosis Rates among California Children by Race/Ethnicity and Census Tract Poverty Level, 2000-2010. 34(8):843–847. ISSN 1532-0987. doi: 10.1097/INF.0000000000000746.

[28] Chengjing Nie, Hairong Li, Linsheng Yang, Gemei Zhong, and Lan Zhang. Socio-economic factors of bacillary dysentery based on spatial correlation analysis in Guangxi Province, China. 9(7):e102020. ISSN 1932-6203. doi: 10.1371/journal.pone.0102020.

[29] Merlin Chowkwanyun and Adolph L. Reed. Racial Health Disparities and Covid-19 — Caution and Context. 383(3):201–203. doi: 10.1056/nejmp2012910. URL https://www.nejm.org/doi/full/10.1056/NEJMp2012910.

[30] Jarvis T Chen, Pamela D Waterman, and Nancy Krieger. COVID-19 and the unequal surge in mortality rates in Massachusetts, by city/town and ZIP Code measures of poverty, household crowding, race/ethnicity,and racialized economic segregation. 19(2), . URL https://www.bostonglobe.com/2020/05/09/nation/disparities-push-coronavirus-death-rates-.

[31] Ankur K. Dalsania, Matthew J. Fastiggi, Aaron Kahlam, Rajvi Shah, Krishan Patel, Stephanie Shiau, Slawa Rokicki, and Michelle DallaPiazza. The Relationship Between Social Determinants of Health and Racial Disparities in COVID-19 Mortality. pages 1–8. doi: 10.1007/s40615-020-00952-y. URL https://doi.org/10.1007/s40615-020-00952-y.

[32] Sean A. P. Clouston, Ginny Natale, and Bruce G. Link. Socioeconomic inequalities in the spread of coronavirus-19 in the United States: A examination of the emergence of social inequalities. 268:113554, . ISSN 1873-5347. doi: 10.1016/j.socscimed.2020.113554.

[33] Ralph Lawton, Kevin Zheng, Daniel Zheng, and Erich Huang. A longitudinal study of convergence between Black and White COVID-19 mortality: A county fixed effects approach. 1. ISSN 2667-193X. doi: 10.1016/j.lana.2021.100011. URL https://www.thelancet.com/journals/lanam/article/PIIS2667-193X(21)00003-X/fulltext.

[34] Daniel T. Lichter, Domenico Parisi, and Michael C. Taquino. Toward a New Macro-Segregation? Decomposing Segregation within and between Metropolitan Cities and Suburbs. 80(4):843–873. doi: 10.1177/0003122415588558. URL http://journals.sagepub.com/doi/10.1177/0003122415588558.

[35] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). 5(4):529–538. ISSN 2397-3374. doi: 10.1038/s41562-021-01079-8. URL https://www.nature.com/articles/s41562-021-01079-8.

[36] Maria Polyakova, Victoria Udalova, Geoffrey Kocks, Katie Genadek, Keith Finlay, and Amy N. Finkelstein. Racial Disparities In Excess All-Cause Mortality During The Early COVID-19 Pandemic Varied Substantially Across States. 40(2):307–316. doi: 10.1377/hlthaff.2020.02142.

[37] Jon Zelner, Rob Trangucci, Ramya Naraharisetti, Alex Cao, Ryan Malosh, Kelly Broen, Nina Masters, and Paul Delamater. Racial Disparities in Coronavirus Disease 2019 (COVID-19) Mortality Are Driven by Unequal Infection Risks. 72(5):e88–e95, . doi: 10.1093/cid/ciaa1723. URL https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa1723/5998295.

[38] Sebastian D. Romano. Trends in Racial and Ethnic Disparities in COVID-19 Hospitalizations, by Region — United States, March–December 2020. 70. ISSN 0149-21951545-861X. doi: 10.15585/mmwr.mm7015e2. URL https://www.cdc.gov/mmwr/volumes/70/wr/mm7015e2.htm.

[39] Grace A. Noppert, Sonia T. Hegde, and John T. Kubale. Exposure, Susceptibility, and Recovery: A Framework for Examining the Intersection of the Social and Physical Environments and Infectious Disease Risk. 192(3):475–482, . ISSN 1476-6256. doi: 10.1093/aje/kwac186.

[40] Philip Blumenshine, Arthur Reingold, Susan Egerter, Robin Mockenhaupt, Paula Braveman, and James Marks. Pandemic Influenza Planning in the United States from a Health Disparities Perspective. 14(5):709–715. ISSN 1080-6040. doi: 10.3201/eid1405.071301. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2600245/.

[41] Christopher T. Rentsch, Farah Kidwai-Khan, Janet P. Tate, Lesley S. Park, Joseph T. King, Melissa Skanderson, Ronald G. Hauser, Anna Schultze, Christopher I. Jarvis, Mark Holodniy, Vincent Lo Re, Kathleen M. Akgün, Kristina Crothers, Tamar H. Taddei, Matthew S. Freiberg, and Amy C. Justice. Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: A nationwide cohort study. 17(9):e1003379–e1003379. doi: 10.1371/journal.pmed.1003379. URL https://dx.plos.org/10.1371/journal.pmed.1003379.

[42] Gbenga Ogedegbe, Joseph Ravenell, Samrachana Adhikari, Mark Butler, Tiffany Cook, Fritz Francois, Eduardo Iturrate, Girardin Jean-Louis, Simon A. Jones, Deborah Onakomaiya, Christopher M. Petrilli, Claudia Pulgarin, Seann Regan, Harmony Reynolds, Azizi Seixas, Frank Michael Volpicelli, and Leora Idit Horwitz. Assessment of Racial/Ethnic Disparities in Hospitalization and Mortality in Patients With COVID-19 in New York City. 3(12):e2026881–e2026881. doi: 10.1001/jamanetworkopen.2020.26881. URL https://jamanetwork.com/.

[43] Donald J. Alcendor. Racial Disparities-Associated COVID-19 Mortality among Minority Populations in the US. 9(8):2442–2442. doi: 10.3390/jcm9082442. URL https://www.mdpi.com/2077-0383/9/8/2442.

[44] Miriam E. Van Dyke, Maria C.B. Mendoza, Wen Li, Erin M. Parker, Brook Belay, Elizabeth M. Davis, Joshua J. Quint, Ana Penman-Aguilar, and Kristie E.N. Clarke. Racial and Ethnic Disparities in COVID-19 Incidence by Age, Sex, and Period Among Persons Aged <25 Years — 16 U.S. Jurisdictions, January 1–December 31, 2020. 70(11):382–388. doi: 10.15585/mmwr.mm7011e1. URL http://www.cdc.gov/mmwr/volumes/70/wr/mm7011e1.htm?s_cid=mm7011e1_w.

[45] Hélène E. Aschmann, Alicia R. Riley, Ruijia Chen, Yea-Hung Chen, Kirsten Bibbins-Domingo, Andrew C. Stokes, M. Maria Glymour, and Mathew V. Kiang. Dynamics of racial disparities in all-cause mortality during the COVID-19 pandemic. 119(40):e2210941119. doi: 10.1073/pnas.2210941119. URL https://www.pnas.org/doi/10.1073/pnas.2210941119.

[46] Alicia R. Riley. Advancing the study of health inequality: Fundamental causes as systems of exposure. 10:100555–100555. doi: 10.1016/j.ssmph.2020.100555. URL https://doi.org/10.1016/j.ssmph.2020.100555.

[47] Kristin N. Nelson, Aaron J. Siegler, Patrick S. Sullivan, Heather Bradley, Eric Hall, Nicole Luisi, Palmer Hipp-Ramsey, Travis Sanchez, Kayoko Shioda, and Benjamin A. Lopman. Nationally representative social contact patterns among U.S. adults, August 2020-April 2021. 40:100605. ISSN 1755-4365. doi: 10.1016/j.epidem.2022.100605. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9242729/.

[48] Joanna Gaitens, Marian Condon, Eseosa Fernandes, and Melissa McDiarmid. COVID-19 and Essential Workers: A Narrative Review of Health Outcomes and Moral Injury. 18(4): 1446. ISSN 1660-4601. doi: 10.3390/ijerph18041446.

[49] Tiana N. Rogers, Charles R. Rogers, Elizabeth VanSant-Webb, Lily Y. Gu, Bin Yan, and Fares Qeadan. Racial Disparities in COVID-19 Mortality Among Essential Workers in the United States. 12(3):311–327. ISSN 2153-2028. doi: 10.1002/wmh3.358. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7436547/.

[50] Yea-Hung Chen, Maria Glymour, Alicia Riley, John Balmes, Kate Duchowny, Robert Harrison, Ellicott Matthay, and Kirsten Bibbins-Domingo. Excess mortality associated with the COVID-19 pandemic among Californians 18–65 years of age, by occupational sector and occupation: March through November 2020. 16(6):e0252454, . ISSN 1932-6203. doi: 10.1371/journal.pone.0252454. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0252454.

[51] Noreen Goldman, Anne R. Pebley, Keunbok Lee, Theresa Andrasfay, and Boriana Pratt. Racial and ethnic differentials in COVID-19-related job exposures by occupational standing in the US. 16(9):e0256085. ISSN 1932-6203. doi: 10.1371/journal.pone.0256085. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0256085.

[52] Amrita Rao, Huiting Ma, Gary Moloney, Jeffrey C. Kwong, Peter Jüni, Beate Sander, Rafal Kustra, Stefan D. Baral, and Sharmistha Mishra. A disproportionate epidemic: COVID-19 cases and deaths among essential workers in Toronto, Canada. 63:63–67. ISSN 1047-2797. doi: 10.1016/j.annepidem.2021.07.010.

[53] The Office of Governor Gretchen Whitmer. Governor Whitmer and Lt. Governor Gilchrist Highlight Significant Reduction in Racial Disparities of COVID-19 Cases and Deaths. URL https://www.michigan.gov/whitmer/0,9309,7-387-90499_90640-540768--,00.html.

[54] prefix=of the useprefix=false family=Commissioner, given=Office. FDA Approves First Treatment for COVID-19, Thu, 10/22/2020 - 17:40. URL https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19.

[55] Emergency Order Under MCL 333.2253 - Gathering Prohibition and Mask Order. URL https://www.michigan.gov/documents/coronavirus/MDHHS_epidemic_order_-_Gatherings_masks_bars_sports_-_FINAL_704287_7.pdf.

[56] Steven Ruggles, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. PUMS USA: Version 11.0 [dataset]. URL https://usa.ipums.org/usa/cite.shtml.

[57] Ken Resnicow, Matthew Jaber Stiffler, and Kristine J. Ajrouch. Looking Back: The Contested Whiteness of Arab Identity. 112(8):1092–1096. ISSN 0090-0036. doi: 10.2105/AJPH.2022.306884. URL https://ajph.aphapublications.org/doi/10.2105/AJPH.2022.306884.

[58] Hubei Province. Michigan State and Local Public Health COVID-19 Standard Operating Procedures. URL https://www.michigan.gov/coronavirus.

[59] Lester R. Curtin. *Direct Standardization (Age-Adjusted Death Rates)*. U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

[60] Nikkil Sudharsanan, Oliver Didzun, Till Bärnighausen, and Pascal Geldsetzer. The Contribution of the Age Distribution of Cases to COVID-19 Case Fatality Across Countries : A Nine-Country Demographic Study. 173(9):714–720. doi: 10.7326/M20-2973. URL https://www.acpjournals.org/doi/10.7326/M20-2973.

[61] W. Holmes Finch and Maria E. Hernández Finch. Poverty and Covid-19: Rates of Incidence and Deaths in the United States During the First 10 Weeks of the Pandemic. 5:47. ISSN 2297-7775. doi: 10.3389/fsoc.2020.00047. URL https://www.frontiersin.org/article/10.3389/fsoc.2020.00047.

[62] Lauren C Zalla, Chantel L Martin, Jessie K Edwards, Danielle R Gartner, and Grace A Noppert. A Geography of Risk: Structural Racism and COVID-19 Mortality in the United States. doi: 10.1093/aje/kwab059. URL https://academic.oup.com/aje/advance-article/doi/10.1093/aje/kwab059/6168675.

[63] Alina S Schnake-Mahl and Usama Bilal. Schnake-Mahl and Bilal Respond to "Structural Racism and COVID-19 Mortality in the US". 190(8):1447–1451. ISSN 0002-9262. doi: 10.1093/aje/kwab058. URL https://doi.org/10.1093/aje/kwab058.

[64] prefix=and useprefix=false family=Sam, given=Ben Goodrich and Jonah Gabry and Imad Ali and Brilleman. Rstanarm: {Bayesian} applied regression modeling via {Stan}. URL https://cran.r-project.org/web/packages/rasterVis/citation.html.

[65] Matthew Kay. Tidybayes: Tidy Data and Geoms for Bayesian Models. R package version 2.3.0. URL http://mjskay.github.io/tidybayes/.

[66] Felix Mölder, Kim Philipp Jablonski, Brice Letcher, Michael B Hall, Christopher H Tomkins-Tinch, Vanessa Sochat, Jan Forster, Soohyun Lee, Sven O Twardziok, Alexander Kanitz, Andreas Wilm, Manuel Holtgrewe, Sven Rahmann, Sven Nahnsen, and Johannes Köster. Sustainable data analysis with Snakemake. 10:33–33. doi: 10.12688/f1000research.29032.1.

[67] Katie Labgold, Sarah Hamid, Sarita Shah, Neel R. Gandhi, Allison Chamberlain, Fazle Khan, Shamimul Khan, Sasha Smith, Steve Williams, Timothy L. Lash, and Lindsay J. Collin. Estimating the Unknown : Greater Racial and Ethnic Disparities in COVID-19 Burden After Accounting for Missing Race and Ethnicity Data. 32(2):157–161. doi: 10.1097/EDE.0000000000001314.

[68] Rob Trangucci, Yang Chen, and Jon Zelner. Modeling racial/ethnic differences in COVID-19 incidence with covariates subject to non-random missingness. URL http://arxiv.org/abs/2206.08161.

[69] Lauren M. Rossen, Amy M. Branum, Farida B. Ahmad, Paul Sutton, and Robert N. Anderson. Excess Deaths Associated with COVID-19, by Age and Race and Ethnicity — United States, January 26–October 3, 2020. 69(42):1522–1527. doi: 10.15585/mmwr.mm6942e2. URL http://www.cdc.gov/mmwr/volumes/69/wr/mm6942e2.htm?s_cid=mm6942e2_w.

[70] Natalie R. Smith, Paul N. Zivich, Leah M. Frerichs, James Moody, and Allison E. Aiello. A Guide for Choosing Community Detection Algorithms in Social Network Studies: The Question Alignment Approach. 59(4):597–605. ISSN 0749-3797. doi: 10.1016/j.amepre.2020.04.015. URL https://www.sciencedirect.com/science/article/pii/S0749379720302166.

[71] Mohammad Nayeem Hasan, Najmul Haider, Florian L. Stigler, Rumi Ahmed Khan, David McCoy, Alimuddin Zumla, Richard A. Kock, and Md Jamal Uddin. The Global Case-Fatality Rate of COVID-19 Has Been Declining Since May 2020. 104(6):2176–2184. ISSN 1476-1645. doi: 10.4269/ajtmh.20-1496.

[72] Usama Bilal, Sharrelle Barber, Loni Tabb, and Ana V. Diez-Roux. Spatial inequities in COVID-19 testing, positivity, incidence and mortality in 3 US cities: A longitudinal ecological study. pages 2020.05.01.20087833–2020.05.01.20087833. doi: 10.1101/2020.05.01.20087833. URL https://doi.org/10.1101/2020.05.01.20087833.

[73] Christina Kamis, Allison Stolte, Jessica S. West, Samuel H. Fishman, Taylor Brown, Tyson Brown, and Heather R. Farmer. Overcrowding and COVID-19 mortality across U.S. counties: Are disparities growing over time? 15:100845. ISSN 2352-8273. doi: 10.1016/j.ssmph.2021.100845. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8219888/.

[74] Luisa N. Borrell and Natalie D. Crawford. All-Cause Mortality Among Hispanics in the United States: Exploring Heterogeneity by Nativity Status, Country of Origin, and Race in the National Health Interview Survey-Linked Mortality Files. 19(5):336–343. ISSN 1047-2797. doi: 10.1016/j.annepidem.2008.12.003. URL https://www.sciencedirect.com/science/article/pii/S1047279708003621.

[75] Sean A.P. Clouston, Marcie S. Rubin, Jo C. Phelan, and Bruce G. Link. A Social History of Disease: Contextualizing the Rise and Fall of Social Inequalities in Cause-Specific Mortality. 53(5):1631–1656, . doi: 10.1007/s13524-016-0495-5. URL http://read.dukeupress.edu/demography/article-pdf/53/5/1631/837637/1631clouston.pdf.

[76] Whitney N. Laster Pirtle. Racial Capitalism: A Fundamental Cause of Novel Coronavirus (COVID-19) Pandemic Inequities in the United States. 47(4):504–508. doi: 10.1177/1090198120922942.

[77] Jon Zelner, Kelly Broen, and Ella August. A guide to backward paper writing for the data sciences. 0(0), . ISSN 2666-3899. doi: 10.1016/j.patter.2021.100423. URL https://www.cell.com/patterns/abstract/S2666-3899(21)00306-8.

[78] Lonnie R. Snowden and Genevieve Graaf. COVID-19, Social Determinants Past, Present, and Future, and African Americans' Health. 8(1):12–20. ISSN 2197-3792. doi: 10.1007/s40615-020-00923-3. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7682952/.

[79] Dirk Witteveen. Sociodemographic inequality in exposure to COVID-19-induced economic hardship in the United Kingdom. 69:100551. ISSN 0276-5624. doi: 10.1016/j.rssm.2020.100551. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7474909/.

[80] Disability Inclusion Overview, . URL https://www.worldbank.org/en/topic/disability.

[81] Matthew M. Coates, Majid Ezzati, Gisela Robles Aguilar, Gene F. Kwan, Daniel Vigo, Ana O. Mocumbi, Anne E. Becker, Julie Makani, Adnan A. Hyder, Yogesh Jain, D. Cristina Stefan, Neil Gupta, Andrew Marx, and Gene Bukhman. Burden of disease among the world's poorest billion people: An expert-informed secondary analysis of Global Burden of Disease estimates. 16(8):e0253073. ISSN 1932-6203. doi: 10.1371/journal.pone.0253073. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0253073.

[82] Yasuhiko Kamiya. Current situation of children with disabilities in low- and middle-income countries. 63(11):1277–1281. ISSN 1442-200X. doi: 10.1111/ped.14904.

[83] Mónica Pinilla-Roncancio, Islay Mactaggart, Hannah Kuper, Carlos Dionicio, Jonathan Naber, G.V.S. Murthy, and Sarah Polack. Multidimensional poverty and disability: A case control study in India, Cameroon, and Guatemala. 11:100591. ISSN 2352-8273.

doi: 10.1016/j.ssmph.2020.100591. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7212179/.

[84] Lena Morgon Banks, Hannah Kuper, and Sarah Polack. Poverty and disability in low- and middle-income countries: A systematic review. 12(12):e0189996. ISSN 1932-6203. doi: 10.1371/journal.pone.0189996. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5739437/.

[85] Maureen Durkin. *CONTROL OF INFECTIOUS CAUSES OF CHILDHOOD DISABILITY IN DEVELOPING COUNTRIES*. National Academies Press (US), . URL https://www.ncbi.nlm.nih.gov/books/NBK83707/.

[86] Maureen Durkin. The epidemiology of developmental disabilities in low-income countries. 8(3):206–211, . ISSN 1080-4013. doi: 10.1002/mrdd.10039.

[87] Anna Dimitrova, Gabriel Carrasco-Escobar, Robin Richardson, and Tarik Benmarhnia. Essential childhood immunization in 43 low- and middle-income countries: Analysis of spatial trends and socioeconomic inequalities in vaccine coverage. 20(1):e1004166. ISSN 1549-1676. doi: 10.1371/journal.pmed.1004166. URL https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1004166.

[88] Deepti Bettampadi, James M. Lepkowski, Ananda Sen, Laura E. Power, and Matthew L. Boulton. Vaccination Inequality in India, 2002–2013. 60:S65–S76. ISSN 0749-3797. doi: 10.1016/j.amepre.2020.06.034. URL https://www.sciencedirect.com/science/article/pii/S0749379720304037.

[89] Maureen Durkin. *CONTROL OF INFECTIOUS CAUSES OF CHILDHOOD DISABILITY IN DEVELOPING COUNTRIES*. National Academies Press (US), . URL https://www.ncbi.nlm.nih.gov/sites/books/NBK83707/.

[90] Andrew J. Tatem, Susana Adamo, Nita Bharti, Clara R. Burgert, Marcia Castro, Audrey Dorelien, Gunter Fink, Catherine Linard, Mendelsohn John, Livia Montana, Mark R. Montgomery, Andrew Nelson, Abdisalan M. Noor, Deepa Pindolia, Greg Yetman, and Deborah Balk. Mapping populations at risk: Improving spatial demographic data for infectious disease modeling and metric derivation. 10(1):8. ISSN 1478-7954. doi: 10.1186/1478-7954-10-8. URL https://doi.org/10.1186/1478-7954-10-8.

[91] Suguru Mizunoya, Sophie Mitra, and Izumi Yamasaki. Disability and school attendance in 15 low- and middle-income countries. 104:388–403. ISSN 0305-750X. doi: 10.1016/j.worlddev.2017.12.001. URL https://www.sciencedirect.com/science/article/pii/S0305750X17303881.

[92] Ahmad Reza Hosseinpoor, Nicole Bergen, Nenad Kostanjsek, Paul Kowal, Alana Officer, and Somnath Chatterji. Socio-demographic patterns of disability among older adult populations of low-income and middle-income countries: Results from World Health Survey. 61(3):337–345. ISSN 1661-8564. doi: 10.1007/s00038-015-0742-3. URL https://doi.org/10.1007/s00038-015-0742-3.

[93] Calistus N. Ngonghala, Mateusz M. Pluciński, Megan B. Murray, Paul E. Farmer, Christopher B. Barrett, Donald C. Keenan, and Matthew H. Bonds. Poverty, Disease, and the Ecology of Complex Systems. 12(4):1–9. doi: 10.1371/journal.pbio.1001827.

[94] Nora Groce, Gayatri Kembhavi, Shelia Wirz, Raymond Lang, Jean-Francois Trani, and Maria Kett. Poverty and Disability – A Critical Review of the Literature in Low and Middle-Income Countries. ISSN 1556-5068. doi: 10.2139/ssrn.3398431. URL https://www.ssrn.com/abstract=3398431.

[95] Pallab K. Maulik and Gary L. Darmstadt. Childhood Disability in Low- and Middle-Income Countries: Overview of Screening, Prevention, Services, Legislation, and Epidemiology. 120:S1–S55. ISSN 0031-4005. doi: 10.1542/peds.2007-0043B. URL https://doi.org/10.1542/peds.2007-0043B.

[96] The World Bank Group. People with Disabilities in India: From Committments to Outcomes. URL https://documents1.worldbank.org/curated/en/577801468259486686/pdf/502090WP0Peopl1Box0342042B01PUBLIC1.pdf.

[97] Baikunth Roy. Temporal and Spatial Discourse of Disability Divides in India: A District-Level Analysis. 5(1):11–23. ISSN 2454–2806. URL http://www.socialspectrum.in/index.php/sp/article/view/146.

[98] Rakhi Dandona, Anamika Pandey, Sibin George, G. Anil Kumar, and Lalit Dandona. India's disability estimates: Limitations and way forward. 14(9):e0222159. ISSN 1932-6203. doi: 10.1371/journal.pone.0222159. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0222159.

[99] Dilip C. Nath and Mousumi Nath. Measuring Disability in India: Spatial and Socio-economic Variations. URL https://paa2006.princeton.edu/papers/60161.

[100] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. URL https://www.R-project.org/.

[101] Edzer Pebesma. Sf. 10(1):439–446. doi: 10.32614/RJ-2018-009. URL https://doi.org/10.32614/RJ-2018-009.

[102] Martijn Tennekes. Tmap. 84(6):1–39. doi: 10.18637/jss.v084.i06.

[103] Luc Anselin. Local Indicators of Spatial Association—LISA. 27(2):93–115. ISSN 1538-4632. doi: 10.1111/j.1538-4632.1995.tb00338.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1538-4632.1995.tb00338.x.

[104] Marcia Caldas de Castro and Burton H. Singer. Controlling the False Discovery Rate: A New Application to Account for Multiple and Dependent Tests in Local Statistics of Spatial Association. 38(2):180–208. ISSN 1538-4632. doi: 10.1111/j.0016-7363.2006.00682.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0016-7363.2006.00682.x.

[105] PTI. Andhra Pradesh, Maharashtra haven't framed rules under Disabilities Act: Union govt. URL newindianexpress.com/nation/2021/jul/28/andhra-pradesh-maharashtra-havent-framed-rules-under-disabilities-act-html#:~:text="According%20to%20the%20information%20available,said%20in%20a%20written%20reply.

[106] Amit Thorat. Rural Poor: Who Are They and Why? A Case Study of Madhya Pradesh.

[107] Sophie Mitra and Usha Sambamoorthi. Disability Estimates in India: What the Census and NSS Tell Us. 41(38):4022–4026. ISSN 0012-9976. URL https://www.jstor.org/stable/4418721.

[108] Johns Hopkins Bloomberg School of Public Health and Interational Vaccine Access Center. Pneumonia & Diarrhea Progress Reports. URL https://www.jhsph.edu/ivac/resources/pdpr/.

[109] Isobel L. Gabain, Anouschka S. Ramsteijn, and Joanne P. Webster. Parasites and childhood stunting – a mechanistic interplay with nutrition, anaemia, gut health, microbiota, and epigenetics. 39(3):167–180. ISSN 1471-4922, 1471-5007. doi: 10.1016/j.pt.2022.12.004. URL https://www.cell.com/trends/parasitology/abstract/S1471-4922(22)00309-9.

[110] Guy Hutton, Laurence Haller, and Jamie Bartram. Global cost-benefit analysis of water supply and sanitation interventions. 5(4):481–502. ISSN 1477-8920, 1996-7829. doi: 10.2166/wh.2007.009. URL https://iwaponline.com/jwh/article/5/4/481/31318/Global-costbenefit-analysis-of-water-supply-and.

[111] Julia Vila-Guilera, Priti Parikh, Hemant Chaturvedi, Lena Ciric, and Monica Lakhanpaul. Towards transformative WASH: An integrated case study exploring environmental, sociocultural, economic and institutional risk factors contributing to infant enteric infections in rural tribal India. 21(1):1331. ISSN 1471-2458. doi: 10.1186/s12889-021-11353-z. URL https://doi.org/10.1186/s12889-021-11353-z.

[112] Oliver Cumming, Benjamin F. Arnold, Radu Ban, Thomas Clasen, Joanna Esteves Mills, Matthew C. Freeman, Bruce Gordon, Raymond Guiteras, Guy Howard, Paul R. Hunter, Richard B. Johnston, Amy J. Pickering, Andrew J. Prendergast, Annette Prüss-Ustün, Jan Willem Rosenboom, Dean Spears, Shelly Sundberg, Jennyfer Wolf, Clair Null, Stephen P. Luby, Jean H. Humphrey, and John M. Colford. The implications of three major new trials for the effect of water, sanitation and hygiene on childhood diarrhea and stunting: A consensus statement. 17(1):1–9. doi: 10.1186/s12916-019-1410-x.

[113] Tom Henfrey, Giuseppe Feola, Gil Penha-Lopes, Filka Sekulova, and Ana Margarida Esteves. Rethinking the sustainable development goals: Learning with and from community-led initiatives. 31(1):211–222. ISSN 1099-1719. doi: 10.1002/sd.2384. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sd.2384.

[114] Gordon McGranahan and Diana Mitlin. Learning from Sustained Success: How Community-Driven Initiatives to Improve Urban Sanitation Can Meet the Challenges. 87:

307–317. ISSN 0305-750X. doi: 10.1016/j.worlddev.2016.06.019. URL https://www.sciencedirect.com/science/article/pii/S0305750X15305441.

[115] Tyler J. VanderWeele and Weihua An. Social Networks and Causal Inference. In Stephen L. Morgan, editor, *Handbook of Causal Analysis for Social Research*, Handbooks of Sociology and Social Research, pages 353–374. Springer Netherlands. ISBN 978-94-007-6093-6 978-94-007-6094-3. doi: 10.1007/978-94-007-6094-3_17. URL http://link.springer.com/10.1007/978-94-007-6094-3_17.

[116] Jessica M. Perkins, S. V. Subramanian, and Nicholas A. Christakis. Social networks and health: A systematic review of sociocentric network studies in low- and middle-income countries. 125:60–78. doi: 10.1016/j.socscimed.2014.08.019. URL http://dx.doi.org/10.1016/j.socscimed.2014.08.019.

[117] Carolina Perez-Heydrich, Jill M. Furgurson, Sophia Giebultowicz, Jennifer J. Winston, Mohammad Yunus, Peter Kim Streatfield, and Michael Emch. Social and spatial processes associated with childhood diarrheal disease in Matlab, Bangladesh. 19:45–52, . ISSN 1353-8292. doi: 10.1016/j.healthplace.2012.10.002. URL https://www.sciencedirect.com/science/article/pii/S1353829212001700.

[118] Sonia T. Hegde, James A. Trostle, Bhramar Mukherjee, and Joseph N. S. Eisenberg. The importance of community during rapid development: The influence of social networks on acute gastrointestinal illness in rural Ecuador. 19:101159. ISSN 2352-8273. doi: 10.1016/j.ssmph.2022.101159. URL https://www.sciencedirect.com/science/article/pii/S2352827322001380.

[119] Tyler J. VanderWeele and Mirjam J. Knol. A Tutorial on Interaction. 3(1):33–72. ISSN 2161-962X. doi: 10.1515/em-2013-0005. URL https://www.degruyter.com/document/doi/10.1515/em-2013-0005/html?lang=en%2F.

[120] Gregory F. Albery, Lucinda Kirkpatrick, Josh A. Firth, and Shweta Bansal. Unifying spatial and social network analysis in disease ecology. 90(1):45–61. ISSN 1365-2656. doi: 10.1111/1365-2656.13356. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1365-2656.13356.

[121] Walter Nicholls. Place, networks, space: Theorising the geographies of social movements. 34(1):78–93. ISSN 1475-5661. doi: 10.1111/j.1475-5661.2009.00331.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-5661.2009.00331.x.

[122] Matthias Kowald, prefix=van den useprefix=true family=Berg, given=Pauline, Andreas Frei, Juan-Antonio Carrasco, Theo Arentze, Kay Axhausen, Diana Mok, Harry Timmermans, and Barry Wellman. Distance patterns of personal networks in four countries: A comparative study. 31:236–248. ISSN 0966-6923. doi: 10.1016/j.jtrangeo.2013.06.006. URL https://www.sciencedirect.com/science/article/pii/S0966692313001166.

[123] Michael Emch, Mohammad Ali, Elisabeth D. Root, and Mohammad Yunus. Spatial and environmental connectivity analysis in a cholera vaccine trial. 68(4):631–637, . ISSN 0277-9536. doi: 10.1016/j.socscimed.2008.11.025.

[124] Carolina Perez-Heydrich, Jill M. Furgurson, Sophia Giebultowicz, Jennifer J. Winston, Mohammad Yunus, Peter Kim Streatfield, and Michael Emch. Social and spatial processes associated with childhood diarrheal disease in Matlab, Bangladesh. 19(1):45–52, . doi: 10.1016/j.healthplace.2012.10.002. URL http://dx.doi.org/10.1016/j.healthplace.2012.10.002.

[125] Jade Benjamin-Chung, Jaynal Abedin, David Berger, Ashley Clark, Veronica Jimenez, Eugene Konagaya, Diana Tran, Benjamin F. Arnold, Alan E. Hubbard, Stephen P. Luby, Edward Miguel, and John M. Colford. Spillover effects on health outcomes in low- and middle-income countries: A systematic review. 46(4):1251–1276. ISSN 1464-3685. doi: 10.1093/ije/dyx039.

[126] David J. Haw, Rachael Pung, Jonathan M. Read, and Steven Riley. Strong spatial embedding of social networks generates nonstandard epidemic dynamics independent of degree distribution and clustering. 117(38):23636–23642. doi: 10.1073/pnas.1910181117. URL https://www.pnas.org/doi/abs/10.1073/pnas.1910181117.

[127] Sonia Hegde. The Science of Human Connection: A Study of the Effect of Social Networks on Acute Gastrointestinal Illness in Rural Ecuadorian Communities. URL https://scholar.google.com/citations?view_op=view_citation&hl=en&user=_KhQKnUAAAAJ&cstart=20&pagesize=80&citation_for_view=_KhQKnUAAAAJ:kNdYIx-mwKoC.

[128] Professor Michael G Baker, Dr Andrea McDonald, Jane Zhang, and Professor Philippa Howden-Chapman. Infectious Diseases Attributable to Household Crowding in New Zealand: A Systematic Review and Burden of Disease Estimate.

[129] Josh M. Colston, Abu S. G. Faruque, M. Jahangir Hossain, Debasish Saha, Suman Kanungo, Inácio Mandomando, M. Imran Nisar, Anita K. M. Zaidi, Richard Omore, Robert F. Breiman, Samba O. Sow, Anna Roose, Myron M. Levine, Karen L. Kotloff, Tahmeed Ahmed, Pascal Bessong, Zulfiqar Bhutta, Estomih Mduma, Pablo Penatero Yori, Prakash Sunder Shrestha, Maribel P. Olortegui, Gagandeep Kang, Aldo A. M. Lima, Jean Humphrey, Andrew Prendergast, Francesca Schiaffino, Benjamin F. Zaitchik, and Margaret N. Kosek. Associations between Household-Level Exposures and All-Cause Diarrhea and Pathogen-Specific Enteric Infections in Children Enrolled in Five Sentinel Surveillance Studies. 17 (21):8078. ISSN 1660-4601. doi: 10.3390/ijerph17218078. URL https://www.mdpi.com/1660-4601/17/21/8078.

[130] Dina Mistry, Maria Litvinova, Ana Pastore y Piontti, Matteo Chinazzi, Laura Fumanelli, Marcelo F. C. Gomes, Syed A. Haque, Quan-Hui Liu, Kunpeng Mu, Xinyue Xiong, M. Elizabeth Halloran, Ira M. Longini, Stefano Merler, Marco Ajelli, and Alessandro Vespignani. Inferring high-resolution human mixing patterns for disease modeling. 12(1):323. ISSN 2041-1723. doi: 10.1038/s41467-020-20544-y. URL https://www.nature.com/articles/s41467-020-20544-y.

[131] Hans Heesterbeek, Roy M. Anderson, Viggo Andreasen, Shweta Bansal, Daniela De Angelis, Chris Dye, Ken T. D. Eames, W. John Edmunds, Simon D. W. Frost, Sebastian Funk, T. Deirdre Hollingsworth, Thomas House, Valerie Isham, Petra Klepac, Justin Lessler, James O. Lloyd-Smith, C. Jessica E. Metcalf, Denis Mollison, Lorenzo Pellis, Juliet R. C. Pulliam, Mick G. Roberts, Cecile Viboud, and Isaac Newton Institute IDD Collaboration. Modeling infectious disease dynamics in the complex landscape of global health. 347(6227):aaa4339. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aaa4339. URL https://www.science.org/doi/10.1126/science.aaa4339.

[132] Stephen A. Mein. COVID-19 and Health Disparities: The Reality of "the Great Equalizer". 35(8):2439–2440. ISSN 1525-1497. doi: 10.1007/s11606-020-05880-5. URL https://doi.org/10.1007/s11606-020-05880-5.

[133] J. S. Peters and Andrea Wolper. *Women's Rights, Human Rights: International Feminist Perspectives*. Routledge. ISBN 978-1-317-32548-2.

[134] Michael Karlberg. Constructive Resilience: The Bahá'í Response to Oppression. 35(2):222–257. ISSN 1468-0130. doi: 10.1111/j.1468-0130.2009.00627.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0130.2009.00627.x.

[135] Zeinabou Niamé Daffé, Yodeline Guillaume, and Louise C. Ivers. Anti-Racism and Anti-Colonialism Praxis in Global Health—Reflection and Action for Practitioners in US Academic Medical Centers. 105(3):557–560. ISSN 0002-9637. doi: 10.4269/ajtmh.21-0187. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8592354/.

[136] Rob Stein. Race, Ethnicity Data To Be Required With Coronavirus Tests In U.S. URL https://www.npr.org/sections/coronavirus-live-updates/2020/06/04/869815033/race-ethnicity-data-to-be-required-with-coronavirus-tests-in-u-s.

[137] Grace A. Noppert and Lauren C. Zalla. Who Counts and Who Gets Counted? Health Equity in Infectious Disease Surveillance. 111(6):1004–1006. ISSN 0090-0036. doi: 10.2105/AJPH.2021.306249. URL https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2021.306249.

[138] J W Buehler, D F Stroup, D N Klaucke, and R L Berkelman. The reporting of race and ethnicity in the National Notifiable Diseases Surveillance System. 104(5):457–465. ISSN 0033-3549. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1579957/.

[139] Grace McCarthy, Sarah Shore, Esra Ozdenerol, Altha Stewart, Arash Shaban-Nejad, and David L. Schwartz. History Repeating—How Pandemics Collide with Health Disparities in the United States. ISSN 2196-8837. doi: 10.1007/s40615-022-01331-5. URL https://doi.org/10.1007/s40615-022-01331-5.

[140] Paula M. Frew, Julia E. Painter, Brooke Hixson, Carolyn Kulb, Kathryn Moore, prefix=del useprefix=true family=Rio, given=Carlos, Alejandra Esteves-Jaramillo, and Saad B.

Omer. Factors mediating seasonal and influenza A (H1N1) vaccine acceptance among ethnically diverse populations in the urban south. 30(28):4200–4208. ISSN 0264-410X. doi: 10.1016/j.vaccine.2012.04.053. URL https://www.sciencedirect.com/science/article/pii/S0264410X12005890.

[141] Evan T Sholle, Laura C Pinheiro, Prakash Adekkanattu, Marcos A Davila, III, Stephen B Johnson, Jyotishman Pathak, Sanjai Sinha, Cassidie Li, Stasi A Lubansky, Monika M Safford, and Thomas R Campion, Jr. Underserved populations with missing race ethnicity data differ significantly from those with structured race/ethnicity documentation. 26(8-9): 722–729. ISSN 1527-974X. doi: 10.1093/jamia/ocz040. URL https://doi.org/10.1093/jamia/ocz040.

[142] Serin Edwin Erayil, M. Kumi Smith, Tsige Gebreslasse, Patricia F. Walker, Erin M. Mann, Syreeta Wilkins, and William M. Stauffer. The Value and Interpretation of Race and Ethnicity Data in the Era of Global Migration: A Change Is in Order. 105(6):1453–1455. ISSN 0002-9637. doi: 10.4269/ajtmh.21-0665. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8641339/.

[143] Shweta Bansal, Gerardo Chowell, Lone Simonsen, Alessandro Vespignani, and Cécile Viboud. Big Data for Infectious Disease Surveillance and Modeling. 214:S375–S379. ISSN 0022-1899. doi: 10.1093/infdis/jiw400. URL https://doi.org/10.1093/infdis/jiw400.

[144] Marcel Salathé, Linus Bengtsson, Todd J. Bodnar, Devon D. Brewer, John S. Brownstein, Caroline Buckee, Ellsworth M. Campbell, Ciro Cattuto, Shashank Khandelwal, Patricia L. Mabry, and Alessandro Vespignani. Digital Epidemiology. 8(7):e1002616. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1002616. URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002616.

[145] Anpotowin Jensen and Victor A. Lopez-Carmen. The "Elephants in the Room" in U.S. global health: Indigenous nations and white settler colonialism. 2(7):e0000719. ISSN 2767-3375. doi: 10.1371/journal.pgph.0000719. URL https://journals.plos.org/globalpublichealth/article?id=10.1371/journal.pgph.0000719.

[146] Shailendra Prasad, Myron Aldrink, Bruce Compton, Judy Lasker, Peter Donkor, David Weakliam, Virginia Rowthorn, Efua Mantey, Keith Martin, Francis Omaswa, Habib Benzian, Erwin Clagua-Guerra, Emilly Maractho, Kwame Agyire-Tettey, Nigel Crisp, and Ramaswami Balasubramaniam. Global Health Partnerships and the Brocher Declaration: Principles for Ethical Short-Term Engagements in Global Health. 88(1):31. ISSN 2214-9996. doi: 10.5334/aogh.3577. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9122001/.

[147] Margaret W. Gichane and Deshira D. Wallace. Dismantling and reimagining global health education. 15(1):2131967. ISSN 1654-9716. doi: 10.1080/16549716.2022.2131967. URL https://doi.org/10.1080/16549716.2022.2131967.

[148] Peter N.-Jonaam Mahama, Amos Tiereyangn Kabo-Bah, Justine I. Blanford, Edmund Ilimoan Yamba, and Prince Antwi-Agyei. Reviewing the Past, Present, and Future Risks of

Pathogens in Ghana and What This Means for Rethinking Infectious Disease Surveillance for Sub-Saharan Africa. 2022:e4589007. ISSN 1687-9686. doi: 10.1155/2022/4589007. URL https://www.hindawi.com/journals/jtm/2022/4589007/.

[149] Jon Zelner, Nina B. Masters, Ramya Naraharisetti, Sanyu Mojola, and Merlin Chowkwanyun. There are no equal opportunity infectors: Epidemiological modelers must rethink our approach to inequality in infection risk. . URL http://arxiv.org/abs/2109.00580.

[150] Yasser Payne. Participatory Action Research and Social Justice. pages 265–280. ISBN 978-0-203-79155-4. doi: 10.4324/9780203791554-14.

[151] Donald E. Morisky, Alfonso Ang, Astou Coly, and Teodora V. Tiglao. A model HIV/AIDS risk reduction programme in the Philippines: A comprehensive community-based approach through participatory action research. 19(1):69–76. ISSN 0957-4824. doi: 10.1093/heapro/dah109. URL https://doi.org/10.1093/heapro/dah109.

[152] Oghenebrume Wariri, Lucia D'Ambruoso, Rhian Twine, Sizzy Ngobeni, prefix=van der useprefix=true family=Merwe, given=Maria, Barry Spies, Kathleen Kahn, Stephen Tollman, Ryan G Wagner, and Peter Byass. Initiating a participatory action research process in the Agincourt health and socio–demographic surveillance site. 7(1):010413. ISSN 2047-2978. doi: 10.7189/jogh.07.010413. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5475314/.

[153] A. Miller, P. D. Massey, J. Judd, J. Kelly, D. N. Durrhein, A. R. Clough, R. Speare, and S. Saggers. Using a participatory action research framework to listen to Aboriginal and Torres Strait Islander people in Australia about pandemic influenza. 15(3):188–197. doi: 10.3316/informit.235204014818711. URL https://search.informit.org/doi/abs/10.3316/informit.235204014818711.

[154] Kathleen A. Crowley, Ronnie Myers, Halley E. M. Riley, Stephen S. Morse, Paul Brandt-Rauf, and Robyn R. M. Gershon. Using Participatory Action Research to Identify Strategies to Improve Pandemic Vaccination. 7(4):424–430. ISSN 1935-7893, 1938-744X. doi: 10.1017/dmp.2013.72. URL https://www.cambridge.org/core/journals/disaster-medicine-and-public-health-preparedness/article/abs/using-participatory-action-research-to-identify-strategies-to-impro
8C10471C4467E26082F0E756F0D07658.

[155] Sue Devlin, Wayne Ross, Richard Widders, Gregory McAvoy, Kirsty Browne, Kerryn Lawrence, David MacLaren, Peter D Massey, and Jenni A Judd. Tuberculosis care designed with barramarrany (family): Participatory action research that prioritised partnership, healthy housing and nutrition. 33(3):724–735. ISSN 2201-1617. doi: 10.1002/hpja.554. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hpja.554.

[156] Vanessa Leonie Sparke. Improving infection control at Atoifi Adventist Hospital, Solomon Islands: A participatory action research approach. URL https://doi.org/10.25903/4tqh-mt24.

[157] Michael Mulondo. Participatory action research approach to address the poor water, sanitation and hygiene conditions in an informal urban settlement in Windhoek, Namibia. URL http://etd.uwc.ac.za/xmlui/handle/11394/7684.

[158] Jo C. Phelan and Bruce G. Link. Is Racism a Fundamental Cause of Inequalities in Health? 41:311–330. doi: 10.1146/annurev-soc-073014-112305.

[159] Jon Zelner, Ramya Naraharisetti, and Sarah Zelner. To make long-term gains against infection inequity, infectious disease epidemiology needs to develop a more sociological imagination. page kwad044, . ISSN 0002-9262. doi: 10.1093/aje/kwad044. URL https://doi.org/10.1093/aje/kwad044.

[160] Alan G. Nyitray, Dańelle Smith, Luisa Villa, Eduardo Lazcano-Ponce, Martha Abrahamsen, Mary Papenfuss, and Anna R. Giuliano. Prevalence of and Risk Factors for Anal Human Papillomavirus Infection in Men Who Have Sex with Women: A Cross-National Study. 201 (10):1498–1508. ISSN 0022-1899. doi: 10.1086/652187. URL https://doi.org/10.1086/652187.

[161] Ann Versporten, Ganna Bolokhovets, Lilit Ghazaryan, Vafa Abilova, Galina Pyshnik, Tijana Spasojevic, Irma Korinteli, Lul Raka, Baktygul Kambaralieva, Lidija Cizmovic, Angela Carp, Vesela Radonjic, Nargis Maqsudova, Hatice Demet Celik, Marina Payerl-Pal, Hanne Bak Pedersen, Nina Sautenkova, and Herman Goossens. Antibiotic use in eastern Europe: A cross-national database study in coordination with the WHO Regional Office for Europe. 14(5):381–387. ISSN 1473-3099, 1474-4457. doi: 10.1016/S1473-3099(14)70071-4. URL https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(14)70071-4/fulltext?cc=y.

[162] A. D. Lopez and C. D. Mathers. Measuring the global burden of disease and epidemiological transitions: 2002–2030. 100(5-6):481–499. ISSN 0003-4983. doi: 10.1179/136485906X97417. URL https://doi.org/10.1179/136485906X97417.

[163] Samuel Roberts. *Infectious Fear: Politics, Disease, and the Health Effects of Segregation*. Univ of North Carolina Press, .

[164] The Viral Underclass, . URL https://us.macmillan.com/books/9781250796639/theviralunderclass.

[165] Nayan Shah. *Contagious Divides: Epidemics and Race in San Francisco's Chinatown*. ISBN 978-0-520-22629-6.

[166] The Hidden Words — Bahá'í Reference Library, . URL https://www.bahai.org/library/authoritative-texts/bahaullah/hidden-words/.

[167] Dipto Sarkar. Spatial Social Networks: Exploring theoretical and methodological challenges.

[168] Tyler J. VanderWeele. On the Distinction Between Interaction and Effect Modification. 20(6):863–871. ISSN 1044-3983. URL https://www.jstor.org/stable/25662776.