Applications of Dynamic Modeling and Statistical Analysis to Infectious Diseases

Alyssa Sholeh Parpia
Yale University Graduate School of Arts and Sciences, alyssa.parpia@gmail.com

Follow this and additional works at: https://elischolar.library.yale.edu/gsas_dissertations

Recommended Citation
Parpia, Alyssa Sholeh, "Applications of Dynamic Modeling and Statistical Analysis to Infectious Diseases" (2022). Yale Graduate School of Arts and Sciences Dissertations. 642.
https://elischolar.library.yale.edu/gsas_dissertations/642

This Dissertation is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Graduate School of Arts and Sciences Dissertations by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
Abstract

Applications of Dynamic Modeling and Statistical Analysis to Infectious Diseases

Alyssa Sholeh Parpia

2022

In this dissertation, I explore the disproportionate burden of infectious disease outbreaks, epidemics, and pandemics and the projected impact of interventions for mitigating their harm on populations in Cameroon and the United States. The spatial heterogeneity in vaccination coverage and access to care in Cameroon creates areas that are highly susceptible to measles transmission. In the US, the HIV epidemic is increasingly concentrated in the southern states in addition to larger cities with variable levels of prevention and linkage to sustained treatment. Disparities in COVID-19 burden exist by race and geography across Michigan, in part due to systemic racism and underlying health burdens.

In Chapter One, I model the spatiotemporal dynamics of a large outbreak of measles in Cameroon by using several multivariate time-series models at the health district, department, and region levels. By assessing the spatiotemporal dynamics at different geographical scales, it was possible to determine the respective contribution of each administrative division to measles transmission throughout the country. The model including long-distance population mobility optimally reflected the spatial spread of measles. Population movement between departments within regions was estimated to contribute to 9.1% of all cases and movement between regions contributed to 18.1% of cases at the health district level. These findings demonstrate the need to improve our understanding of the roles of population mobility and local heterogeneity of vaccination coverage in the spread and control of measles in Cameroon.

In Chapter Two, I develop a mathematical model of HIV transmission and progression to evaluate the impact of expanding HIV prevention, diagnosis, treatment, and viral suppression levels in 57 priority counties and states in the United States, as identified by
the federal government initiative “Ending the HIV Epidemic”. This plan aims to increase access to diagnosis, linkage to treatment, maintenance of treatment and pre-exposure prophylaxis uptake in high-incidence counties as well as states with high burdens of disease in rural areas between 2020 and 2030. I project that the number of annual new infections could be reduced by 58% and that over 157,000 cumulative new infectious could be averted over the next decade nationwide upon successful implementation of this initiative. Despite the substantial benefit incurred by this HIV care continuum expansion, additional concerted efforts beyond its scope such as community-specific interventions benefiting disproportionately affected populations, stigma erasure, HIV criminalization elimination, and ending systemic oppression will be needed to truly stop HIV transmission in the US.

In Chapter Three, I examined racial disparities in COVID-19 mortality in Michigan, US, stratified by age, sex, and comorbidity prevalence. Using individual-level linked death certificate and surveillance data on all COVID-19 deaths statewide, I calculated that the mortality rate for Black populations overall was 3.6 times that of White populations, with heterogeneity across neighbourhoods. Strikingly, the mortality rate for Black individuals under 65 years lacking comorbidities was 12.6 times that of their White counterparts. Prevalence of comorbidities, age, and sex did not account for the elevated mortality rate experienced by Black individuals in Michigan. Even after accounting for demographic and underlying health characteristics, my work highlights that disparities across race resulting from systemic racism are compounded in crises.

This dissertation contributes to our understanding of the inequitable impacts of epidemics on under-resourced or historically marginalized communities within the United States and Cameroon, with analyses focused on racial and geographic disparities. Addressing the root causes of these disparities through elimination of systemic racism, improved access to care, and healthcare reform is necessary to prevent further infections and deaths. Furthermore, these changes have the capacity to reduce the impact of future infectious disease epidemics on the populations that are consistently and disproportionately affected.
Applications of Dynamic Modeling and Statistical Analysis to Infectious Diseases

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
In Candidacy for the Degree of
Doctor of Philosophy

by
Alyssa Sholeh Parpia

Dissertation Director: Alison P. Galvani

May 2022
Table of Contents

List of Tables vi
List of Figures vii
Acknowledgements ix
Dedication x
Introduction 1

Chapter One:
  Spatiotemporal dynamics of measles in Cameroon 5
  Supplement 29

Chapter Two:
  Turning the tide on HIV in the United States 36
  Supplement 53

Chapter Three:
  Racial disparities in COVID-19 mortality across Michigan, United States 123
  Supplement 150

Conclusion 158
List of Tables

Table 1.1: Average contribution (%) of transmission within a health district, department, region, and the rest of the country at the region, department, and health district level. ......................25

Table 3.1 Demographic characteristics of individuals who died from COVID-19 in Michigan overall and by race. ........................................................................................................139

Table S.3.1: Residents/Workers at High Risk or Congregate Living Facilities .........................152

Table S.3.2: COVID-19 Deaths in Michigan by age group, sex, number of comorbidities, and race from March 16 to October 26, 2020. ..............................................................................................................153

Table S.3.3. Calculating Proportion of population that has no comorbidities and 1 comorbidity. ........................................................................................................................................154

Table S.3.4: Exploration of missing data for Time to Hospitalization by race..............................156
List of Figures

Figure 1.1: Measles Incidence and vaccination coverage in Cameroon.................................21

Figure 1.2: Second-order model fitted to weekly measles cases from 2011-2012 at the region level..........................................................22

Figure 1.3: Proportion of cumulative measles cases from 2011-2012 in each department that are attributable to A) Department-level, B) Region-level, and C) Country-level transmission. Grey areas indicate departments with 50 or fewer cases over the two year period..................................................23

Figure 1.4: Proportion of cumulative measles cases from 2011-2012 in each health district that are attributable to A) Health District-level, B) Department-level, C) Region-level, and D) Country-level transmission. Grey areas indicate health districts within departments with 50 or fewer cases over the two year time-period..................................................24

Figure S.1.1: Weekly measles cases and cumulative cases across Cameroon from 2011 to 2012........................................................................33

Figure S.1.2: Weekly incident measles cases from 2011 to 2012 in Cameroon by region.........34

Figure S.1.3: Measles cases across Cameroon in 2012 at the (A) Region, (B) Department, and (C) Health District levels.................................................................35

Figure 2.1. Cumulative new HIV infections between 2020 and 2030 in priority jurisdictions (A) under maintenance of status quo and (B) implementing the Ending the HIV Epidemic (EHE) intervention..............................................................49

Figure 2.2. Cumulative new HIV infections per 10,000 capita between 2020 and 2030 in priority jurisdictions (A) under maintenance of status quo and (B) implementing the Ending the HIV Epidemic (EHE) intervention..............................................................50

Figure 2.3. Percent reduction in annual new HIV infections in 2030 compared to 2020 upon enactment of the Ending the HIV Epidemic (EHE) intervention..............................................................51

Figure 2.4. Percent reduction in annual new infections in 2030 compared to 2020 with the Ending the HIV Epidemic (EHE) and UNAIDS initiatives..............................................................52

Figure 2.5. National projections of the number of people living with HIV, incidence per 10,000 capita, and cumulative new HIV infections across the US between 2020 and 2030 upon maintaining the status quo (blue) and successfully implementing the Ending the HIV Epidemic (EHE) and UNAIDS initiatives (orange) ..............................................................52

Figure S.2.1: Compartmental model diagram..............................................................54

Figure S.2.2: Projections of the median and interquartile range of the number of people living with HIV (PLHIV), HIV incidence per 100,000 capita, and cumulative new HIV infections across the United States between 2020 and 2030 upon maintaining the status quo (blue) and achieving the Ending the HIV Epidemic (EHE) and UNAIDS goal (orange) for each jurisdiction modelled in the United States..............................................................72

Figure 3.1 COVID-19 mortality per 10,000 population in Michigan by age, sex, number of comorbidities, and race..............................................................141
Figure 3.2: COVID-19 Mortality rates per 100,000 population among Black and White Michigan residents and Michigan residents overall by ZIP Code Tabulation Area (ZCTA).............................................................................................................................................142

Figure 3.3: Mortality rate ratio for Black to White populations at the ZIP Code Tabulation Area (ZCTA) level of residence in Michigan..................................................................................................................143

Figure 3.4: Date of symptom onset and date of death over time among individuals who died from COVID-19 in Michigan by race. .........................................................................................................................144

Figure S.3.1: COVID-19 related deaths by ZIP Code Tabulation Area (ZCTA) and race........150

Figure S.3.2: Black population by Zip Code Tabulation Area (ZCTA) in Michigan.............151

Figure S.3.3: Negative Binomial distribution fit to time from symptom onset to hospitalization for Black and White decedents. ...................................................................................................................155
Acknowledgements

My deepest thanks go to Dr. Alison Galvani, my dissertation committee chair, for her unfailing encouragement throughout my studies. She is responsible for teaching me how to be a scientist, nurturing my inquisitiveness, and ensuring I keep high standards for my work. She has taught me everything from dynamic modeling to scientific writing, and most importantly she has taught me how to correctly frame and break down problems to help fill voids in our collective knowledge. Through her, I have had the opportunity to work on projects with true population-level impact and for that I will be forever grateful.

My committee members were endlessly patient with me and offered invaluable support. Dr. Daniel Weinberger took time to explain various forecasting methods as many times as was necessary and has provided numerous opportunities for my growth as a researcher. Dr. Joshua Warren made complex statistical concepts accessible to me with utmost humility and was always exceptionally generous with his time and brain. Dr. Tyler Sharp expanded my horizons by teaching me infectious disease epidemiology with immense care. The reading periods with each of my committee members during my proposal development are times I will always cherish. They truly prioritized my learning at every step of the way on this journey.

I have been extremely fortunate to learn mathematical modelling from Dr. Abhishek Pandey. He has been an irreplaceable mentor and teacher throughout my doctoral studies and to him I am eternally grateful. Dr. Martial Ndeffo-Mbah taught me to identify biases and grew my mathematical and statistical skills immensely. Dr. Jorge Alfaro-Murillo taught me to be meticulous with my methods and thoughtfully consider all possibilities. Dr. Meagan Fitzpatrick has made me a better writer and critical appraiser of literature. I am also grateful to all members of the Center for Infectious Disease Modeling and Analysis for their kindness and support over the years.

I am infinitely appreciative of Dr. Christian Tschudi, Dr. Virginia Pitzer, Dr. Robert Dubrow, and Melanie Elliot for their ever-present academic and personal support throughout my studies. I would also like to thank Dr. David Paltiel, Dr. Gregg Gonsalves, and Dr. Travis Porco, for taking the time to read and comment on my dissertation.

I have been generously supported by the Yale Climate Change and Health Initiative through a grant from the Overlook International Foundation, PEO International, the Yale MacMillan Center, the National Science Foundation, and the National Institutes of Health.

My family has supported me unconditionally throughout this ride. My mom, Rumina Jivraj, always reassured me that the stars are in reach. My dad, Zahir Parpia, instilled in me a deep desire to answer questions that no one else has. My nanima, Roshan Jivraj, ensured I prioritized education and learning at every turn. Lastly, my love, Benoit Cyrenne, encouraged me daily to just do it, and because of you, it is done.
This dissertation is dedicated to
Mom, Dad, and Nanima who have believed in me from the beginning and
Benoit, my rock, who keeps me going.
Introduction

Infectious diseases, both those that have recently emerged and those that have persisted for decades or longer, continue to spread despite advancements in vaccination and pre-exposure prophylaxis. Underlying their transmission is inequitable and inadequate provision of these resources.[1,2]

Measles

Measles, likely originating in the 11th century,[3] is among the most contagious diseases in the world. The development of a safe, effective, and inexpensive vaccine has failed to prevent measles outbreaks globally and particularly in sub-Saharan Africa. As sporadic outbreaks continue to arise, optimizing the implementation of local, regional, and national measles vaccination programs is essential for curbing spread. In 2012, Cameroon experienced a devastating measles outbreak of over 14,000 cases. In northern Cameroon, only half of the population had been vaccinated against measles, while the western regions had coverage exceeding 85%,[4] in addition to double the number per capita of health care providers.[5] These factors are likely attributable to the recurrent outbreaks arising disproportionately in the historically marginalized northern part of the country and subsequently spreading nationwide.[6–9] In Chapter One, I evaluate the drivers of measles transmission in Cameroon by analyzing the spatiotemporal trends of this outbreak.

HIV/AIDS

Historical injustices have perpetuated the 40-year HIV epidemic which today is both highly geographically and racially concentrated in the United States as a result of a failure to equitably provide access to prevention and treatment services.[10] Since the mid-1980s peak of the HIV epidemic in the US, advances in antiretroviral therapies and diagnostic testing for HIV have led to a 73% reduction in annual new infections.[11,12] However, despite reaching pre-exposure prophylaxis (PrEP) coverage among 23% of those eligible,[13] diagnosing 87% of people living with HIV, retaining 76% of those diagnosed in care, and achieving viral
suppression among 86% of those in treatment, [14] success in the HIV care continuum is highly heterogeneous nationwide. Understanding how improvements in prevention and treatment will affect jurisdictions with varying levels of viral suppression is essential for projecting the potential to turn the tide on the HIV epidemic. To accelerate progress towards HIV control in the US, the federal government has launched Ending the HIV Epidemic (EHE), a geographically focused plan that aims to improve the HIV care continuum in 57 high-incidence counties and states nationwide. [15] In Chapter Two, I evaluate the impact of successfully implementing this initiative on HIV incidence across the country.

COVID-19

The most recent pandemic, driven by SARS-CoV-2 which emerged in 2019, [16] spread rapidly across the world, magnifying health care system inadequacies and a lack of equity in our societies more broadly. The COVID-19 pandemic has had and continues to have a significantly greater impact on Black and Latinx/Hispanic individuals than White individuals as measured by infections per population in the US. In total, over 46 million infections and 750,000 deaths have occurred nationally. [17] Examining the extent to which race influences COVID-19 mortality, after adjusting for known risk factors, will emphasize the need for restructuring of our healthcare system and elimination of structural racism. In Chapter Three, I uncover whether disparities by race in COVID-19 mortality are explained by demographic characteristics and underlying health conditions alone.

The patchwork healthcare system that plagues the US is one factor at the root of the magnitude of both the HIV and COVID-19 epidemics nationally. Similarly in Cameroon, inequitable access to care and vaccination predisposes specific localities to measles susceptibility. Furthermore, systemic racism and the resulting inequities, including economic deprivation, underscore the disproportionate impact of these preventable infectious diseases. Addressing the underlying causes of preventable disease spread is essential for ameliorating disparities in their burden on already underserved communities. The following
chapters use data on disease incidence and deaths to quantify differential transmission and mortality risk as well as drivers of infectious disease spread.

References:


Chapter 1:

Spatiotemporal dynamics of measles outbreaks in Cameroon

Alyssa S. Parpia, M.P.H¹, Laura A. Skrip, Ph.D², Elaine O. Nsoesie, Ph.D³, Moise C. Ngwa, Ph.D⁴, Aristide S. Abah Abah, M.D⁵, Alison P. Galvani, Ph.D¹, Martial L. Ndeffo-Mbah, Ph.D⁶,⁷

¹ Department of Epidemiology of Microbial Disease, Yale School of Public Health, New Haven, CT, 06510, USA
² Institute for Disease Modeling, Bellevue, WA 98005, USA
³ Department of Global Health, Boston University School of Public Health, Boston, MA 02118, USA
⁴ Department of International Health, Johns Hopkins Bloomberg School of Public Health, MD, 21205, USA
⁵ Department of Epidemiological Surveillance, Ministry of Health, Yaoundé, Cameroon
⁶ Department of Veterinary and Integrative Biosciences, Texas A&M College of Veterinary Medicine and Biomedical Sciences, College Station, TX 77845, USA
⁷ Department of Epidemiology and Biostatistics, Texas A&M School of Public Health, College Station, TX 77845, USA

This manuscript has been previously published as:

Background:

Despite the existence of a safe, inexpensive, and highly efficacious vaccine that provides lifelong immunity, measles is an important infectious diseases causing an estimated 110,000 deaths worldwide in 2017 [1]. Sub-saharan Africa remains the most affected region with an estimated mortality of 4.23 deaths per 100,000 population and 355.68 disability adjusted life years per 100,000 population attributable to measles [2]. In 2011, the World Health Organization (WHO) African Region (AFR) established the goal of eliminating measles in Africa by 2020 [3]. It was initially expected that this goal could be achieved by both increasing national and health district level vaccination coverage to more than 95% and reducing measles incidence to less than 1 per 1 million population in all countries [3]. This ambitious goal appears elusive due to sporadic outbreaks across Africa. To optimize the impact of local, regional, and national measles vaccination efforts, it is paramount to improve our understanding of the spatiotemporal dynamics of measles transmission in order to inform the design of optimal targeted interventions in each endemic country.

Despite a general decline in measles incidence since the implementation of vaccination campaigns [4,5], Cameroon persistently experiences sporadic measles outbreaks [6–9], the most devastating of which was the 2012 outbreak that rapidly spread across the country to infect over 14,000 people, as reported by the integrated disease surveillance system. Vaccination coverage in Cameroon remains well below the WHO target of 95%, with only 77% coverage in 2017 at the national level [10], and substantial heterogeneity at the regional level. Vaccination coverage prior to the 2012 outbreak varied from 51.7% and 52.4% in the North and Far North regions to 85.1% and 93.7% in Littoral and the North West regions, respectively [11]. This spatial heterogeneity of vaccination coverage across the country is a likely culprit for the recurrent measles outbreaks in Cameroon. Heterogeneity of vaccination coverage may be conducive to source-sink type dynamics [12] whereby pockets of the country are more susceptible to new infections. The potential of these sources to
generate and sustain local or national outbreaks depends on spatio-temporal trends in disease spread.

Recurrent measles outbreaks arise disproportionately in northern Cameroon[8,13–15], throughout which poverty is widespread. The Far North and North regions have lower access to healthcare than the rest of the country, with half as many nurses and doctors per capita (0.53 per 1000) than the region where the nation's capital is located (1.10 per 1000 population in Center)[16]. Furthermore, these regions have over half their population in the lowest national economic quintile (54.8% and 51.7%, respectively), with female illiteracy rates (61.0% and 50.4%, respectively) double the national average (25.5%)[11]. The proximity of the Far North region to the southern Chadian and northern Nigerian borders, terrorized by recent activities of the Boko Haram militant group, has made it difficult to implement vaccination campaigns. This ongoing security issue exacerbates the risk of disease outbreaks in northern Cameroon which was already precariously positioned.

To evaluate the drivers of measles transmission and identify hotspots for disease spread in Cameroon, we analyze the spatiotemporal trends that characterized the 2011-2012 measles outbreak, the most devastating in recent years. In particular, we analysed the spatio-temporal dynamics of the outbreak at multiple geographical scales to determine the respective contribution of each scale to measles transmission.

Methods:

Data Sources:
The Cameroon Ministry of Health mandates measles reporting with case follow-up as part of a passive surveillance system[17] operated by the Epidemiology Service in the National Disease Control Department. Measles case and death data are collected through health district-level passive surveillance conducted by clinics, health centers, and hospitals, which are then reported to the Ministry of Health. We analyzed surveillance data from the Ministry of Health on weekly measles cases and deaths reported in 2011 and 2012, from each of the 183 health districts in Cameroon. Health districts are the smallest geographic scale for
healthcare services and decision-making in Cameroon. To evaluate the robustness of our analysis at larger geographic scales, we also aggregated data at the departmental and regional levels. There are 60 departments and 10 regions across the country. Data on measles vaccination coverage at the region level were obtained from the 2011 Cameroon Demographic and Health Survey (Enquête Démographique et de Santé et à Indicateurs Multiples: EDS-MICS) [11].

Spatial mapping of measles at the region, department, and health district levels was conducted using R, with a base map of health districts obtained from previous work on the spatial mapping of infectious diseases in Cameroon [18].

**Case Definition:**

Suspected measles cases are defined as an illness characterized by a generalized maculopapular rash and fever, as well as one or more of the following symptoms: cough, coryza, and conjunctivitis; or as any case suspected to be measles by a clinician. For suspected measles cases identified within 30 days after the onset of symptoms, a laboratory confirmation test for measles-specific immunoglobulin (Ig) M antibody is requested by the Ministry of Health. Measles case confirmation is either defined by a positive result for measles IgM antibody testing or by epidemiological linkage. The data available from the Disease Control Department of the Ministry of Health were suspected cases. No information with which to classify them as laboratory confirmed or epidemiologically linked cases was available. These data were collected through the integrated disease surveillance system rather than the measles case-based reporting system.

**Analysis:**

Spatial autocorrelation of measles incidence across health districts was assessed by calculating Moran’s *I* statistic using Monte Carlo simulation (n=1000). While Moran’s *I* statistic provides a measure of spatial autocorrelation across all health districts, local Moran’s *I* statistics were calculated to assess spatial autocorrelation for each health district.
To analyse the spatio-temporal trends of the 2011-2012 measles outbreak, we fitted a range of multivariate time-series models to identify which model most effectively describes the data, and used this model to assess the spatio-temporal dynamics of the outbreak and evaluate the contribution of population mobility to disease transmission. The analysis was successively conducted at the region, department, and health district levels. The R package “surveillance” was used for development of the multivariate time-series model [19–21].

We generated multivariate time-series model for measles case counts, $Y_{i,t}$, in geographic localities ($i = 1, ..., I$) during weekly time periods ($t = 1, ..., T$) for a total of two years [22,23]. The model accounts for (1) the contribution of local transmission to disease outbreak and (2) the spatial spread of disease from individuals’ movement between localities. This model was used to assess the spatio-temporal trends of the 2011-2012 measles outbreak in Cameroon at spatial scales of decreasing size: region ($i_r = 1$ to 10), department ($i_d = 1$ to 60), and health district ($i_h = 1$ to 183).

Conditional on past observations, $Y_a$ is assumed to have a negative binomial distribution with mean incidence $\mu_a$ and variance of $\sigma_a = \mu_a(1+\psi_\mu \mu_a)$, with an overdispersion parameter $\psi > 0$:

$$\mu_a = \lambda_a Y_{1,t-1} + \varphi_{1,a} \sum w_{ji} Y_{j,t-1} \text{ with } j \neq i.$$  

The mean incidence was decomposed into a within-locality ($\lambda_a Y_{1,t-1}$) component (i.e., the generation of new measles cases from cases within locality $i$), and a spatial spread (i.e., between-locality interaction through population movement) component ($\varphi_{1,a} \sum w_{ji} Y_{j,t-1}$) which accounts for the contribution of another locality, $j$, to disease transmission in locality $i$. $Y_{j,t-1}$ indicates the number of cases observed in locality $j$ at time $t-1$. The spatial component includes weights ($w_{ji}$) which reflect the flow of infections from locality $j$ to locality $i$, and a within-locality parameter, $\varphi_{1,a}$, which captures the contribution of $Y_{j,t-1}$, where $j \neq i$, on $Y_{1,t}$. The epidemic components indicate a strict dependence between events driven by the observed past cases, and captures occasional outbreaks in the time-series data.
We used several models to capture the flow of infections between areas: first-order (nearest neighbour interaction), second-order, gravity, and power law models. Each model represents a different scenario for population movement between localities. An adjacency order matrix was generated to depict whether or not localities are neighbours based on sharing a common border, which was used to determine the weights. At each geographic level, we identified the type of model that exhibited the optimal fit to measles data.

The first order model assumes that spatial spread of disease may only arise as a result of cases in directly adjacent localities. Here, $w_{ji} = 1$ if $j$ and $i$ are adjacent and 0 otherwise. The second order model includes weights that decay with distance for first and second-order neighbours only, eliminating the possibility of transmission of cases from localities more than 2 neighbours away. $w_{ji} = 1 \cdot 1(o_{ji} = 1) + e^{\omega_2} \cdot 1(o_{ji} = 2)$, where $1$ is the indicator function and $o_{ji}$ is the adjacent order between localities $j$ and $i$. The spatial gravity model involves having attraction to a locality scale with its population size to reflect commuter-driven disease spread. This model is developed by multiplying the ability of a locality to import cases from neighbouring localities, i.e. susceptibility ($\phi$), by the population of the given locality ($e_i \beta_{pop}$). This model accounts for the idea that humans tend to travel further and preferentially to densely populated metropolises. The power law model accounts for potential long-distance transmission events between all localities. Weights ($w_{ji}$) are defined as a function of the adjacency order ($o_{ji}$) between localities ($w_{ji} = o_{ji}^{-\delta}$), and are normalized such that $\sum w_{ji} = 1$.

At each level of geographic analysis, Akaike information criterion (AIC), an estimator of the relative quality of a model based on the maximum likelihood, was calculated for each model. The model with the lowest AIC, which indicates a combination of optimal goodness of fit with discouragement for overfitting, was identified at the region, department, and health district levels. Correlation coefficients between data and the model with the lowest AIC were calculated for all geographic levels of analysis in order to assess goodness of model fit to the data. At the health district level, the proportion of cases attributable to disease transmission within the district, within the department in which the district is located, within the region in which the district is located, and within the rest of the country (excluding the
Results:

Descriptive analysis

A total of 14,806 measles cases and 73 measles-attributable deaths were reported during 2012 in Cameroon. The vast majority (97.4%) of cases were reported within the first six months of the year, with a peak in cases during week 12 (Figure S.1.1). The North region reported the highest measles incidence, with 182.5 cases per 100,000 population, and the South West region had the lowest at 17.7 cases per 100,000 population. Health districts with the lowest cumulative measles incidence per 100,000 population (Figure 1.1A) are located in the Western (West, North West, South West, and Littoral) regions, which are correspondingly the regions with highest measles vaccination coverage (Figure 1.1B). Regions with low vaccination coverage have highly heterogeneous measles incidence rates (Figure 1.1A). The Adamaoua region in particular, which has a relatively low overall measles vaccination coverage of 64.0%, has starkly different cumulative incidences of measles across its health districts, ranging from 3.1 to 103.3 cases per 100,000 population. The East (74.4%) and South (69.5%) regions also have moderate to low vaccination coverages overall, with highly variable cumulative incidence in 2012 by health district.

The Moran’s I coefficient (index = 0.206, p = 0.001) suggests that the measles outbreak in Cameroon was characterized by statistically significant spatial autocorrelation. This indicates spatial clustering of cumulative measles incidence across health districts. Upon calculation of local Moran’s I statistics for each health district, we identified two statistically significant clusters of measles incidence in districts within the North (Bibemi, Garoua I, Garoua II, Lagdo, Ngong, and Rey Bouba) and Far North (Bogo and Vele) regions.
**Spatio-temporal analysis**

**Region Level**

At the region level, the second-order model had the lowest AIC. Compared to the first-order (nearest neighbour) model, the gravity model, and the power law model, the second-order model had an AIC that was 80.6, 17.4, and 0.42 lower, respectively. The small difference between the AIC of the second-order model and the power law model suggests that there is substantial support for both models. Therefore, we could not preferentially choose one of the models over the other. At this geographic scale, a second-order neighbour model is synonymous with a long-distance movement model given that there are on average three degrees of separation (sharing the same border) between regions. Correlation between the reported cases and the second-order neighbour model exceeded 0.5 for all regions and exceeded 0.75 for 80% of regions, indicating that the model provides a good fit to the data. The model showed that the reported measles cases in each region were predominantly the result of within-region transmission with very little contribution from other regions (Figure 1.2). Measles transmission between regions had its highest contribution to measles outbreaks in the Littoral and North West regions, 36.2% and 17.8% of cases, respectively (Table 1.1). On average, 9.1% of cases were attributable to transmission from outside of the region (Table 1.1).

**Department Level**

At the department level, we only considered the 38 departments that cumulatively reported more than 50 cases from 2011-2012. This restriction was made to minimize the impact of underreporting in our analysis. At this geographic scale, the power law model provided the best fit to the data. Compared to the first-order model, the gravity model, and the second-order model, the power law model had an AIC that was 31.07, 32.38, and 17.34 less, respectively. Correlation between the data and the power law model exceeded 0.5 for 84.6% of departments and exceeded 0.7 for 41.0% of departments. Our analysis showed that across
the country, disease transmission was primarily driven by within-department transmission (Figure 1.3A). On average, within-department transmission contributed to 76.3% of cases, within-region transmission contributed to 11.5% of cases, and transmission from other regions contributed to 12.2% of cases.

Transmission within-departments predominantly contributed to cases in the North (average: 82.9%, range: 50.0% in Mayo-Louti to 96.2% in Benoue), Littoral (82.3%, 32.6% in Mouno to 96.1% in Wouri), and Far North (76.0%, 47.0% in Mayo-Kani to 89.3% in Diamare) regions (Figure 1.3A and Table 1.1). Transmission between departments in the same region had the greatest impact in the Far North and Adamaua regions where it contributed to 18.3% and 16.4% of cases, respectively (Table 1.1). In the Far North region, transmission between departments contributed to over 30% of cases in the Logon-Chari and Mayo-Kani departments (Figure 1.3B). In Adamaua, it contributed to over 20% of cases in the Djerem and Faro-et-Deo departments (Figure 1.3B). Transmission from other regions had the greatest impact in departments of the South and South West regions (Figure 1.3C and Table 1.1).

**Health District Level**

At the health district level, the power law model was also shown have the best fit to the data. A total of 141 health districts are situated within the 38 departments reporting more than 50 cases from 2011-2012. In comparison to the first-order model, the gravity model accounting for commuter-driven travel, and the second-order model, the power law model at the health district level had an AIC that was 247.1, 136.46, and 69.12 less, respectively. Among the 92 health districts that reported over 20 cases from 2011-2012, correlation between the data and the power law model results exceeded 0.5 in 67.4% of health districts and exceeded 0.7 in 25.0% of health districts. As observed at the region (Figure 1.2) and department levels (Figure 1.3), measles cases were also predominantly explained by within-district transmission at the health district level (Figure 1.4A). However, the contribution of between-district interaction in driving cases through population movement was higher at the
health district level (Figure 1.4B–D), than in the department level of analysis (Figure 1.3). On average, transmission within districts contributes to 66.8% of cases in health districts, transmission between health districts in the same department contributes to 6.0% of cases, transmission between health districts in the same region contributes to 9.1% of cases, and other regions contribute to 18.1% of cases (Table 1.1).

Transmission between districts within the same department had its greatest contribution to disease transmission in the South-West, with on average 10.1% (6.8% in Limbe to 27.8% in Konye) of cases attributable to population movement within a department (Figure 1.4B and Table 1.1). Transmission between districts within a region had its greatest impact on measles cases in the Far North region (Figure 1.4C and Table 1.1). It contributed to more than 27% of cases in 17 of the 29 health districts in the region, including Tokombe and Meri, where within-region transmission contributed to 49.1% and 48.9% of cases, respectively (Figure 1.4C). Transmission from other regions contributed on average to 31.3% of cases in health districts in the South region (varying from 15.0% in Kribi to 75.8% in Sangmelima) and to 29.9% in the East region (varying from 10.9% in Batouri to 90.4% in Moloundou) (Figure 1.4D and Table 1.1).

Population movement across localities plays a greater role in disease transmission at the health district level than the department and region level. Between-district transmission contributed to more than 50% of cases in 45.4% of health districts, compared to 18.4% of departments in which between-department transmission contributed to more than 50% of measles cases. Across the country, between-locality transmission contributed to 9.1%, 23.7%, and 33.2% of measles cases at the region, department, and health district levels, respectively (Table 1.1). At the region level, the proportion of cases attributable to transmission from outside a given region ranged from 0.1% in the South to 36.2% in Littoral (Table 1.1). Other regions contributed to 12.2% of cases at the department level on average, ranging from 5.6% of cases in the Far North to 29.7% of cases in the South (Table 1.1). Finally, at the health district level, other regions excluding the region in which the health district is located
contributed to 18.1% of cases on average, ranging from 8.9% of cases in the Far North to 31.3% of cases in the South (Table 1.1).

Overall, the majority of cases in areas reporting large outbreaks (Figure S.1.3) were driven by within-area transmission. In departments with over 200 cases, an average of 72.3% of cases were due to within-department transmission, and in health districts with over 100 cases, 66.6% of cases on average were due to within-district transmission.

**Discussion:**

The power law model accounting for long-distance population travel was found to have the best fit at the health district and department levels. The second order model, which at the region level also inherently considers long-distance travel, was the best fit at the region level. These models showed that local transmission was the main mechanism for disease transmission at the region, department, and health district levels. As expected, the contribution of between-locality interaction through population movement was shown to increase as the geographic scale was refined; however, the Littoral and Adamaoua regions are exceptions to this trend, where the contribution of population mobility to disease transmission appears to be independent of spatial scale. On average, 76.3% (range: 29.8–97.4%) of cases in departments were attributable to within-departments transmission in contrast to 66.8% (range: 0.0–94.1%) of cases in health districts that were attributable to within-health district transmission. Interaction between departments contributed to 23.7% (range: 2.6–70.2%) of cases in departments and between-district interaction contributed to 33.2% (range: 5.9–100.0%) of cases in health districts, on average. These results indicate that population mobility between localities is an important risk factor for large scale measles outbreaks in Cameroon.

As Littoral and Center contain the two most populated cities in the country, the economic and political capitals, Douala and Yaoundé, we find that substantial travel in and out of these cities, and in turn the regions, might be driving the measles outbreak far more so than in other regions in Cameroon, and that further work to elucidate the reasons behind this
are necessary. Littoral in particular had a high vaccination coverage (85.1%) in 2011 compared to the country as a whole (70.6%), yet had a moderate regional measles incidence rate in 2012 (44.3 per 100,000 population) compared to the rest of the country (17.7 in the South to 182.5 per 100,000 population in the North). Our findings of the substantial contribution (17.7 to 37.6%) of between-locality transmission in driving measles cases in Littoral indicates the importance of maintaining high vaccination coverage country-wide in preventing measles cases in this highly traveled region. We also found this to be true in departments and health districts reporting large outbreaks (Figure S.1.3). Our results agree with previous modeling studies that demonstrate that the epidemic trajectory of large-scale measles outbreaks in a community is virtually unaffected by immigrant infection [24,25]. These studies showed that epidemic dynamics were predominantly driven by within-locality transmission [24,25].

In Cameroon, within-locality movement is driven by numerous socio-economic factors including trade, farming, education, and family ties among others [26,27]. In some rural areas, children must travel regularly between their health district of residence and the nearest high school. Thus, schools may become hubs for the spread of diseases between the neighbouring health districts. The need to travel to regional and weekly markets also contributes significantly to rural-to-rural and urban-to-rural population movements within departments and regions [26]. Population movement for family-related reasons, such as visiting one’s locality of origin or relatives in other parts of the country, are a major cause of the movement of children during school breaks in Cameroon. This population mobility compounds the low vaccination coverage, especially in the northern regions, to exacerbate the risk of measles outbreaks in Cameroon. Low vaccination coverage is driven by a combination of poverty, inaccessibility to healthcare services, parents’ level of education, religious beliefs, and parents attitudes towards vaccination [28,29]. While our study focuses on the impact of population mobility within national borders on measles outbreak dynamics, cross-border movement is also an important instigator of the spread of measles [30]. Although the impact of cross-border movement was not assessed in our model, the current security crisis in the
northern regions of Cameroon has increased the risk of measles outbreak in the country and its subsequent spread to Nigeria and Chad through mass migration.

Our descriptive analysis demonstrates that the measles outbreak started in Adamaoua and the North, where the population is relatively dispersed, before moving to more densely populated regions near the national capital, Yaoundé, and the economic capital, Douala. This is consistent with northern regions being hotspots for measles outbreaks due to their lower vaccination coverage, health care providers to inhabitants ratio, income, and female education levels, than other regions. Moreover, the start of the 2012 outbreak, during late 2011 and early 2012, coincides with the Christmas to New Year holiday season which is marked by high population movement between urban, higher vaccination coverage areas and rural, lower vaccination coverage areas. While vaccination coverage is greater around the large cities in the western part of the country than in northern Cameroon, the higher population density in Yaoundé and Douala places more people at risk of measles infection, furthering the spread of the epidemic. Heterogeneity of vaccination coverage creates environments for outbreaks to take off [31] and weakens impact of herd immunity [32].

Supplementary immunization activities (SIAs) have been conducted regularly in Cameroon in order to improve measles vaccination coverage. However, some of these SIAs have been localized and have historically missed key areas in Cameroon with particularly low vaccination coverage [33]. Before the 2012 measles outbreak, catch-up SIAs for children 9 to 59 months old or follow-up SIAs targeting children born since the most recent SIA had been conducted in Cameroon in 2001, 2002, 2006, 2007, 2009, and 2012 [6,34,35] The SIA in 2012 reached over 3.5 million children nationally such that 78% of targeted districts had ≥95% coverage [6]. However, SIA were initiated in April well after the start of the outbreak and did not target children aged 6 to 14 years old who have a low measles vaccination coverage and are at high risk of infection [6,36,37].

In Cameroon, population movement tends to occur in the direction of city centers [38], which informed our consideration of gravity models for fitting the epidemic trajectory. However, our findings have indicated that pull of individuals towards high-density areas,
accounted for in the gravity model, does not sufficiently explain this measles outbreak. Accounting for long-distance transmission was shown to better explain the spread of measles in Cameroon at all geographic levels of analysis. Particularly as the resolution of data increases from the region to health district level, population movement was shown to play an increasingly important role in disease transmission. In the absence of population mobility data, the power law model provides a useful way to measure the influence of long-distance travel on disease spread [39].

A study on cases of measles in Cameroon in 2000-2001 identified distinct patterns of measles in the three northern regions compared to the seven southern ones, with annual major epidemics in the north and major epidemics only occurring in the south every 3 years. [13] This study identified that higher cumulative region-specific incidence rates were associated with higher birth and lower routine vaccination rates [13]. The Benakuma health district, located in North West Cameroon, experienced a measles outbreak in 2015 and was identified as a hotspot for future measles outbreaks due to a combination of poor vaccination levels, low socio-economic status overall, and environmental factors limiting ease of vaccination [40]. We identified that 41.6% of cases in the 2011-2012 outbreak in Benakuma were attributable to between-district interaction. These findings indicate the importance of identifying and then improving health district-specific vaccination coverage for curtailing future measles outbreaks.

The measles surveillance system in Cameroon is mainly based on passive case diagnosis and reporting by local health facilities. Given inadequate infrastructure, the low number of health facilities, and the low ratio of health care providers to inhabitants in many health districts, case diagnosis and timely reporting is challenging in many parts of the country. This situation has surely resulted in underreporting of measles cases in some health districts. For example, in the Far North region, which is especially vulnerable to infectious disease outbreaks given low vaccination coverage and conflict, the reports of very low annual measles incidence (<2 cases per 100,000 population) in some health districts neighbouring high incidence districts brings into question the ability of such districts to accurately diagnose
and report cases. The potentially high level of underreporting of measles cases in some Far North health districts is likely to affect our results on the relative contribution of the different transmission routes on disease outbreaks not only within the health districts but also in some of their neighbouring districts.

Overall, our analysis shows that considering the role of population movement from higher order neighbours, in addition to directly neighbouring areas, is essential in understanding transmission dynamics of measles in Cameroon. On average 23.7 to 33.2% of cases occurring at the department or district level originated from population movement outside of the residential locality. This indicates that substantial health benefits of improving measles vaccination within a particular district or department are likely to be realized beyond that locality’s borders. Improving vaccination coverage in rural and high risk transmission areas such as the North and Far North regions of Cameroon would not only benefit these regions but would also provide benefits to urban areas and the rest of the country.

Conclusions:

This study found that the 2011-2012 measles outbreak in Cameroon was driven by a combination of local and long-distance transmission factors. The contribution of population movement to disease transmission was shown to be highly heterogeneous across the country. Improving our understanding of vaccination coverage at health district and department levels will be essential in mitigating measles cases originating due to long-distance movement of populations across artificial, administrative boundaries.

Authors' contributions: ASP and MLNM designed the study; ASAA acquired the data; ASP performed the analysis; ASP, EON, and MLNM interpreted the data. MCN contributed the health district level shapefile. ASP, LS, APG, and MLNM drafted the manuscript. All authors read and provided edits on the final paper.
Acknowledgements: The authors would like to acknowledge the provision of data from the Cameroon Ministry of Health.

Funding: This study was funded by the National Institutes of Health. Martial L. Ndeffo-Mbah and Alison P. Galvani are supported by NIH grant U01 GM087719. Elaine O. Nsoesie is supported by grant K01ES025438-04.
Figures and Tables:

**Figure 1.1:** Measles Incidence and vaccination coverage in Cameroon.

(A) Cumulative incidence of measles per 100,000 population in Cameroon by health district in 2012. Red indicates a higher incidence and blue indicates a lower incidence of measles, with colour divisions by quantiles. Yaoundé (star), the capital, and Douala (circle), the economic capital, are the most populated cities in Cameroon. (B) Measles Vaccination Coverage (%) in Cameroon by region in 2011 [11]. Red indicates lower vaccination coverage and blue indicates higher vaccination coverage.
Figure 1.2: Second-order model fitted to weekly measles cases from 2011-2012 at the region level.

The model-generated numbers of cases estimated to be driven by interaction with first and second-order neighbouring regions are represented in orange and the numbers of cases attributable transmission within each region are in purple. Black dots represent weekly measles cases at the region level as reported by the Ministry of Health.
Figure 1.3: Proportion of cumulative measles cases from 2011-2012 in each department that are attributable to A) Department-level, B) Region-level, and C) Country-level transmission. Grey areas indicate departments with 50 or fewer cases over the two-year period.
Figure 1.4: Proportion of cumulative measles cases from 2011-2012 in each health district that are attributable to A) Health District-level, B) Department-level, C) Region-level, and D) Country-level transmission. Grey areas indicate health districts within departments with 50 or fewer cases over the two year time-period.
Table 1.1: Average contribution (%) of transmission within a health district, department, region, and the rest of the country at the region, department, and health district level.

<table>
<thead>
<tr>
<th>Region</th>
<th>Region-level (%)</th>
<th>Department-level (%)</th>
<th>Health District-level (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Region</td>
<td>Department</td>
<td>Region</td>
</tr>
<tr>
<td>Adamaoua</td>
<td>99.7</td>
<td>65.5</td>
<td>16.4</td>
</tr>
<tr>
<td>Center</td>
<td>87.0</td>
<td>73.0</td>
<td>14.1</td>
</tr>
<tr>
<td>East</td>
<td>99.6</td>
<td>76.4</td>
<td>11.8</td>
</tr>
<tr>
<td>Far North</td>
<td>86.9</td>
<td>76.0</td>
<td>18.3</td>
</tr>
<tr>
<td>Littoral</td>
<td>63.8</td>
<td>82.3</td>
<td>4.5</td>
</tr>
<tr>
<td>North</td>
<td>98.1</td>
<td>82.9</td>
<td>4.8</td>
</tr>
<tr>
<td>North West</td>
<td>82.2</td>
<td>67.8</td>
<td>6.1</td>
</tr>
<tr>
<td>West</td>
<td>96.6</td>
<td>73.0</td>
<td>9.0</td>
</tr>
<tr>
<td>South</td>
<td>99.9</td>
<td>69.0</td>
<td>1.3</td>
</tr>
<tr>
<td>South West</td>
<td>95.3</td>
<td>60.2</td>
<td>11.5</td>
</tr>
<tr>
<td>Cameroon</td>
<td>90.9</td>
<td>76.3</td>
<td>11.5</td>
</tr>
</tbody>
</table>

The economic capital, Douala, is located in Littoral and the political capital, Yaoundé, is located in the Center region.
References:


29. Wiysonge CS, Uthman OA, Ndumbe PM, Hussey GD. Individual and contextual factors


Chapter 1 Supplement:

Methods

We generated an endemic-epidemic multivariate time-series model for measles case
counts, \( Y_{i,t} \), in geographic localities \( (i = 1, \ldots, I) \) during weekly time periods \( (t = 1, \ldots, T) \) for a total of two years[23,24]. This model was used to assess the spatio-temporal
trends of the 2011-2012 measles outbreak in Cameroon at spatial scales of decreasing size:
region \( (i_r = 1, \ldots, 10) \), department \( (i_d = 1, \ldots, 60) \), and health district \( (i_h = 1, \ldots, 183) \)
geographic area levels.

Our model accounts for within-locality transmission as well as transmission from
neighbouring localities. Conditional on past observations, \( Y_{it} \), the time series of measles case
counts, is assumed to have a negative binomial distribution with mean incidence:

\[
\mu_{it} = \lambda_{it} Y_{i,t-1} + \phi_{it} \sum_{j=1,j \neq i}^{I} w_{ij} Y_{j,t-1} + e_{it} v_{it}
\]

- \( \mu_{it} \): Mean measles incidence
- \( \lambda_{it} Y_{i,t-1} \): Within-area components for capturing occasional outbreaks.
- \( \phi_{it} \): Between-area driven effects, the influence of area \( j \) on area \( i \)
- \( w_{ij} \): Weight matrix reflecting flow of infections from area \( j \) on area \( i \)
- \( Y_{j,t-1} \): Number of cases observed in area \( j \) at time \( t-1 \)
- \( v_{it} \): Endemic component which represents baseline cases and is temporally stable
- \( e_{it} \): Population fraction offset in each area \( i \) at time \( t \)

Past cases in other neighbouring areas are explanatory covariates in the model. \( Y_{j,t-1} \)
is the number of cases observed in area \( j \) at time \( t - 1 \), and \( w_{ij} \) are weights reflecting flow of
infections from area \( j \) to area \( i \). The mean incidence, \( \mu_{it} \), decomposes additively into a within-area
transmission \( (\lambda_{it} Y_{i,t-1}) \) component (i.e., the reproduction of measles within area \( i \)) and a
between-area transmission parameter \( \phi_{it} \) affects the influence of \( Y_{j,t-1}, j \neq i \), on \( Y_{i,t} \) (i.e., the
transmission of measles from other areas \( j \) to area \( i \)). The conditional variance of \( Y_{it} \) is
\( \mu_{it} (1 + \psi_i \mu_{it}) \), with an overdispersion parameter \( \psi_i > 0 \).
The unit of area, $i$, refers to the region ($i_r$), department ($i_d$), or health district ($i_n$), depending on which geographic level of analysis is being considered, and the fraction of population living in area $i$ is used as the endemic offset. In the endemic components, risk of new events are driven by external factors independent of the history of the epidemic process such as seasonality, population density, and vaccine coverage. This component of the model explains a baseline rate of cases with a stable temporal pattern.

The endemic component, $v_{it}$, represents the baseline cases and is stable over time:

$$
\log(v_{it}) = \alpha_{i}^{(v)} + \beta_{i}^{(v)} t + \left\{ \sum_{s=1}^{S} y_{ls} \sin(\omega_{ls} t) + \delta_{ls} \cos(\omega_{ls} t) \right\}
$$

- $\alpha_{i}$: Intercept
- $\beta_{i}$: trend parameters
- $y_{ls}, \delta_{ls}$: unknown parameters for modeling seasonal variation
- $\omega_{ls}$: $2\pi s / 52$, Fourier frequencies
- $S$: number of terms in the truncated Fourier Series

The other two log-linear components are the within-area ($\lambda_{it}$) and between-area driven ($\phi_{it}$) effects [44]:

$$
\log(\lambda_{it}) = \alpha_{i}^{(\lambda)} + \beta_{i}^{(\lambda)^T} z_{it}^{(\lambda)},
\log(\phi_{it}) = \alpha_{i}^{(\phi)} + \beta_{i}^{(\phi)^T} z_{it}^{(\phi)}
$$

- $\alpha_{i}$: Intercept
- $\beta_{i}$: Regression parameters. Area-level effects which account for heterogeneity between areas, such as vaccination coverage, number of health centers per locality, and the proportion of received measles case reports out of the total expected per locality from the national surveillance system.
- $z_{it}$: Exogenous covariate vector of time effects

Intercepts of all predictors are assumed to be identical across localities.

An $I \times I$ adjacency order matrix was generated to depict whether areas are neighbours based on sharing a common border. This matrix which was used in determining between-area interaction weights. At the regional, departmental, and health district levels, we generated first-order, second-order, gravity, and power law, models and identified the type of model at each resolution of measles data that demonstrated the best fit. Each of these four
models were fit with covariates for measles vaccination coverage, the number of health centers per population, and the proportion of received reports on weekly measles cases among those expected by the national surveillance system were incorporated into the auto-regressive component of all models. Vaccination coverage and the number of health centers per population vary by locality, while the proportion of received reports varies by locality and week in the year. For each of the four models detailed below, model selection using Akaike’s Information Criterion was used to determine the model and covariates that provide the best fit to the data.

First-order Model

The first-order model involves terms for endemic log-linear transmission \( v_{it} \) with the multiplicative effect of seasonality, between-area driven effects \( \phi_{it} \), and within-area transmission effects \( \lambda_{it} \). The endemic component, \( v_{it} \), represents the baseline level of cases following a stable temporal pattern. This model assumes epidemics can only arise in a given area as a result of cases in directly adjacent areas. In this model, developed for all geographic levels of analysis, all areas have the same coefficient for case importation from directly neighbouring areas (same susceptibility). Therefore, the sum of all case counts in adjacent areas enters the model as an explanatory variable. Conditional on past observations, \( Y_{it} \) is assumed to have a negative binomial distribution. The within-area epidemic parameter \( \lambda_{it} = \exp(\alpha_i^{(I)}) \) and between-area driven epidemic parameter \( \phi_{it} = \exp(\alpha_i^{(\phi)}) \) are both homogenous across areas and remain constant over time. Here, the weights, \( w_{ji} \), are such that the epidemic can only arise as a result of cases in directly adjacent districts \( w_{ji} = 1(j \sim i) = 1(\alpha_{ji} = 1) \). First-order (or nearest neighbour) models have been used previously in network modeling to simulate the links between individuals [23,45].

Second-order model

The second-order model includes weights that decay with distance for first and second-order neighbours only, eliminating the possibility of transmission of cases from areas
more than two neighbours away. \( w_{ji} = 1 \cdot I(o_{ji} = 1) + e^{o_{ji}} \cdot I(o_{ji} = 2) \), where \( I \) is the indicator function. Therefore, travel between adjacent areas and neighbours of neighbours are considered. Second order models have been used to represent the spread of dengue epidemics between cities [46].

Spatial-interaction Gravity Model

The spatial interaction gravity model involves having attraction to an area scale with population size to reflect commuter-driven disease spread. This model is developed by multiplying the area’s ability to import cases from neighbouring areas, i.e. susceptibility \( (\phi_{it}) \), by the fraction of the total population of the given area \( (\epsilon_{it} \cdot \text{prop}) \) where \( w_{ji} = I(j \sim i) \).

This model accounts for the idea that humans tend to travel further and preferably to metropolitan, densely populated areas, and was developed for all geographic areas of analysis. The gravity model has been used in studies of human migration and mobility in relation to transmission of infectious diseases in Sub-Saharan Africa [47–49].

Power law model

The power law model accounts for potential transmission of cases as a result of long-distance travel events and involves estimating a parameter for the decay in epidemiological coupling as distance between areas increases. Weights \( (w_{ji}) \) are estimated as a function of the adjacency order \( (o_{ji}) \) between areas \( w_{ji} = o_{ji}^{-d}, \) for \( j \neq i \) and \( w_{jj} = 0 \), where \( d \) is decay parameter which represents the deterioration in the impact of locality \( j \) on cases in locality \( i \) as the number of localities between the two areas increases, and are normalized such that \( \Sigma_j w_{ji} = 1 \). This model was developed for all geographic areas of analysis. The power-law model has been used previously to model human travel [50].
Figure S.1.1: Weekly measles cases and cumulative cases across Cameroon from 2011 to 2012.
Figure S.1.2: Weekly incident measles cases from 2011 to 2012 in Cameroon by region.

The vast majority of all cases during this time period originated from the North and Far North regions.
Figure S.1.3: Measles cases across Cameroon in 2012 at the (A) Region, (B) Department, and (C) Health District levels.
Chapter 2:

Turning the Tide on the HIV Epidemic in the United States

Alyssa S. Parpia, MPH\textsuperscript{1}, Abhishek Pandey, PhD\textsuperscript{1}, Meagan C. Fitzpatrick, PhD\textsuperscript{1,2}, Brandon Marshall, PhD\textsuperscript{3}, William Goedel, PhD\textsuperscript{3}, Burton H. Singer, PhD\textsuperscript{4}, Alison P. Galvani, PhD\textsuperscript{1}

\textsuperscript{1} Center for Infectious Disease Modeling and Analysis (CIDMA), Yale School of Public Health, 135 College St., Suite 200, New Haven, CT 06510
\textsuperscript{2} Center for Vaccine Development and Global Health, University of Maryland School of Medicine, 685 W Baltimore St, Baltimore, MD 21201
\textsuperscript{3} Department of Epidemiology, School of Public Health, Brown University, 121 South Main Street, 2nd Floor, Providence, RI 02912, USA
\textsuperscript{4} Emerging Pathogens Institute, University of Florida, P.O. Box 100009, Gainesville, FL 32610
Introduction

Since HIV/AIDS was first reported, the US has reported more annual new HIV infections than any other high-income country.\textsuperscript{1,2} Currently, there are 1.2 million people living with HIV (PLHIV) in the US, of whom 87\% are aware of their status.\textsuperscript{3} Among those diagnosed, 76\% are in care, and of those being treated, 86\% are virally suppressed (87–76–86).\textsuperscript{4} Viral suppression not only greatly improves longevity, but nearly eliminates the possibility of HIV transmission. Conversely, the remaining 43\% of PLHIV whose infections are not virally suppressed may continue to transmit HIV. To curb HIV transmission on a population level, public health initiatives to increase testing, treatment, and engagement in care are urgently needed. Consequently, the federal government initiated “Ending the HIV Epidemic: A Plan for America” (EHE) to improve access to HIV prevention, diagnosis, and treatment services with a focus on high-incidence jurisdictions.\textsuperscript{5}

The Joint United Nations Programme on HIV/AIDS (UNAIDS) goal for HIV care continuum improvement is to reach 95\% diagnosed, 95\% treated among diagnosed, and 95\% virally suppressed among treated by 2030 (95–95–95 by 2030).\textsuperscript{6} The federal EHE plan involves accelerating achievement of the diagnosis and treatment components of the UNAIDS goal by five years. Concomitantly, the EHE aims for viral suppression of all treated individuals (95–95–100 by 2025), thereby achieving viral suppression among 90.3\% of PLHIV across targeted jurisdictions. Additionally, by 2025 the plan aims to prescribe post-exposure prophylaxis (PrEP) to at least 50\% of those eligible,\textsuperscript{7–10} a major expansion given that current coverage is only 23\%.\textsuperscript{11} The priority jurisdictions, including 48 counties as well as Washington, DC and San Juan Municipio, Puerto Rico, collectively account for over half of all new HIV diagnoses that occurred in recent years. The EHE also prioritizes seven southern states where more than 10\% of new infections originated from rural regions and at least 75 people had been diagnosed annually (see Appendix for a list of all priority jurisdictions).\textsuperscript{12,13}

While some studies have evaluated the impact of the EHE plan in reducing HIV incidence, analyses exclusively considered the priority counties\textsuperscript{14–17} or did not incorporate
transmission dynamics. To understand the effectiveness of the EHE initiative in reducing incidence in the US overall by 2030, it is essential to track the dynamic HIV epidemiological trajectory in both priority and non-priority jurisdictions. Therefore, we developed a mathematical model of HIV transmission that assesses the EHE initiative combined with continuing progress towards UNAIDS goals in the remainder of the country. We projected trajectories of HIV incidence and forecast the reduction in annual new HIV infections, number of PLHIV, and cumulative new infections averted by 2030. In priority counties, for every state, and nationally, we find that these HIV care continuum and prevention efforts will substantially reduce new infections, reversing the growth of the epidemic at all spatial levels.

Methods

We developed a dynamic transmission and progression model of HIV, with epidemiological states including susceptible to infection, acute infection, undiagnosed chronic infection, diagnosed, treated, viral suppression and living with AIDS (SI Figure 1). Epidemiological parameters were informed from literature and the model was tailored to each jurisdiction using demographic data (SI Table 1, SI Section 2 Data Sources and Model initialization). County-specific and state-specific transmission rates were fitted to historical trajectories of incidence and prevalence in each jurisdiction (SI Section 3 Model Fitting). In total, we independently calibrated to 101 jurisdictions: 48 priority counties, Washington, DC, San Juan Municipio and seven targeted states for the EHE goal as well as 43 untargeted states and Puerto Rico for the UNAIDS goal. Initial conditions for unprioritized states that contained at least one priority county were adjusted to exclude the priority counties.

Under status quo projections, we maintained the 2019 levels of the HIV care continuum for diagnosis, treatment, and viral suppression as well as 2020 PrEP coverage in all model jurisdictions. For the EHE scenario in priority jurisdictions, current levels of the HIV care continuum were linearly increased to reach 95% of all PLHIV being aware of their HIV status, 95% of all diagnosed individuals being on treatment, and 100% of all individuals
on treatment achieving viral suppression by 2025 (95–95–100 by 2025), as well as a linear scale-up in PrEP coverage to 50% among eligible individuals. In the rest of the country, we assume the UNAIDS goal of 95–95–95 is achieved by 2030 with no increase in PrEP coverage.

For each jurisdiction, we sampled the transmission rate from its fitted distribution and the rest of the model parameters from empirical distributions to simulate 1,000 HIV trajectories under both the status quo and EHE scenarios. We summarized our results with median and interquartile range for annual incidence and the number of PLHIV from 2020 to 2030. By aggregating the projections of priority and non-priority regions of a state, we generated projections for each state of the US. We present the impact of achieving the EHE goal in the US at the county and state levels, as well as the impact nationally by aggregating our results (see SI for additional details).

**Results**

*Priority Jurisdiction Projections*

We found that if status quo interventions are maintained across all jurisdictions, a total of 190,133 [interquartile range: 174,714, 206,574] new HIV infections would be expected over the next decade in those prioritized under the EHE plan (Figure 1A). The successful implementation of the EHE plan is projected to reduce cumulative new infections by 57.3% [55.4%, 59.2%] in these jurisdictions, averting 108,927 [103,510, 114,374] infections (Figure 1B). If the EHE goals are achieved, we project a 72.1% [67.5%, 76.6%] reduction in 2030 annual incidence compared to 2020 across all priority jurisdictions.

*Projections for priority states*

There is substantial heterogeneity in the impact of EHE on cumulative new infections in the priority jurisdictions. Among all priority states, South Carolina is projected to avert the highest number of cumulative new infections (4080 [3772, 4387]) in contrast to Arkansas which is projected to avert the fewest (1832 [1565, 2164]) (Figure 2, see SI for projections of
all metrics for each analyzed jurisdiction in the US). However, the population of Arkansas is much smaller than that of South Carolina. At the per capita level, the impact is more similar between the two states (7.3 vs 9.2 per 10,000). In fact, the percent reduction in annual new infections between 2020 and 2030 for Arkansas is the greatest of any state at 80.9% [75.6%, 85.3%]. From the per capita perspective, the EHE plan would avert the highest rate of new infections in Mississippi (12.4 per 10,000 capita), and the lowest rate in Missouri (4.8 per 10,000 capita) (Figure 2).

Projections for priority counties

At the county level, Harris County, TX is expected to avert the most cumulative new infections (7158 [6468, 7911]) while San Francisco is projected to avert the fewest over the next decade (356 [276, 434]). Per capita, the change in new infections is highest for Fulton County, GA (28.6 per 10,000). Fulton County, GA is also the priority jurisdiction with the highest per capita annual incidence at the onset of the model (5.3 per 10,000).

The greatest reduction in annual new infections in 2030 compared to 2020 is expected for San Juan Municipio, PR (87.7% [84.4%, 90.4%]) and the smallest reduction for San Francisco County, CA (45.3% [36.4%, 54.4%]) (Figure 3). While the cumulative infections averted in San Juan Municipio and San Francisco County are projected to be similar (499 [480, 517] and 356 [276, 434], respectively), per capita infections averted are more divergent (18 per 10,000 capita; 4.4 per 10,000 capita, respectively).

State and territory-level projections under EHE and UNAIDS goals

We also projected annual new infections and cumulative new infections over a decade for non-prioritized jurisdictions (non-targeted states as well as the non-targeted regions of states that contain targeted counties) under status quo versus the UNAIDS 95–95–95 goal by 2030. In the scenario when the UNAIDS goal is implemented in non-prioritized regions, the EHE plan is simulated in prioritized regions.
Across all states, a median of 2.8 [1.3, 5.3] per 10,000 capita cumulative new infections from 2020 to 2030 are projected to be averted under the EHE/UNAIDS scenario compared to maintaining the status quo. Vermont and Montana are projected to avert the lowest absolute numbers of cumulative infections (35 [27, 42] and 37 [29, 45], respectively), as well as the lowest per capita (0.6 and 0.4 per 10,000 capita). By contrast, Texas is projected to avert the highest absolute number (21,601 [19,071, 24,509]), and relatively high per capita infections at 8.6 per 10,000, although five other states would experience greater per capita reductions: Nevada (16.3 per 10,000), Georgia (14.4 per 10,000), Mississippi (12.4 per 10,000), Florida (9.3 per 10,000), and South Carolina (9.2 per 10,000).

The median percent reduction at the state level in annual new infections from 2020 to 2030 with the EHE plan is 47.0% [38.2%, 54.7%] (Figure 4). Montana, a non-priority state without any priority counties, is projected to have the smallest percent reduction in annual new infections in 2030 compared to 2020 for the EHE/UNAIDS scenario (23.2% [15.4%, 30.8%]). Montana currently has the lowest HIV incidence rate in the country at 0.24 per 10,000 capita. As noted above, the priority state of Arkansas is expected to benefit from the most dramatic reduction in annual new infections (80.9% [75.6%, 85.3%]). Across all non-priority states that contain at least one priority county, Nevada, is projected to achieve the highest percent reduction in annual incidence by 2030 (71.6% [64.3%, 77.6%]).

Across the non-prioritized states that do not contain any priority counties, annual new infections are projected to increase under status quo interventions from 2020 to 2030 by 10.3%. This trajectory is predominantly attributable to rising incidence in Kansas and West Virginia, with increases of 38.8% [32.1%, 46.5%] and 41.5% [18.4%, 77.3%]. Under the UNAIDS goal, annual new infections would fall by 39.7% overall in non-prioritized states lacking priority counties. In particular, Kansas and West Virginia would experience 33.9% [22.9%, 43.6%] and 32.5% [15.2%, 46.0%] reductions, respectively. Further, if EHE prioritization were extended to Kansas and West Virginia, the reductions achieved are projected to be 73.2% [66.5%, 79.0%] and 78.3% [72.5%, 83.1%], respectively.
National projections

Nationally, we project that achieving the EHE plan of expanding diagnosis, treatment, viral suppression, and PrEP coverage in priority jurisdictions would achieve a 57.9% [49.8%, 64.6%] reduction in annual new HIV infections for 2030 compared to 2020. If the status quo is maintained, we predict that there will be 1,253,443 [1,220,525, 1,287,866] PLHIV in the US in 2030 (Figure 5). With the EHE intervention, we project a decline in PLHIV to 1,156,670 [1,133,738, 1,182,087] PLHIV by 2030, thereby reversing the growth in the epidemic that has been perpetuated throughout the last 40 years. The prevention components of EHE will more than compensate for the longer life expectancies of PLHIV through expanded viral suppression. The improvements on the HIV care continuum in combination with PrEP under the EHE plan are projected to reduce the national incidence rate from 1.14 [1.10, 1.18] per 10,000 capita under the status quo in 2030 to 0.43 [0.36, 0.51] per 10,000 capita. We project that 369,614 [363,228, 377,541] cumulative new HIV infections will occur between 2020 and 2030 under status quo interventions. Successful implementation of the EHE plan would avert 157,607 [141,467, 172,062] of those infections.

Discussion

The EHE initiative has potential to accelerate the stagnant rate of progress towards HIV control by improving HIV prevention and treatment engagement through a geographically-focused approach targeting high-incidence jurisdictions across the US. In priority regions, annual new infections would be reduced by 72.1% [IQR: 67.5%, 76.6%] from 2020 to 2030. In addition, if the UNAIDS goal is concomitantly achieved in non-prioritized regions, a 57% reduction in annual new infections is expected nationally by 2030.

Our results show that the EHE initiative can turn the tide on HIV incidence and prevalence. We project sufficient reduction in the number of annual new infections while increasing viral suppression such that the total number of PLHIV, both diagnosed and undiagnosed, will decline for the first time since HIV emerged in the US. The number of
PLHIV has increased steadily from 1.00 million in 2010 to 1.19 million in 2019. If improvements are not made beyond status quo interventions, we project that there will be over 1.25 million [1.22, 1.29 million] PLHIV by 2030. However, successful implementation of the EHE initiative is projected to substantially suppress transmission and avert more than 157,000 infections by 2030.

Our model also showed that EHE implementation would reduce the share of annual infections among the priority jurisdictions from 51.9% in 2020 to 39.7% in 2030, thus reducing the disproportionate HIV burden in these regions. The benefit relative to status quo associated with the EHE improvements in viral suppression depends on the current state of HIV treatment continuum and PrEP coverage for each jurisdiction. As an example, 67.9% of PLHIV are already virally suppressed and 63.7% of eligible individuals are prescribed PrEP in San Francisco, and the county is projected to experience among the lowest reductions in annual new infections as a result of EHE (45.2% [37.7%, 53.2%]). In contrast, the most substantial reduction in annual new infections (80.9% [75.6%, 85.3%]) is projected to occur in Arkansas, where currently only 35.8% of PLHIV in Arkansas are virally suppressed and only 12.4% of eligible individuals are prescribed PrEP.

Among all non-prioritized states, achieving the UNAIDS goal would bring about a reversal from an increase in annual new infections to a significant decrease overall. Although Kansas and West Virginia both currently report low annual new infections, their incidences per capita are comparable to that of priority counties and have been rising steadily over the last decade. While achieving the UNAIDS goal can lower annual infections by over 30% in both these states, extending the EHE initiative to them would achieve reductions of 73.2% and 78.3%, respectively. Therefore, expansion of the EHE initiative should be considered to ensure concomitant progress in the rest of the country.
In addition to the geographic heterogeneity of the HIV epidemic in the US, there are several racial, ethnic, sexual, and gender communities with disproportionately high rates of new infections, demonstrating inequity in prevention and care initiatives to date. Men who have sex with men (MSM), transgender women who have sex with men, Black, Latinx, certain indigenous populations, and people who inject drugs are at increased risk of HIV due to systemic oppression, racism, and stigma. For example, MSM account for 70% of new HIV infections in the US, of which Black MSM represent 37% of all new diagnoses. Many of the priority jurisdictions in the EHE plan are where these disparities are most pronounced. Currently, community and faith-based organizations, government agencies, and other local entities are all working to improve local services given inadequate access to HIV prevention and care. Further investment in community-specific, culturally congruent programs will allow for tailored responses to the unique challenges each setting faces, in combination with the overarching goals of the EHE plan.

Amidst the ongoing HIV epidemic is the COVID-19 pandemic, which has affected the US disproportionately among high-income countries. The flaws with the inadequate American health care system were magnified, and many of the same racial and ethnic disparities observed in the HIV epidemic were mirrored by the COVID-19 pandemic. While roll out of the EHE plan has begun, the COVID-19 pandemic has posed implementation challenges due to redeployment of front-line workers in HIV/AIDS control and prevention, which has limited the ability to expand access to diagnosis and care. Telehealth as well as funding via the Coronavirus Aid, Relief and Economic Security Act have been used to mitigate some of these challenges, yet sustained efforts are necessary to ensure that progress towards the successful implementation of the EHE plan is not jeopardized.
Improving HIV diagnosis and treatment engagement is critical to ending the HIV epidemic in the US. Our findings suggest that the EHE plan is poised to make a significant impact in the health of Americans by galvanizing expansion of HIV prevention and treatment towards HIV control and ultimately underscore the importance of the EHE initiative as a national priority.
References


14. Fojo AT, Schnure M, Kasaie P, Dowdy DW, Shah M. What Will It Take to End HIV in


Figure 2.1. Cumulative new HIV infections between 2020 and 2030 in priority jurisdictions (A) under maintenance of status quo and (B) implementing the Ending the HIV Epidemic (EHE) intervention.

Figure 2.2. Cumulative new HIV infections per 10,000 capita between 2020 and 2030 in priority jurisdictions (A) under maintenance of status quo and (B) implementing the Ending the HIV Epidemic (EHE) intervention.

Figure 2.3. Percent reduction in annual new HIV infections in 2030 compared to 2020 upon enactment of the Ending the HIV Epidemic (EHE) intervention.

Percent reduction in annual new HIV infections across all targeted jurisdictions in the US.

Figure 2.4. Percent reduction in annual new infections in 2030 compared to 2020 with the Ending the HIV Epidemic (EHE) and UNAIDS initiatives.

Figure 2.5. National projections of the number of people living with HIV, incidence per 10,000 capita, and cumulative new HIV infections across the US between 2020 and 2030 upon maintaining the status quo (blue) and successfully implementing the Ending the HIV Epidemic (EHE) and UNAIDS initiatives (orange).
Supplement: Chapter 2

Methods:

We developed a dynamic model of HIV transmission and progression that is fit to county and state-level data across the United States. Simulations of our mathematical model, project the number of individuals in each HIV-related stratum from 2020 to 2030. We evaluated the impact of achieving the US Government’s Ending the HIV Epidemic: A Plan For America (EHE) goals, which aim to improve the performance of 48 counties and seven states in addition to Washington, D.C. and San Juan, Puerto Rico on the HIV care continuum.

Data and model parameters were acquired from a combination of the Centers for Disease Control, the US Department of Health & Human Services, and peer-reviewed literature (Section S.2.2). County and state transmission rates were fitted to estimates of historical trajectories of HIV incidence and prevalence. For our intervention, we conducted 1000 model simulations in which we sampled values of model parameters from empirical distributions of each simulation and reported the results with both median values and percentiles. We assessed the impact of achieving the EHE goal in the US at the county and state levels, as well as the impact nationally by aggregating the results.
S.2.1 Mathematical Model

For each jurisdiction considered, our continuous-time, compartmental model stratifies individuals aged 15 years and above into 7 health states: susceptible to HIV infection (S), acute HIV infection (A), undiagnosed HIV infection (U), diagnosed but untreated HIV infection (D), treated without achieving viral suppression (T), virally suppressed (V), and having AIDS (W) (Figure S.1). Individuals in A, U, D, T, V and W compartments together comprise the people living with HIV (PLHIV). Transition rates between compartments are governed by a series of differential equations parameterized using rates estimated from other studies and from county, municipality, and state-specific incidence and prevalence data (Table S.2.1). Uncertainty in the model parameters was incorporated by running the model 1000 times with samples from the parameter distributions.

Figure S.2.1: Compartmental model diagram.

People susceptible to HIV (S) acquire HIV infection and move to the acute phase (A) according to a force of infection. This force of infection is dependent on HIV transmission rate, which is estimated from jurisdiction-specific incidence and prevalence data (Section S.2.3).
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>Background death rate</td>
<td>Jurisdiction specific</td>
<td>(1)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Recruitment rate</td>
<td>Jurisdiction specific</td>
<td>(1)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Rate of progression from acute infection</td>
<td>Triangular(2, 4, 14, 9.6)</td>
<td>(2)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Rate of developing AIDS among undiagnosed and diagnosed individuals who are not under treatment nor virally suppressed</td>
<td>0.08/year</td>
<td>(3)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Rate of developing AIDS among treated and virally suppressed individuals</td>
<td>Eq. S.2.1</td>
<td>–</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Rate of viral suppression</td>
<td>Uniform(1, 1.714)/year</td>
<td>(4)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Death rate from AIDS</td>
<td>0.25/year</td>
<td>(5)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>Reduction in life expectancy with viral suppression</td>
<td>Triangular(5, 8)</td>
<td>(6, 7)</td>
</tr>
<tr>
<td>$\tau_A$</td>
<td>Transmissibility per coital act during acute phase</td>
<td>Triangular(0.0039, 0.0082, 0.0150)</td>
<td>(8, 9)</td>
</tr>
<tr>
<td>$\tau_U$</td>
<td>Transmissibility per coital act after acute phase</td>
<td>Triangular(0.00077, 0.0014, 0.00251)</td>
<td>(10)</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Relative transmissibility per coital act with viral suppression</td>
<td>$\beta(0.08, 0.002, 0.57)$</td>
<td>(11)</td>
</tr>
<tr>
<td>$n$</td>
<td>Coital acts per year</td>
<td>Uniform(96, 108)</td>
<td>(9, 12)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Diagnosis rate</td>
<td>Eq. S.15</td>
<td>–</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Treatment rate</td>
<td>Eq. S.15</td>
<td>–</td>
</tr>
</tbody>
</table>

Table S.2.1: Parameters used in model. (T: triangular distribution, U: uniform distribution, $\beta$: beta distribution).
As the duration of acute infection is approximately 3 months, we assume that individuals remain undiagnosed during this phase and at least some of the subsequent period where they move to the undiagnosed compartment \((U)\). Individuals in the undiagnosed compartment get diagnosed at a rate \((\alpha)\) determined by the jurisdiction’s initial diagnosis level. Once diagnosed \((D)\), people may transition to the treated compartment when they start ART at a rate of \(\phi\), and then to the virally suppressed compartment \((V)\) after achieving viral suppression through continued treatment at a rate of \(\gamma\). Disengagement from treatment moves people who are on treatment \((T\text{ and } V)\) back to the diagnosed but untreated compartment \((D)\) at a rate of \(\psi\). Rates of movement between the undiagnosed \((U)\), diagnosed \((D)\), treated \((T)\) and virally suppressed \((V)\) compartments are determined by the proportions of PLHIV diagnosed \((p_D)\), proportion of diagnosed on ART \((p_T)\), and the proportion of individuals on ART who have achieved viral suppression \((p_V)\) (More details in S.4.1). PLHIV who are diagnosed but are not on treatment \((D)\) progress to AIDS \((W)\) at an average rate of \(\sigma\), whereas people on treatment \((T, V)\) take longer to develop AIDS \((1/\theta)\). The rate of developing AIDS among those who are receiving treatment \((\theta)\) is

\[
\theta = \frac{1}{1/\mu - \omega - 1/\nu - \mu},
\]

where \(\mu\) is the background mortality rate, \(\omega\) is the reduced life expectancy associated with PLHIV on treatment, and \(\nu\) is the mortality rate when living with AIDS. Differential equations were added to track the cumulative number of new infections \((Y)\) during the simulation period.

The model equations are
\[
\begin{align*}
\frac{dS}{dt} &= \kappa N - \lambda S - \mu S, \\
\frac{dA}{dt} &= \lambda S - \delta A - \mu A, \\
\frac{dU}{dt} &= \delta A - \alpha U - \mu U, \\
\frac{dD}{dt} &= \alpha U + \psi T + \psi V - \phi D - \mu D - \sigma D, \\
\frac{dT}{dt} &= \phi D - \psi T - \gamma T - \mu T - \theta T, \\
\frac{dV}{dt} &= \gamma T - \psi V - \mu V - \theta V, \\
\frac{dW}{dt} &= \sigma D + \theta T + \theta V - \nu W, \\
\end{align*}
\]

with force of infection

\[
\lambda = \frac{\xi[\beta_A A + \beta_U (U + D + T) + \beta_V V]}{N},
\]

and sexually active population size of \( N = S + A + U + D + T + V \). We assumed that individuals with AIDS (\( W \)) were too ill to be sexually active. The relative transmission rates per unit of time (\( \beta \)) from acute (\( A \)), unsuppressed (\( U, D, \) and \( T \)) and virally suppressed individuals (\( V \)) are:

\[
\begin{align*}
\beta_A &= 1 - (1 - \tau_A)^n, \\
\beta_U &= 1 - (1 - \tau_U)^n, \\
\beta_V &= 1 - (1 - \epsilon \tau_U)^n,
\end{align*}
\]

where \( \tau_A \) and \( \tau_U \) are the relative transmissibilities per coital act during acute and later phase of infection of infection, \( n \) is the annual number of coital acts, and \( \epsilon \) is the relative transmissibility per coital act under viral suppression. We calculated county and state-
specific transmission rates (ξ) from longitudinal trends in HIV prevalence and incidence (Section S.2.3).

S.2.2 Data Sources and Model initialization

Population

Population data at county and state levels on individuals aged 15 and above was acquired from CDC Wonder for years 2010 to 2019 (1). Population data for Puerto Rico and San Juan Municipio from 2010 to 2019 were acquired from the American Community Survey (13). The population growth rate was calculated by taking the difference between county and state-level birth and death rates (1), acquired from CDC Wonder.

Incidence and Prevalence

State and county level data on estimated HIV incidence and estimated HIV prevalence among individuals aged 13 and older were obtained from CDC Atlas Plus for 2010-2019 and 2017-2019, respectively (14).

Diagnosis, Treatment, and Viral Suppression Levels

Diagnosis, treatment, and viral suppression levels at the county and state levels were acquired from CDC Atlas Plus for 2017-2019 (14). These levels informed the initial conditions for the model for the number of people who have been diagnosed with HIV, who are on anti-retroviral therapy, and who have achieved viral suppression or have been retained on treatment for at least 12 months.
Pre-exposure Prophylaxis (PrEP) uptake Levels

Jurisdiction-specific proportions of the eligible population who are regularly taking pre-exposure prophylaxis for HIV (PrEP) were acquired from America’s HIV Epidemic Analysis Dashboard (AHEAD) for 2020 (15). Individuals with indications for PrEP include those with any risk factors for acquiring HIV such as condomless sex with a partner with a detectable or unknown HIV-RNA level, recent sexually transmitted infection, or injection drug use.

According to surveillance data, San Francisco County, CA and King County, WA are the prioritized counties with the smallest HIV care continuum gap at the onset of the forecast period with 67.9% and 66.8% of PLHIV having achieved viral suppression in the two counties in 2019, respectively.(15) New York County, NY (72.1%) and San Francisco County, CA (63.7%) have the highest PrEP coverage among eligible individuals of all priority counties. Among the unprioritized jurisdictions, Vermont (68.8%) and Montana (68.2%) have the highest proportions of PLHIV with suppressed viral loads while Pennsylvania (25.2%) has the lowest. Arkansas has the lowest proportion of PLHIV virally suppressed among the priority states (35.8%), and Hamilton County, OH has the lowest proportion among the priority counties (36.9%). PrEP coverage among those eligible is lowest in San Juan Municipio, PR (2.9%) with Kentucky reporting the lowest PrEP coverage at the forecast onset among all priority states (9.6%).

Initialization

short duration of this phase, we assumed that there were no acute infections initially in the model. We assumed that the proportion of diagnosed but untreated individuals \((D)\) who have AIDS \((W)\) in 2019 was:

\[
p_A = \frac{\sigma}{\sigma + \nu},
\]

which is the equilibrium fraction of diagnosed individuals who have AIDS in the absence of treatment and background mortality. The model initial conditions are as follows:

\[
\]

\[
A(2019) = 0,
\]

\[
\]

\[
\]

\[
\]

\[
\]

\[
\]

\[
Y(2019) = 0.
\]

Parameterized differential equations were solved numerically from 2019 through 2030 using the LSODA routine (16, 17). We computed cumulative new infections, HIV incidence and PLHIV at each time point using the outcomes from the solutions.

Our results from the targeted jurisdictions, untargeted portions of states that contain one or more targeted jurisdictions, and untargeted states were aggregated to national levels by adding the numbers of people in each compartment over time and then cumulative new infections, annual new infections, incidence per population, and PLHIV were calculated from the aggregates.
S.2.3 Model Fitting

To tailor our model to each of the jurisdiction considered, we used available historical jurisdiction specific estimates of prevalence and incidence, and published estimates of transmissibility per coital act to derive an average jurisdiction-specific transmission rate.

Assuming that there is no difference in transmission risk among PLHIV in any phase of the disease or treatment status, we approximated the force of infection (Equation S.2.3) with a jurisdiction specific time-dependent aggregate transmission rate $\beta(t)$ as

$$\lambda(t) = \frac{\xi[\beta_A A(t) + \beta_U U(t) + D(t) + T(t) + \beta_V V(t)]}{N(t)} \approx \beta(t) \frac{I}{N}, \quad (S.2.7)$$

where $I = A + U + D + T + V + W$ is the total number of PLHIV. Therefore, the per-capita incidence can be written as

$$i(t) = \lambda(t) \frac{S(t)}{N(t)} \approx \beta(t) \frac{I(t)}{N(t)} \frac{S(t)}{N(t)} = \beta(t)p(t)(1 - p(t)), \quad (S.2.8)$$

where $p(t)$ is the prevalence. Thus, we can estimate the transmission rate at each historical time point using the incidence and prevalence data,

$$\beta(t) \approx \frac{i(t)}{p(t)(1 - p(t))} \quad (S.2.9)$$

To reflect the uncertainty in $\beta(t)$, we assumed it follows a lognormal distribution such that its parameters $\mu$ and $\sigma^2$ were informed by exponentially weighted mean and variance (with a half-life of 1 year) of log $\beta(t)$ for the year 2019. That is

$$\mu = \frac{\sum_{i=1}^{n} 2^{t_i} \log \beta(t_i)}{\sum_{i=1}^{n} 2^{t_i}} \quad (S.2.10)$$

$$\sigma^2 = \frac{\langle \sum_{i=1}^{n} 2^{t_i} \rangle \left[ \sum_{i=1}^{n} 2^{t_i}(\log \beta(t_i) - \mu)^2 \right]}{\langle \sum_{i=1}^{n} 2^{t_i} \rangle^2 - \sum_{i=1}^{n} 2^{2t_i}}$$
Using the mode of this lognormal distribution as aggregate transmission rate $\bar{\beta}$, we rearranged Equation S.7 to derive the jurisdiction-specific transmission rate as

$$\xi = \frac{\bar{\beta} I(2019)}{\beta_A A(2019) + \beta_U (U(2019) + D(2019) + T(2019)) + \beta_W V(2019)}$$  \hspace{1cm} (S.2.11)

and applied the model initial conditions on proportion of PLHIV divided among acute infections, undiagnosed, diagnosed, or treated infections; and virally suppressed infections.

### S.2.4 Analysis and Implementation

#### S.2.4.1 Calculating rates based on DTV

The proportion of PLHIV who have been diagnosed at time $t$ is

$$p_{D}(t) = \frac{D(t) + T(t) + V(t) + W(t)}{A(t) + U(t) + D(t) + T(t) + V(t) + W(t)}. \hspace{1cm} (S.2.12)$$

The proportion of diagnosed individuals who are receiving treatment at time $t$ is

$$p_{T}(t) = \frac{T(t) + V(t) + W(t)}{D(t) + T(t) + V(t) + W(t)}. \hspace{1cm} (S.2.13)$$

and the proportion with viral suppression at time $t$ is

$$p_{V} = \frac{V(t)}{T(t) + V(t)}. \hspace{1cm} (S.2.14)$$

The proportion of PLHIV who are diagnosed, are on treatment, and are virally suppressed over time is determined by the scenario under evaluation such as maintaining status quo or increasing it to a target level over time. We represent these desired targets for proportion of PLHIV diagnosed, proportion of diagnosed individuals on treatment, and proportion
of individuals on treatment with viral suppression by $p^*_D(t)$, $p^*_T(t)$ and $p^*_V(t)$ respectively. To achieve these targets over time, we adjust the diagnosis rate over time ($\alpha(t)$), rate of treatment initiation over time ($\phi(t)$) and the rate of people relapsing to untreated over time ($\psi(t)$) according to the functions:

$$
\alpha(t) = \alpha_{\text{max}} H(p^*_D(t) - p_D(t)),
$$

$$
\phi(t) = \phi_{\text{max}} H(p^*_T(t) - p_T(t)),
$$

$$
\psi(t) = \psi_{\text{max}} H(p_V(t) - p^*_V(t)),
$$

where $\alpha_{\text{max}} = 1 \text{ y}^{-1}$, $\phi_{\text{max}} = 10$, and $\psi_{\text{max}} = 1$. $H(x)$ is the Heaviside-like function

$$
H(x) = \begin{cases} 
0 & \text{if } x < 0, \\
x/\chi & \text{if } 0 \leq x \leq \chi, \\
1 & \text{if } x > \chi, 
\end{cases}
$$

with $\chi = 0.001$. The functional form allows us to avoid computational challenges that may occur when using discontinuous functions by rapidly switching from $H = 0$ to $H = 1$ when $x$ changes sign from negative to positive. The functional form in Equation S.15 allows new diagnoses ($\alpha(t) > 0$) or new treatment ($\phi(t) > 0$), when the current levels are below the target diagnosis ($p^*_D(t) > p_D(t)$) or treatment ($p^*_T(t) > p_T(t)$) level. When the rates are at or above their target, they stop new diagnoses ($\alpha(t) = 0$) and stop new treatment initiations ($\phi(t) = 0$) respectively. As viral suppression rates depend on the time on treatment, we varied the relapse rate to achieve the target level of viral suppression. Therefore, when the current level of viral suppression is below the target ($p^*_V(t) > p_V(t)$), relapses are prevented ($\psi(t) = 0$), whereas relapses are permitted ($\psi(t) > 0$) when the current level of
viral suppression is at or above the target levels.

S.2.4.2 Implementation

We run our model under different scenarios for each of the jurisdictions considered. Jurisdictions considered include 48 targeted counties, Washington, D.C., San Juan Municipo, P.R., and all 50 US states. For the states that were not targeted, but contained at least one targeted county, we adjusted its initial conditions such that projections reflect results for the remainder of the state excluding the targeted counties, assuming there was no interaction between the targeted counties and the rest of the state in which it is located. For each scenario, we run the model from 2019 to 2030.

Scenarios considered are as follows:

Status Quo

For the status quo scenario, we run the model for each jurisdiction such that that status quo levels of proportions diagnosed, on treatment, and virally suppressed remain constant from 2019 to 2030. Therefore, the status quo scenario maintains the proportions $p_D$, $p_T$, and $p_V$ at their initial levels going forward: $p_D - p_T - p_V$. This scenario is implemented by computing diagnosis, treatment and relapse rates over time (Equation S.2.15) with initial proportions assigned as the targets over time:

$$
p^*_D(t) = p_D(2019),
$$
$$
p^*_T(t) = p_T(2019),
$$
$$
p^*_V(t) = p_V(2019).
$$

Current jurisdiction-specific levels of PrEP coverage among the eligible population were maintained in the Status Quo scenario, which is implicitly accounted for in the estimated
jurisdiction-specific transmission rate (Equation S.2.11, Equation S.2.20).

### EHE Goal

The U.S. Department of Health and Human Services' Ending the HIV Epidemic in the U.S. (EHE) initiative aims to provide 57 geographic focus areas with an infusion of additional resources, expertise, and technology to develop and implement locally tailored plans. We evaluated the impact of achieving the EHE goal in targeted jurisdictions by modeling the expansion of diagnosis, treatment, and viral suppression rates as well as PrEP coverage such that:

1. 95% of all PLHIV know their HIV status by 2025,

2. 95% of all diagnosed individual are on treatment by 2025,

3. 100% of all individuals on treatment achieve viral suppression by 2025 (which is equivalent to 95% of all individuals who receive a diagnosis of HIV achieving viral suppression), and

4. 50% individuals with indications for PrEP are receiving PrEP by 2025.

After achieving 95–95–100 by 2025, the diagnosis, treatment, and viral suppression proportions, in addition to the 50% PrEP coverage, are maintained from 2025 through 2030.

### Diagnosis and treatment expansion for targeted jurisdictions

We implement the EHE goal for targeted jurisdictions such that targeted jurisdictions reach 95–95–100 by 2025 and maintain those levels of intervention until 2030. To implement the EHE goal, we compute the diagnosis, treatment and relapse rates over time (Equation S.2.15) with targets for the proportion of PLHIV diagnosed, proportion of diagnosed individuals on treatment,
and proportion of individuals on treatment with viral suppression varying over time as follows:

\[ p^*_D(t) = \begin{cases} 
F(t, 2019, p_D(2019), 2025, 95) & \text{if } p_D(2019) < 0.95, \\
p_D(2019) & \text{if } p_D(2019) \geq 0.95.
\end{cases} \]  

\[ p^*_T(t) = \begin{cases} 
F(t, 2019, p_T(2019), 2025, 95) & \text{if } p_T(2019) < 0.95, \\
p_T(2019) & \text{if } p_T(2019) \geq 0.95.
\end{cases} \]  

\[ p^*_V(t) = \begin{cases} 
F(t, 2019, p_V(2019), 2025, 100) & \text{if } p_V(2019) < 1, \\
\end{cases} \]  

where

\[ F(t, t_0, x_0, t_1, x_1) = \begin{cases} 
x_0 & \text{if } t < t_0, \\
x_0 + (x_1 - x_0) \frac{t - t_0}{t_1 - t_0} & \text{if } t_0 \leq t < t_1, \\
x_1 & \text{if } t \geq t_1,
\end{cases} \]  

is constant at \( x_0 \) for \( t < t_0 \), linearly connects \( x_0 \) at \( t_0 \) to \( x_1 \) at \( t_1 \) for \( t_0 \leq t < t_1 \), and is constant at \( x_1 \) for \( t \geq t_1 \).

**Scaling up PrEP coverage in targeted jurisdictions**  Under the EHE goal, targeted jurisdictions aim to achieve a 50% PrEP coverage among eligible populations by 2025.

\[ \xi = (1 - R(P_0))\xi_0 \]  

where \( \xi_0 \) is the jurisdiction specific transmission rate in the absence of any PrEP coverage, \( P_0 \) represents the current level of PrEP coverage among eligible individuals and \( R \) is the function that evaluates the population-level reduction in transmission for a given PrEP
coverage among eligible individuals. Similarly, achieving a targeted PrEP coverage \( P_T \) would result in a reduction of \( R(P_T) \) to the jurisdiction specific transmission rate in the absence of any PrEP coverage, \( \xi_0 \). Therefore, jurisdiction-specific transmission rate after achieving targeted scale-up in PrEP coverage can be written as

\[
\xi_T = \frac{1 - R(P_T)}{1 - R(P_0)} \xi. \tag{S.2.21}
\]

Current levels of PrEP coverage for each targeted jurisdiction were informed by the proportion of individuals who have been prescribed PrEP among those with indications for PrEP from the CDC. \((18)\) For each targeted jurisdiction, if the current level of PrEP coverage \( P_0 \) already exceeds the goal of 50% coverage (e.g. New York County, NY; San Francisco County, CA) the current coverage was maintained going forward. Otherwise, we implemented a scale-up in PrEP coverage by increasing the PrEP coverage linearly from its current level \( P_0 \) to the targeted coverage \( P_{2025} = 50\% \) by 2025, after which it remains constant.

\[
P_t = \begin{cases} 
P_0 & t \leq 2020 \\
P_0 + (0.5 - P_0) \frac{t - 2020}{2025 - 2020} & 2020 < t \leq 2025 \\
0.5 & t \geq 2025.
\end{cases} \tag{S.2.22}
\]

Therefore, for each targeted jurisdiction, the PrEP coverage scale-up component of the EHE goal is implemented by a reduction factor \( r_t \) multiplied to the the jurisdiction-specific transmission rate \( (\xi, \text{Equation S.2.11}) \), given by

\[
r_t = \frac{1 - R(P_t)}{1 - R(P_0)} \tag{S.2.23}
\]

where \( P_0 \) is jurisdiction-specific current level of PrEP coverage, \( P_t \) is coverage at time \( t \), and \( R \) is the functional form, determining population-level reduction in transmission.
Kasaie et al. 2017 (19) estimated the population level reduction in overall HIV incidence in Baltimore based on varying coverage of PrEP among eligible populations. We fitted a logistic regression to their data to determine the function form of $R$, which is given by:

$$R(P) = 0.2463 \times \log(P) + 0.70737$$  \hspace{1cm} (S.2.24)

where $P$ is the PrEP coverage among individuals with indications for PrEP.

**Diagnosis and treatment rates in non-targeted jurisdictions**  For the HHS goal, we assume that non-targeted jurisdictions reach 95–95–95 by 2030. This assumption was made as it is consistent with goals of the Joint United Nations Programme on HIV/AIDS (UNAIDS) for 2030 worldwide and is implemented by computing the rates of diagnosis ($\alpha(t)$), treatment initiation ($\phi(t)$) and viral suppression ($\psi(t)$) using the targeted proportions over time in Equation S.2.15.

$$p^*_D(t) = \begin{cases} 
F(t, 2019, p_D(2019), 2030, 95) & \text{if } p_D(2019) < 0.95, \\
p_D(2019) & \text{if } p_D(2019) \geq 0.95.
\end{cases}$$  \hspace{1cm} (S.2.25)

$$p^*_T(t) = \begin{cases} 
F(t, 2019, p_T(2019), 2030, 95) & \text{if } p_T(2019) < 0.95, \\
p_T(2019) & \text{if } p_T(2019) \geq 0.95.
\end{cases}$$

$$p^*_V(t) = \begin{cases} 
F(t, 2019, p_V(2019), 2030, 95) & \text{if } p_V(2019) < 0.95, \\
p_V(2019) & \text{if } p_V(2019) \geq 0.95.
\end{cases}$$

where function $F$ is given by Equation S.2.19.
S.2.6 Definitions

Targeted Jurisdictions: 48 counties, San Juan, PR, and Washington, D.C., which contained over half of all new HIV diagnoses in 2016 and 2017, as well as 7 states with notably high rates of rural new HIV infections (> 10%) and at least 75 new diagnoses state-wide in 2016 and 2017 were selected by HHS as the jurisdictions in which expertise and resources will be infused (15). These jurisdictions are as follows:

Targeted Counties and Territories: Maricopa County, Alameda County, Los Angeles County, Orange County, Riverside County, Sacramento County, San Bernardino County, San Diego County, San Francisco County, Broward County, Duval County, Hillsborough County, Miami-Dade County, Orange County, Palm Beach County, Pinellas County, Cobb County, DeKalb County, Fulton County, Gwinnett County, Cook County, Marion County, East Baton Rouge Parish, Orleans Parish, Baltimore City, Montgomery County, Prince George’s County, Suffolk County, Wayne County, Clark County, Essex County, Hudson County, Bronx County, Kings County, New York County, Queens County, Mecklenburg County, Cuyahoga County, Franklin County, Hamilton County, Philadelphia County, Shelby County, Bexar County, Dallas County, Harris County, Tarrant County, Travis County, King County, Washington, DC, San Juan Municipio.

Targeted States: Alabama, Arkansas, Kentucky, Mississippi, Missouri, Oklahoma, South Carolina.

References


Figure S.2.2: Projections of the median and interquartile range of the number of people living with HIV (PLHIV), HIV incidence per 100,000 capita, and cumulative new HIV infections across the United States between 2020 and 2030 upon maintaining the status quo (blue) and achieving the Ending the HIV Epidemic (EHE) and UNAIDS goal (orange) for each jurisdiction modelled in the United States.
Essex, NJ

Florida

Status Quo

Intervention

PLHIV

Incidence per 100,000 population

Cumulative New Infections

0 10 20 30

2020 2022 2024 2026 2028 2030

125,000 130,000 135,000

0 10 20

2020 2022 2024 2026 2028 2030

0 10,000 20,000 30,000 40,000
Chapter 3

Racial disparities in COVID-19 mortality across Michigan, United States

Alyssa S. Parpia, MPH,1 Isabel Martinez, PhD,1 Abdulrahman M. El-Sayed, MD, DPhil,2 Chad R. Wells, PhD,1 Lindsey Myers, BBA,3 Jeffrey Duncan, PhD,3 Jim Collins, MPH,4 Meagan C. Fitzpatrick, PhD,1,5 Alison P. Galvani, PhD1, Abhishek Pandey, PhD1

1 Center for Infectious Disease Modeling and Analysis (CIDMA), Yale School of Public Health, New Haven, CT, USA
2 Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, USA; Department of Public Health, Wayne State University, Detroit, MI, USA
3 Division for Vital Records and Health Statistics, Michigan Department of Health and Human Services, Detroit, MI, USA
4 Division of Communicable Diseases, Michigan Department of Health and Human Services, Detroit, MI, USA
5 University of Maryland School of Medicine, Baltimore, MD, USA

This manuscript has been previously published as:

Introduction:

Throughout history, epidemics have inequitably affected vulnerable populations in our societies,\textsuperscript{1,2} and the COVID-19 pandemic is no exception. In particular, Black and Hispanic/Latinx populations are disproportionately experiencing severe COVID-19 morbidity and mortality in the United States (US).\textsuperscript{3,4} According to data from the Centers for Disease Control, 53\% and 23\% of all COVID-19 deaths in the US are among White and Black individuals, respectively, while these races represent 42.1\% and 17.0\% of the population.\textsuperscript{5,6} In Michigan, the disparities are even starker: while Black Americans represent 14.1\% of the total population,\textsuperscript{7} 35.0\% of state-wide COVID-19 deaths as of November 5, 2020 and 23.1\% of state-wide deaths as of January 21, 2021 occurred in this group.\textsuperscript{8} Underlying this disparate burden are systemic inequities in socio-economic conditions, health, and access to care by race, which impact infection exposure and survival\textsuperscript{9,10}.

Such inequalities in health are often perpetuated by systemic racism, which can result in reduced access to healthcare and increased risks for health.\textsuperscript{11,12} Systemic racism refers to the embedded power in institutions and other structures led by those with both known and unrecognized biases that are reflected in regulations, policies and practices, perpetuating inequitable access to resources and opportunities.\textsuperscript{13} Amongst Black communities, longstanding marginalization has led to higher rates of housing instability, financial insecurity, and essential service inaccessibility compared to White communities.\textsuperscript{14–16} For instance, Black families in Michigan report 2.5 times higher rates of poverty\textsuperscript{17} and 43\% lower household incomes\textsuperscript{18} than White families. With lower socioeconomic status, health related costs can create significant challenges for obtaining medical care, medical insurance, and often delay diagnoses and treatments.\textsuperscript{19,20} Flint and Detroit, the residents of which are predominantly Black, are the second and fourth poorest cities in the US, respectively.\textsuperscript{21} With the confluence of housing, urbanization and socioeconomic disadvantage, Black communities are uniquely placed to face the brunt of contagious and widespread infections. Socioeconomic disparities can influence community-level COVID-19 transmission, as smaller living space,
greater household sizes, and reliance on public transportation each impact the potential risk for COVID-19 infection. In addition to disparities in access to healthcare, other aspects of structural racism such as environmental injustice and inequities in built environment, contribute to higher rates of chronic health conditions in Black populations such as diabetes, obesity, asthma, and cardiovascular disease in comparison to White populations, which affect the severity of disease following infection.

This multitude of factors creates a challenge for individuals in general, and especially those with underlying medical conditions, to protect themselves from infection. The “stay at home” public health guidelines implemented during the COVID-19 pandemic are in tension with the need for stable employment, wages, and housing. Pandemic related job losses have been experienced predominantly by those in lower income brackets, those in Black and Hispanic/Latinx communities, and those unable to perform work tasks at home. When a certain segment of the population is fundamentally unable to avoid risk and placed in a position to ignore infection prevention guidelines, disparities in morbidity and mortality inevitably arise.

The extent to which Black individuals are being disproportionately killed by the pandemic is not well-quantified. In this study, we aim to uncover whether disparities by race in COVID-19 mortality can be explained by demographic and underlying health characteristics alone. We analyzed individual-level data on people who died from COVID-19 in Michigan, stratified by demographic characteristics, chronic conditions, and geographical location. Then, we assessed disparities in mortality risk by race within each ZIP Code Tabulation Area between March 16 to October 26, 2020, thereby evaluating the interdependent crises of the COVID-19 epidemic and systemic racism in Michigan. Further, we used statistical analyses to assess disparities across races among individuals dying from COVID-19 in Michigan. We found that Black individuals have 3.6 times the risk of dying from COVID-19 as White individuals in Michigan overall (p<0.001). Among those with no comorbidities under the age of 65, Black individuals have 12.6 times the mortality rate of White individuals (p<0.001).
Methods:

Data Sources:

*Linked Individual-level COVID-19 Surveillance and Death Certificate Data*

Yale IRB deemed the study as not Human Subjects research and as such, IRB approval was not required. As the data obtained for research purposes were de-identified, patient consent was also not required. Linked death certificate and COVID-19 surveillance data were obtained from the Michigan Department of Health and Human Services. The Division of Vital Records and Health Statistics provided death certificate data for individuals with an International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code for COVID-19 (U07.1 or U07.2) as an underlying or related cause of death and the Communicable Diseases Division provided data from the Michigan Disease Surveillance System (MDSS) on COVID-19 deaths occurring between March 16 and October 26, 2020. As per the Data Use Agreement, ASP, AP, IM, and LM were given access to the linked data on November 10, 2020 for analysis. All COVID-19 related deaths occurred in the state of Michigan. Datasets were linked based on string values of first name, last name, and date of birth. Names and dates of birth were formatted uniformly and combined into string variables, and then linked using a Generalized edit distance technique which measures dissimilarity between two strings. Generalized edit distance scores of 0 were considered exact matches while scores ≥325 were automatically excluded as non-matches. Linkages with scores ranging from 1 to 324 were subject to individual manual review where commonality of first or last name, suffixes, sensible misspellings, and single differences in date of birth were considered.

Individual level data were collected on age, sex, race, ZIP Code Tabulation Area (ZCTAs) of residence, underlying and related causes of death, other medical conditions of interest, pre-existing conditions, immunosuppressive medications, retirement or unemployment status, and residence or employment at a high risk or congregate living facility. ZCTAs are generalized area representations of the US Postal Service Zip Code
service areas. Our dataset exclusively contains those who were diagnosed with and whose death was attributed to COVID-19 in Michigan.

Comorbidities were identified by combining information on underlying causes of death, related causes of death, and other medical conditions of interest from death certificates with information on pre-existing conditions and medications from the Surveillance System case report forms. Comorbidities were categorized as follows: asthma or reactive airway disease, cardiovascular disease, cancer, chronic lung disease, diabetes mellitus, neurologic disease, chronic liver disease, chronic renal disease, other immunosuppressive conditions, and other chronic diseases. Other immunosuppressive conditions included rheumatoid arthritis and bullous pemphigoid, among others, as well as those unspecified but noted as immunosuppressive. Other chronic diseases included anemia, depression, and chronic venous thromboembolism as well as those unspecified (See Supplement). Cardiovascular disease corresponded to ICD-10 codes I00-178, and includes heart attacks and strokes. Comorbidity of cancer represented any history of cancer.

ZCTA of residence was obtained primarily from death certificate data and from surveillance system case reports in instances of missing ZCTA on the death certificate. Only individuals who resided in Michigan were included in the mapping. Individuals who died from COVID-19 were categorized by their status of being residents or employees of high risk or congregate living facilities, which include: long-term care homes, skilled nursing facilities, assisted living facilities, homeless shelters, federal prisons, Michigan Department of Corrections prisons, county jail, juvenile justice facilities, foster care, and others, including senior, retirement, and group homes.

Population-level Demographic Data

Data on Michigan population demographics by age group, sex, and race were obtained from the Michigan Department of Health and Human Services. The prevalence of comorbidities in the Michigan population by age group, sex, and race were estimated using a combination of 2017 Michigan Medicare data and National Health Interview Survey data.
on national comorbidity distribution by age group and sex\textsuperscript{35} (Table S.3.3). Given data available to us on morbidity burden in the Michigan population, we assumed that the distribution of age and race and the distribution of sex and race were independent. Populations by ZCTA for all races were obtained from the United States Census Bureau Decennial Census.\textsuperscript{36}

**Analysis**

With a two-sample Z-test, we assessed equality of proportions of decedents that are White and Black with continuity corrections in the matched and unmatched data. Using a Chi-Square test, we determined whether a statistically significant difference ($\alpha = 0.05$) existed between the proportions of deaths which occurred among Black and White individuals in Michigan, and the proportions of the overall population who are Black and White in Michigan.

We also performed univariate and bivariate analyses on COVID-19 mortality by demographic characteristics and presence of single or multiple comorbidities. Specifically, Chi-square and Kruskal-Wallis tests were used to identify differences in the following variables by race: age, sex, number of comorbidities, presence of specific COVID-19 related comorbidities and comorbidity combinations, employment status, and high-risk or congregate living facility exposure as resident or staff. Age was analyzed as both a continuous variable, categorical variable (<40, 40-64, 65-79, $\geq$80 years), and as a binary variable (<65, $\geq$65 years). In addition, COVID-19 mortality rates and mortality rate ratios were calculated for the overall sample, as well as by race, age, sex, and number of comorbidities both separately and in combination. Chi-square tests were performed to identify differences in these mortality rates by race.

We combined the individual-level data on COVID-19 deaths with ZCTA level data on total population and population by race to evaluate geographic heterogeneity in mortality. For each ZCTA, we calculated the COVID-19 mortality rate overall and by race, as well as the mortality rate ratio of Black to White deaths. A one-tailed Wilcoxon signed-rank test was
used to compare whether the mortality rates among Black individuals were greater than those of White individuals across ZCTAs throughout Michigan.

We conducted descriptive analyses of the dates of COVID-19 symptom onset, hospitalization, and death due to COVID-19, stratified by race, to characterize epidemic progression. We also fit distributions to the time from symptom onset to hospitalization by race. All data were analyzed using Python 3.7.4 and R-3.5.1.

Role of the Funding Source: The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results:

Demographics of COVID-19 Decedents

Between March 16 and October 26, 2020, a total of 6,065 COVID-19 related deaths were recorded in the linked surveillance and death certificate databases across Michigan, of which 96.3% (5,838) occurred among Black or White individuals (Figure S.3.1). Of the total decedents, the races and ethnicities were as follows: 3497 (57.7%) White, 2341 (38.6%) Black, 18 (0.3%) Native American, 28 (0.5%) South Asian, 25 (0.4%) Southeast Asian, 11 (0.2%) East Asian, 30 (0.5%) other Asian ethnicity, and 115 (1.9%) were decedents of other races or did not have a race reported. A total of 159 (2.6%) decedents were Hispanic/Latinx, 5798 (95.6%) were non-Hispanic/Latinx, and 108 (1.8%) had missing data for this variable. Among White decedents, 1.4% were Hispanic/Latinx in contrast to 0.4% of Black decedents. We focused our analyses on individuals who were White and Black, given the small proportions of individuals of other races and ethnicities. Among the 865 unmatched records, 469 (54.2%) were White decedents and 349 (40.3%) were Black decedents. Proportions of White and Black decedents were not significantly different between matched and unmatched records (White: p=0.061, Black: p=0.342).
Black individuals represent 15.7% of the combined Black and White population of Michigan yet accounted for 40.1% of the COVID-19 deaths during our study period. The proportion of deaths reported among Black individuals in Michigan is significantly higher than the proportion of deaths we would expect based on population representation alone (p<0.001).

Black individuals who died from COVID-19 were significantly younger (median [IQR]: 72 [63, 81], p<0.001) than White individuals (81 [72, 89]) and reported lower rates of being to be retired or unemployed (12.9%, p<0.001) than White decedents (22.2%) (Table 3.1). Among all COVID-19 deaths, 44.0% were either residents or employees of high-risk or congregate living facilities, such as long-term care and senior homes, homeless shelters, and prisons. Black decedents (29.7%) had lower rates of living or working in high-risk or congregate living facilities than White individuals (54.8%, p<0.001). Among all decedents who worked or resided in these facilities, 90.6% were aged 65 and older (Table S.3.1).

The most common comorbidity dyads among all COVID-19 deaths were diabetes and cardiovascular disease (24.4%), and cardiovascular and neurologic disease (21.6%). Black decedents had significantly higher rates of reporting asthma (p=0.016), diabetes (p=0.002), and chronic renal disease (p=0.004) than White individuals who died from COVID-19, but significantly lower rates of cardiovascular disease (p<0.001), cancer (p<0.001), chronic lung disease (p<0.001), and neurologic disease (p<0.001). Additionally, Black decedents had higher rates of to have the combination of diabetes mellitus, cardiovascular disease, and chronic renal disease (p=0.003) than White decedents, and lower rates of the combinations of cardiovascular and chronic lung disease (p<0.001) and cardiovascular and neurologic disease (p<0.001).

**Mortality Rates**

Between March 16 to October 26, 2020, we calculated that the COVID-19 mortality rate in Michigan was 5.4 per 10,000 population. Stratifying by race, the mortality rate was 3.6 times higher for Black populations (15.6 per 10,000 population) than White populations.
(4.3 per 10,000 population, p<0.001). Stratifying by age, the mortality rate for Black individuals under 40 years of age (0.50 per 10,000 population) is 7.4 times that for White individuals in the same age group (0.07 per 10,000 population, p<0.001). Among those aged 40 to 69 years, the mortality rate for Black individuals (18.4 per 10,000) was 8.5 times that of White individuals (2.2 per 10,000, p<0.001). Mortality risk for both races increased with age, where Black individuals aged 70 years and older had a mortality rate of 121.1 per 10,000 population and White individuals had a mortality rate of 27.7 per 10,000 population (p<0.001).

While the mortality rate in the White population did not vary significantly by sex (4.5 per 10,000 population for males and 4.2 per 10,000 for females, p = 0.167), Black males have a significantly greater (31.2%) mortality rate than their female counterparts (p<0.001). Comparing males and females under the age of 40, the Black male mortality rate (0.77 per 10,000 population) is 10.8 times that of White males (0.07 per 10,000 population, p<0.001), and the Black female mortality rate (0.23 per 10,000 population) was 3.56 times that of White females in the same age group (0.06 per 10,000 population, p<0.001).

When stratified by age, sex, and number of comorbidities, the mortality rate for the Black population was significantly higher than that for the White population for every pairwise comparison (p<0.001). Among those with no comorbidities under the age of 65, Black individuals have 12.6 times the mortality rate of White individuals (p<0.001). Black males under the age of 65 years with no comorbidities had a COVID-19 mortality rate of 6.2 per 10,000 population, while White males with no comorbidities in the same age group experienced a lower mortality rate of 0.53 per 10,000 (p<0.001, Figure 3.1). Among males aged 65 and older with no comorbidities, the COVID-19 mortality rate was 367.5 per 10,000 for Black decedents and 31.6 per 10,000 for White decedents (p<0.001). Black females without comorbidities aged 65 and older had mortality rates of 363.1 per 10,000 in contrast to white females in the same stratum who experienced mortality rates of 35.9 per 10,000 (p<0.001). The relative difference in mortality rate is also stark when comparing Black and White individuals with comorbidities. Black males aged 65 and older with multiple
comorbidities had a mortality rate that was 3.5 times that of White males with multiple comorbidities (p<0.001).

*Geographical Distribution*

Across the 198 ZCTAs in which at least one Black and at least one White individual died from COVID-19, the median mortality rate for Black populations was 167.9 per 100,000 [IQR: 89.4-251.8] compared to the median mortality rate of 64.0 per 100,000 population [IQR: 34.5-97.0] among white populations (Figure 3.2). Among all 569 ZCTAs in which a COVID-19 related death took place, 253 ZCTAs reported ≥5 deaths.

We found that mortality rates were higher among Black individuals in 84.8% of these 198 ZCTAs (Figure 3.2, Figure 3.3). For those ZCTAs in which individuals of both races died, mortality rates were significantly higher among Black individuals, as determined by Wilcoxon signed-rank test (p<0.001). In these ZCTAs, Black individuals had a median of 2.3 times [IQR:1.3-4.8] greater odds of dying from COVID-19 than White individuals (Figure 3.3). Detroit (524,634) and Flint (51,686) have the largest Black populations in Michigan and were both devastated by COVID-19 mortality (Figure S.3.2).

*Date of Symptom Onset, Hospitalization, and Death*

The transmission intensity driven by exposure risk within a population is reflected by the timing of the peaks in symptom onset and mortality. Date of symptom onset was recorded for 4110 of the individuals in our study (67.8%), with missing data on 1955 individuals (32.2%). Among White and Black individuals, 28.2% and 39.1% of individuals are missing symptom onset date data, respectively. Among all COVID-19 decedents in Michigan, dates of symptom onset and death were earlier for Black populations in comparison to White populations (Figure 4), indicating a steeper escalation in risk for Black individuals. The peak in symptom onset for Black individuals in Michigan was 6 days earlier (March 26, 2020) than White individuals (April 1, 2020). Duration of hospitalization was 2 days longer for Black individuals (median of 10 days vs. 8 days) compared to White (p<0.001). The median onset
date and median date of death were 13 days earlier for Black individuals who died than White, and peak daily deaths occurred 3 days earlier for Black individuals than White (April 8, 2020 vs April 11, 2020).

Prompt medical attention is an important component of recovery. The time between symptom onset and hospitalization was a median of 1 day longer for Black decedents compared to White (4 days vs 3 days). Upon fitting negative binomial distributions to the time from symptom onset to hospitalization, we observed a median of 2 days (mean: 4.51 days) for White individuals and a median of 3 days (mean: 5.05 days) for Black individuals (Figure S.3.3).

Discussion:

Our highly granular analyses demonstrate that at the individual, ZCTA and state levels, racial disparities in the burden and impact of COVID-19 are salient and reflective of pervasive inequities. By accounting for pre-existing comorbidities, we show that these underlying health conditions alone do not explain the racial disparities in COVID-19 mortality. In fact, when adjusting for age and health status, the gap becomes even more striking. Overall, Black populations in Michigan have 3.6 times the mortality rate of White populations. When subset to individuals under 65 years old and without comorbidities, Black individuals have 12.6 times the mortality rate of White individuals. As we consider the substantial disparities highlighted by the COVID-19 pandemic in Michigan, we can see similar patterns elsewhere in the US and around the world.37 In the United Kingdom (UK), a study on over 17 million adults documented increased COVID-19-related mortality risk among those of Black and South Asian race compared to White people (Hazard ratio 1.48, 1.30–1.69 and 1.44, 1.32–1.58, respectively), even after adjusting for relevant risk factors.38 Our findings are consistent with race-based COVID-19 mortality risk results from other studies in the UK and the US.10,39–41

The racial disparity in mortality rate is most pronounced at younger ages. The significantly younger median age of Black deaths (72 vs. 81 years) results in even more years
of life lost in this community than the differential mortality rate alone would suggest. Among deceased individuals, peak incidence occurred six days earlier among Black individuals. Although this data is a subset of all exposures and there is incomplete data on symptom onset date, our finding suggests that Black individuals may experience heightened exposure to COVID-19. We found that White decedents overall and particularly those of older age were more likely to have been living in high-risk congregate living facilities than Black decedents. This may indicate that White decedents were exposed primarily due to their living circumstances in long-term care and assisted living facilities. As Black decedents were younger overall, this population's exposure may be attributed to the vulnerabilities of individuals unable to participate in stay-at-home orders during the pandemic. For instance, the flexibility to work from home is often linked to higher income and to industries in which Black individuals are underrepresented. While 47 to 49% of White individuals report being able to work from home, only 34 to 39% of Black individuals have the same privilege. Furthermore, Black Americans are disproportionately employed in low-wage and high-contact essential service industries within which sick leave is often discouraged and uncompensated. Compounding these issues, low-wage earners have more children with an average of 2.4 dependent family members. In the absence of sick leave benefits under these circumstances, infectious individuals in essential service industries are more likely to spread COVID-19 to their disproportionately Black coworkers. Similar factors may be driving the higher mortality rate observed in Black males compared to Black females, in addition to the significantly higher risk of COVID-19 mortality experienced by men compared to women and the higher baseline mortality rate experienced by Black men.

Our results regarding the duration between symptom onset and hospitalization indicate that Black COVID-19 patients do not receive medical attention as promptly as their White counterparts, a factor which is known to influence survival. Financial barriers and distrust deterring care-seeking, test scarcity, and racial bias among healthcare providers may be driving this delay, among other factors. In the US, Black and Hispanic/Latinx
populations represent 12.3% and 17.6% of the total employed population, and make up 25% to 40% of those employed in industries experiencing the highest proportion of job losses due to COVID-19.30,51 The disproportionate burden of unemployment in these communities, and the precarious linkage of health insurance to employment in this country, may be serving as a deterrent to care-seeking when COVID-19 symptoms arise. As families struggle to endure job losses with limited resources, the COVID-19 pandemic places the most vulnerable in the precarious situation of having to decide between staying at home and safeguarding against infection or seeking out employment as a means for financial survival. Furthermore, with the longstanding history of unethical treatment in public health and medicine, new challenges may arise as medical advances become available. Individuals may distrust the medical system due to historical unethical and inhumane treatment of marginalized populations, which could delay treatment.43 Upon seeking medical attention, those living in densely populated inner cities have faced inadequate access to testing.52 These issues interplay with the racial composition of an area. Detroit, the most populous city in Michigan, is also home to a larger proportion of Black individuals than any other city in the United States. Additionally, the SARS-CoV-2 test used in Detroit during the early months of the pandemic was found to miss 45% of positive cases.53,54 Patients receiving these false negative results are likely to have received delayed or inappropriate care, compounding the other factors described above. Finally, physician racial biases may affect patient health outcomes.50,55

Racial disparities in access to healthcare, educational opportunities and economic security predate the COVID-19 epidemic. At least 90% of the ZCTAs within the predominantly Black cities of Detroit and Flint have rates of child poverty that are higher than the national average.56 The anti-Black systemic racism in our political system is paralleled by the racialized nature of COVID-19.43 Inadequate governance nationwide has hindered pandemic response, analogous to the ongoing mishandling of the Flint water crisis that has arisen from a series of misguided governmental, social, and economic policies.57 The effects of COVID-19 shown here highlight a need for corrective resource allocation by federal and local governments that would mitigate the toll of public health crises on vulnerable
populations. Universal healthcare, living wages for all workers, and paid sick leave are essential public health interventions for addressing racial inequities in the US.

Moreover, public health must also concentrate efforts to target and eliminate pathways to inequality. In the face of the COVID-19 pandemic, institutionalized drivers of inequality such as the prison industrial complex have highlighted the glaring disparities affecting racial minorities. In 2018, the State of Michigan had 8,604 commitments. While breakdowns for specific races are not available, a total of 47.0% of the male prison population and 32.2% of the female prison population were non-White despite the state’s population being 79.2% White. Estimates from research centers suggest Black individuals comprise 53% of the prison population and 37% of the jail population while only comprising 15% of the state population. With overcrowding, inability to maintain social distancing, reduced access to soap or sanitizers, and reduced availability of personal protective equipment, incarceration presents an avenue for widespread transmission of disease in the Black prison, jail and detention center populations as a whole. As public health practitioners, we must work diligently to eliminate such drivers of inequality and work to fundamentally restructure prison systems as a public health intervention to reduce disparities and promote health.

Over the course of the epidemic in Michigan, several public health measures were implemented to curb transmission. A statewide stay-at-home order was placed between March 24 and June 1, 2020, alongside business restrictions such as prohibiting dining at restaurants and large gatherings. Other factors such as household size, employment in essential services, use of public transportation, and adherence to mask-wearing and stay-at-home orders can also impact infection exposure as well as mortality risk. In Detroit, the most impacted area in Michigan, social distancing in April was adhered to by 48.3% of non-Black individuals and 41.7% of Black individuals. These measures likely contributed towards limiting exposure to infection which may have impacted the eventual decline in mortality after mid-April. However, due to lack of sufficient data, we were unable to integrate these characteristics into
our study framework. We were only able to deduce time to hospitalization for a subset of all
decedents due to missing data on symptom onset date, hospitalization date or both. There was
also heterogeneity in missingness of this data across races, indicating a possible association
with delays in receiving care. By conducting a Chi-Square test (Table S.3.4), we compared
the demographic profile of decedents for whom time to hospitalization could not be
ascertained from data with the rest of the decedents. The demographic profile of individuals
with missing time to hospitalization best resembled individuals who were hospitalized within
3 days of symptom onset. Therefore, it may be unlikely that missing data is associated with
delays in hospitalizations, but a deeper understanding of this relationship requires additional
evidence.

There is also a complex relationship between economic deprivation and systemic
racism in terms of COVID-19 mortality risk. These underlying factors could provide deeper
insight into the drivers of racial disparities in COVID-19 mortality across Michigan, as well
as other US states. Under-reporting of COVID-19 may also be driven by racial disparities,
which would exacerbate the current situation. To address this issue, further analysis and data
is required to investigate the relationship between COVID-19 testing availability and
healthcare accessibility throughout the state. While our analysis of COVID-19 deaths is
specific to Michigan, our findings of increased mortality rates after stratifying for key
demographic and health variables are consistent with the increased mortality burden
experienced by Black individuals nationwide.4 The mortality of COVID-19 is skewed towards
the elderly and those with one or more comorbidities, as also observed in our data. Our study
was limited by the absence of population-level comorbidity data for those under age 65 in the
state of Michigan, and thus national-level estimates were used instead. However, given the
similar comorbidity burdens observed nationally and in Michigan for those 65 and older,33
this is unlikely to alter our overall findings.

As the COVID-19 pandemic continues to unfold, it is imperative to ameliorate
disparities as both a pandemic response and prevention measure. Systemic oppression of
Black individuals in the US underlies circumstances leading to racial disparities in COVID-19 exposure, the timeliness of treatment and case fatality rates. While structural racism as an infectious disease risk factor is increasingly being recognized, the US response to this pandemic demonstrates our society’s continued negligence toward the wellbeing of all Americans. It is not only the longstanding increased risk of comorbidities in the Black population that drives racial disparities in COVID-19 mortality, but socio-political and economic factors that impact COVID-19 exposure and medical care. Rectifying these inequities is urgent both to limit the devastation of COVID-19 and to protect against future public health crises.

Acknowledgements:
We would like to thank Seth Eckel and Adam Hart for linking Vital Records and MDSS data, and to remember all those in Michigan who lost their lives to COVID-19.

Author Contributions:
ASP and APG conceived the study. LM, JD, and JC collected the data. ASP, AP, IM, and CRW designed the study. ASP conducted data analysis. ASP drafted the manuscript with contributions from AP, IM, CRW, and AES. MCF and APG revised the manuscript. AP and APG supervised the research. All authors reviewed the manuscript and agreed to be responsible for all aspects of the work.

Data Sharing:
Individual-level data used in this study cannot be made publicly available but aggregated COVID-19 data provided by the Michigan Department of Health and Human Services is freely available to the public. ASP, AP, IM, and LM had access to the data used in the study.

Declaration of interests: All authors have nothing to disclose.

Funding: ASP, AP and APG were funded by NSF Expeditions grant 1918784, NIH grant 1R01AI151176-01, and NSF Rapid Response Research for COVID-19 grant RAPID-2027755. MCF was supported by NIH grant K01AI141576.
Table 3.1 Demographic characteristics of individuals who died from COVID-19 in Michigan overall and by race.

<table>
<thead>
<tr>
<th>Race</th>
<th>Black (n = 2341 (38.6%))</th>
<th>White (n = 3497 (57.7%))</th>
<th>Overall* (n=6065)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1274 (54.4)</td>
<td>1773 (50.7)</td>
<td>3186 (52.5)</td>
<td>0.006</td>
</tr>
<tr>
<td>Female</td>
<td>1067 (45.6)</td>
<td>1724 (49.3)</td>
<td>2879 (47.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Age, years (median [IQR])</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>44 (1.9)</td>
<td>26 (0.7)</td>
<td>74 (1.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>40-65</td>
<td>602 (25.7)</td>
<td>399 (11.4)</td>
<td>1053 (17.4)</td>
<td></td>
</tr>
<tr>
<td>65-80</td>
<td>1011 (43.2)</td>
<td>1170 (33.5)</td>
<td>2283 (37.6)</td>
<td></td>
</tr>
<tr>
<td>≥80</td>
<td>684 (29.2)</td>
<td>1902 (54.4)</td>
<td>2655 (43.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Congregate Living Facility Resident or Employee</strong></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Retired or Unemployed</td>
<td>302 (12.9)</td>
<td>776 (22.2)</td>
<td>1128 (18.6)</td>
<td></td>
</tr>
<tr>
<td>Number of Comorbidities</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>None</td>
<td>675 (28.8)</td>
<td>566 (16.2)</td>
<td>1293 (21.3)</td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>428 (18.3)</td>
<td>636 (18.2)</td>
<td>1116 (18.4)</td>
<td></td>
</tr>
<tr>
<td>Two or more</td>
<td>1238 (52.9)</td>
<td>2295 (65.6)</td>
<td>3656 (60.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Comorbidities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma or Reactive Airway Disease</td>
<td>149 (6.4)</td>
<td>170 (4.9)</td>
<td>341 (5.6)</td>
<td>0.016</td>
</tr>
<tr>
<td>Cardiovascular Disease</td>
<td>1235 (52.8)</td>
<td>2293 (65.6)</td>
<td>3647 (60.1)</td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td>191 (8.2)</td>
<td>425 (12.2)</td>
<td>639 (10.5)</td>
<td></td>
</tr>
<tr>
<td>Chronic Lung Disease</td>
<td>369 (15.8)</td>
<td>826 (23.6)</td>
<td>1226 (20.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diabetes Mellitus</td>
<td>744 (31.8)</td>
<td>978 (28.0)</td>
<td>1819 (30.0)</td>
<td></td>
</tr>
<tr>
<td>Neurologic Disease</td>
<td>460 (19.6)</td>
<td>1242 (35.5)</td>
<td>1741 (28.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Chronic Liver Disease</td>
<td>64 (2.7)</td>
<td>76 (2.2)</td>
<td>149 (2.5)</td>
<td>0.199</td>
</tr>
<tr>
<td>Chronic Renal Disease</td>
<td>498 (21.3)</td>
<td>637 (18.2)</td>
<td>1181 (19.5)</td>
<td>0.004</td>
</tr>
<tr>
<td>Other Immunosuppressive Conditions</td>
<td>134 (5.7)</td>
<td>204 (5.8)</td>
<td>348 (5.7)</td>
<td>0.906</td>
</tr>
<tr>
<td>Other Chronic Diseases</td>
<td>215 (9.2)</td>
<td>505 (14.4)</td>
<td>741 (12.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Relevant Comorbidity Combinations</strong>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes Mellitus, Cardiovascular, and Chronic Renal Disease</td>
<td>214 (9.1)</td>
<td>244 (7.0)</td>
<td>488 (8.0)</td>
<td>0.003</td>
</tr>
<tr>
<td>Diabetes Mellitus and Cardiovascular Disease</td>
<td>576 (24.6)</td>
<td>828 (23.7)</td>
<td>1480 (24.4)</td>
<td>0.435</td>
</tr>
<tr>
<td>Cardiovascular and Chronic Lung Diseases</td>
<td>281 (12.0)</td>
<td>660 (18.9)</td>
<td>964 (15.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cardiovascular and Neurologic Disease</td>
<td>353 (15.1)</td>
<td>935 (26.7)</td>
<td>1312 (21.6)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

* Races and ethnicities of decedents presented in the column labeled Overall (n=6065) are White (57.7%), Black (38.6%), Native American (0.3%), South Asian (0.5%), Southeast Asian (0.4%), East Asian (0.2%), other Asian ethnicity (0.5%), and other not specified races and individuals for whom race was not reported (1.9%).

** The number of individuals living in congregate living facilities by age group, sex, and race is presented in Table S.3.1.

*** Comorbidity combinations included here represent those typical among patients with the following medical and behavioural conditions in order: uncontrolled diabetes, obesity, long-term smoker, and dementia. These conditions have been found to be associated with COVID-19 mortality.46
Data are median [IQR], n (%), or n/N (%), and p-values were calculated by Chi-Square tests or Kruskal-Wallis tests as appropriate. The study population overall includes deaths among individuals of all races, of which Black and White individuals make up 96.3%.
Figure 3.1 COVID-19 mortality per 10,000 population in Michigan by age, sex, number of comorbidities, and race.

The population was first stratified by individuals under 65 years (left) and 65 years and older (right) then by sex (male, female), number of comorbidities (none, one, multiple), and race (White and Black). These mortality rates are based on 6,065 COVID-19 deaths that occurred in the state of Michigan between March 16 and October 26, 2020. Differences between the mortality rate of Black and White populations are statistically significant for every comparison (p<0.001). Absolute numbers of deaths in each stratum are presented in Table S.3.2.
Figure 3.2: COVID-19 Mortality rates per 100,000 population among Black and White Michigan residents and Michigan residents overall by ZIP Code Tabulation Area (ZCTA).

Mortality rate per 100,000 population ranges from 0 (green) to 100+ (red). The highest mortality rate per 100,000 population is 5263. Dark grey regions indicate ZCTAs where no COVID-19 deaths for a particular race occurred and light grey regions indicate ZCTAs where no COVID-19 related deaths took place. These mortality rates are based on 6027 COVID-19 deaths among Michigan residents spread across the state between March 16 and October 26, 2020, of whom 5809 individuals are either White or Black. Total includes individuals of all races. The inset map represents the Detroit Metropolitan Area and Flint. See Supplement for the number of COVID-19 related deaths by ZCTA (Fig S1) and the population of Black individuals by ZCTA (Fig S2).
Figure 3.3: Mortality rate ratio for Black to White populations at the ZIP Code Tabulation Area (ZCTA) level of residence in Michigan.

There were 30 ZCTAs where deaths among Black individuals occurred yet the White COVID-19 mortality rate exceeded the Black mortality rate (green). The Black COVID-19 mortality rate exceeds White mortality rate in 178 ZCTAs (yellow to purple). Dark grey regions indicate ZCTAs where COVID-19 deaths occurred only among one race, Black or White. Light grey regions indicate ZCTAs where no COVID-19 related deaths took place. The inset map represents the Detroit Metropolitan Area and Flint. These mortality rates are based on 5809 COVID-19 deaths among Black and White Michigan residents spread across the state between March 16 and October 26, 2020.
Figure 3.4: Date of symptom onset and date of death over time among individuals who died from COVID-19 in Michigan by race.

Date of symptom onset (n=4110) and death (n=6065) are displayed from February 23 to October 23, 2020 and March 16 to October 26, 2020, respectively. Date of symptom onset was missing for 1955 individuals.
References


51 Employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity. 2020; published online Jan 22. https://www.bls.gov/cps/cpsaat11.htm (accessed June 1, 2020).


60 Simpson PL, Butler TG. Covid-19, prison crowding, and release policies. BMJ. 2020; 369: m1551.


Supplement: Chapter 3

Figure S.3.1: COVID-19 related deaths by ZIP Code Tabulation Area (ZCTA) and race.

Deaths displayed occurred between March 16 to October 26, 2020.
**Figure S.3.2:** Black population by Zip Code Tabulation Area (ZCTA) in Michigan.

Dark Grey represents ZCTAs with no black individuals and light grey indicates ZCTAs where no race-based population data was available.
Table S.3.1: Residents/Workers at High Risk or Congregate Living Facilities

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Sex</th>
<th>Number of Comorbidities</th>
<th>Deaths among Residents of and Workers at High Risk or Congregate Living Facilities</th>
<th>Race</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Black</td>
<td>White</td>
</tr>
<tr>
<td>&lt;65</td>
<td>Male</td>
<td>0</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>39</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>34</td>
<td>47</td>
</tr>
<tr>
<td>65+</td>
<td>Male</td>
<td>0</td>
<td>54</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>45</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>165</td>
<td>560</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0</td>
<td>54</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>46</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>217</td>
<td>759</td>
</tr>
</tbody>
</table>

* High Risk or Congregate Living Facilities include: long-term care homes, skilled nursing facilities, assisted living facilities, homeless shelters, federal prisons, Michigan Department of Corrections prisons, county jail, juvenile justice facilities, foster care, and others, including senior, retirement, and group homes.
Table S.3.2: COVID-19 Deaths in Michigan by age group, sex, number of comorbidities, and race from March 16 to October 26, 2020.

<table>
<thead>
<tr>
<th>Race</th>
<th>Sex</th>
<th>Number of Comorbidities</th>
<th>&lt;65</th>
<th>65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (n=6065)</td>
<td>Male</td>
<td>None</td>
<td>187</td>
<td>494</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One</td>
<td>172</td>
<td>429</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>352</td>
<td>1552</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>None</td>
<td>98</td>
<td>514</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One</td>
<td>84</td>
<td>431</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>234</td>
<td>1518</td>
</tr>
<tr>
<td>Black (n=2341)</td>
<td>Male</td>
<td>None</td>
<td>118</td>
<td>246</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One</td>
<td>96</td>
<td>156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>190</td>
<td>468</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>None</td>
<td>70</td>
<td>241</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One</td>
<td>48</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>124</td>
<td>456</td>
</tr>
<tr>
<td>White (n=3497)</td>
<td>Male</td>
<td>None</td>
<td>58</td>
<td>227</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One</td>
<td>67</td>
<td>253</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>143</td>
<td>1025</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>None</td>
<td>25</td>
<td>256</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One</td>
<td>31</td>
<td>285</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2+</td>
<td>101</td>
<td>1026</td>
</tr>
</tbody>
</table>
Table S.3.3. Calculating Proportion of population that has no comorbidities and 1 comorbidity.

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age Group (years)</th>
<th>0</th>
<th>1</th>
<th>≥1&lt;sup&gt;50&lt;/sup&gt;</th>
<th>≥2&lt;sup&gt;50&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>55 to 64</td>
<td>32.3%</td>
<td>35.4%</td>
<td>67.7%</td>
<td>32.3</td>
</tr>
<tr>
<td></td>
<td>≥65</td>
<td>17.0%</td>
<td>31.6%</td>
<td>83.0%</td>
<td>51.4</td>
</tr>
<tr>
<td>Female</td>
<td>55 to 64</td>
<td>28.9%</td>
<td>29.6%</td>
<td>71.1%</td>
<td>41.5</td>
</tr>
<tr>
<td></td>
<td>≥65</td>
<td>12.4%</td>
<td>28.2%</td>
<td>87.6%</td>
<td>59.4</td>
</tr>
</tbody>
</table>

Using data from the National Health Interview Survey on the percentage of adults with one or more and two or more chronic conditions by age group and sex,<sup>50</sup> we were able to calculate the percentage of individuals in each age group and sex with no comorbidities and with only one comorbidity. For individuals aged 65 and older, we used this information in combination with Medicare Beneficiary data on those 65 years and older from Michigan on the proportion of beneficiaries with 0 or 1 comorbidity and 2 or more comorbidities by race. We assumed the distribution of comorbidities for each sex do not vary by race and we created estimates for Michigan comorbidity distribution by age, sex, and race. For individuals under age 65 years, we calculated race-specific proportions by applying Medicare data by race to the National Health Interview Survey data by sex and number of comorbidities. Medicare data for the under 65 population was not used as the primary source for comorbidity burden distribution due to strict enrollment criteria in this age group (disability, end-stage renal disease, and Amyotrophic Lateral Sclerosis).<sup>51</sup>

As data on comorbidity distribution at the population level was not stratified by, age, sex and race simultaneously, we assumed same comorbidity distribution by sex for each race. In our COVID-19 death dataset, this distribution was consistent with our assumption. Specifically, among those who died from COVID-19 who were Black, the proportion of those who were male who had comorbidities (71.1%) was very similar to the proportion of those who were female who had comorbidities (70.3%). Likewise, among White COVID-19 descendants, the proportion with comorbidities among males (83.9%) and females (83.7%) were similar.
Figure S.3.3: Negative Binomial distribution fit to time from symptom onset to hospitalization for Black and White decedents.

Length of time from symptom onset to hospitalization (N=2955) was acquired only from individuals for whom both date of symptom onset (N=4110) and date of hospital admission (N=4110) were available. In the table below, we display the number of Black and White individuals below 65 years of age and aged 65 and older, who were noted as having less than 3 days, 3 to 4 days, 5 or more days, and a missing number of days between symptom onset and hospitalization (Time to Hospitalization). Using a Chi-Square test, we determined that there are significant differences between race and time to hospitalization-stratified groups. Both White and Black individuals with missing data on time to hospitalization appear to have
the most similar age structure to those individuals who had the shortest time (<3 days) between symptom onset and hospital admission. Using a z-test of equal proportions, we have found that the proportion of White and Black individuals aged 65 and older with missing time to hospitalization is not significantly different from the proportions in each race with a time to hospitalization of less than 3 days (White: p=0.9883 and Black: p=0.7223). We found that the proportion of male White and Black individuals with missing data on time to hospitalization is not significantly different from that of individuals with a time to hospitalization of less than 3 days (White: p=0.1199 and Black: p=0.6107) as well as 3 to 4 days (White: p=0.0812 and Black: p=0.08).

Table S.3.4: Exploration of missing data for Time to Hospitalization by race.

<table>
<thead>
<tr>
<th>Time to Hospitalization</th>
<th>&lt;3 days</th>
<th>3-4 days</th>
<th>5+ days</th>
<th>Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age group (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Black: 104 (17.1)</td>
<td>Black: 150 (33.0)</td>
<td>Black: 191 (10.8)</td>
<td>Black: 315 (25.3)</td>
</tr>
<tr>
<td>65+</td>
<td>White: 753 (89.3)</td>
<td>White: 327 (75.7)</td>
<td>White: 234 (85.4)</td>
<td>White: 134 (63.8)</td>
</tr>
<tr>
<td></td>
<td>Black: 104 (17.1)</td>
<td>Black: 150 (33.0)</td>
<td>Black: 191 (10.8)</td>
<td>Black: 315 (25.3)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>White: 428 (50.8)</td>
<td>White: 221 (51.2)</td>
<td>White: 146 (53.3)</td>
<td>White: 125 (59.5)</td>
</tr>
<tr>
<td></td>
<td>Black: 359 (58.9)</td>
<td>Black: 272 (59.8)</td>
<td>Black: 840 (47.4)</td>
<td>Black: 656 (52.7)</td>
</tr>
<tr>
<td>Female</td>
<td>White: 415 (49.2)</td>
<td>White: 211 (48.8)</td>
<td>White: 128 (46.7)</td>
<td>White: 85 (40.5)</td>
</tr>
<tr>
<td></td>
<td>Black: 250 (41.1)</td>
<td>Black: 183 (40.2)</td>
<td>Black: 931 (52.6)</td>
<td>Black: 588 (47.3)</td>
</tr>
</tbody>
</table>
Related causes of death and comorbidities included under the categories of:

*Other Chronic Conditions:*

Anemia, Chronic venous thromboembolism (deep vein thrombosis and pulmonary embolism),
Bipolar disorder, Osteoarthritis, Depression, Osteoporosis & osteopenia, Alcohol Abuse,
Fragile X syndrome, Hemochromatosis, Hepatitis C & Hepatitis B, Schizophrenia, Addison’s
Disease, Chronic pancreatitis, Chronic Wounds, Sickle Cell Disease.

*Other Immunosuppressive Conditions:*

Rheumatoid arthritis, Bullous Pemphigoid, HIV, Polymyalgia rheumatica, Inflammatory
bowel disease (Crohn’s disease & Ulcerative Colitis), Sarcoidosis, IgG4 Disease, Systemic
Lupus Erythematosus, Giant Cell Arteritis (or Temporal arteritis), IgA Deficiency, Psoriatic
arthritis, Solid Organ Transplant.

**Supplement References: Chapter 3**

50. Health Policy Data Requests - Percent of U.S. Adults 55 and Over with Chronic
Conditions. 2019; published online Feb 7.
https://www.cdc.gov/nchs/health_policy/adult_chronic_conditions.htm (accessed July 2,
2020).

51. Medicare for People Under 65. https://medicareadvocacy.org/under-65-project/ (accessed
Conclusion

In this dissertation, I used statistical and mathematical methods to assess the impact of outbreaks, epidemics, and pandemics on disproportionately affected groups in the United States and Cameroon. I analyse the spatiotemporal dynamics of a measles outbreak, provide insights on the projected impact of an ongoing intervention to improve HIV prevention and treatment uptake, and highlight the inequitable COVID-19 mortality burden experienced by Black individuals in Michigan.

In Chapter One, I find that population movement is an important driver of large-scale measles outbreaks in Cameroon, and that its contribution to disease transmission leads to highly heterogeneous incidence. These findings demonstrate the need to improve our understanding of the roles of population mobility and local heterogeneity of vaccination coverage in the spread and control of measles in Cameroon.

In Chapter Two, I determine that achieving the federal government goals for prevention, diagnosis, treatment, and viral suppression in priority jurisdictions will significantly reduce HIV incidence. A combination of improving performance on the HIV care continuum and increasing PrEP uptake as proposed in the EHE plan will be essential for turning the tide on the HIV epidemic. With concurrent efforts to erase stigma, restructure criminalization and health insurance, and undo our previous and ongoing oppression of racial, ethnic, sexual, and gender minority communities that bear the brunt of the epidemic, ending the national epidemic can become a reality.

In Chapter Three, I examined characteristics of individuals who died from COVID-19 in Michigan by race stratified by their age, sex and comorbidity prevalence to illustrate and understand disparities in mortality risk. I found that prevalence of comorbidities, age, and sex alone do not account for disparities in mortality inflicted upon Black individuals in Michigan. Ultimately, I highlight that the racial disparities in the burden and impact of COVID-19 are reflective of pervasive historical and ongoing inequities.

Future research will aim to evaluate the impact of targeted increase of vaccination coverage on reducing measles outbreaks and spread of disease between regions, departments,
and health districts across Cameroon. For the HIV epidemic in the United States, future work will involve calculating the cost-effectiveness and performing an analysis of lives and years of lives saved upon improving performance on the HIV care continuum. Furthermore, this model could be expanded to account for sub-populations with differing risk of HIV infection and access to care specific to each jurisdiction. This stratification would allow us to account for variability in incidence and treatment uptake by age, race, injection drug use status, and other factors associated with disproportionate risk. As the COVID-19 pandemic continues to unfold, we must consider how we can address inequity as both a pandemic response and prevention measure. Future work will involve quantifying the proportion of cases attributable to COVID-19 mortality risk factors. Additionally, accounting for the ability of individuals who died from COVID-19 to have physically distanced prior to infection, considering their employment and essential worker status is an essential next step. Investigating the comorbidities listed in hospitalization data of COVID-19 decedents and survivors would help identify high-risk populations for targeted interventions such as priority vaccine dissemination. Lastly, a study of mortality risk focused on other vulnerable populations in Michigan must be conducted, specifically centred on incarcerated individuals and long-term care homes.

Ultimately this body of research demonstrates the importance of reducing differential risk in infectious disease burden. Coordinated efforts to expand preventative services to the areas in greatest need, data-driven investment in enhancing access to treatment programs for those infected, and restructuring the systems that maintain disparities will collectively reduce the impact of outbreaks, epidemics, and pandemics.