Assessing local health outcomes using spatially-resolved health surveillance data



Nathaniel Henry

St. Hugh's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2021

Abstract

Complete and accurate health information systems are necessary inputs for effective health policy. Although many countries maintain civil registration and infectious disease surveillance systems, variation in data completeness often impedes analysis of aggregated health records at the spatial level where policymaking occurs, discouraging greater investment in these systems.

This thesis aims to integrate health information systems into local health decision-making using spatial modelling approaches. In four case studies, I introduce a class of spatial statistical models based on incomplete vital and health surveillance records that offer insights into the health of a country that would be impossible to derive from other sources.

I demonstrate how civil registration and vital statistics data (CRVS) can be synthesised with supplementary spatial data sources to estimate both neonatal mortality and CRVS completeness by municipality across Mexico. This spatial modelling strategy can be applied to a wide array of health outcomes, including infectious diseases. I demonstrate how an analogous model can combine data from a tuberculosis (TB) prevalence survey and TB case notifications to estimate TB prevalence across Uganda.

Complete registration of births and deaths ensures that all citizens receive the same legal and health protections. I take a holistic approach to analyse India's three health surveillance systems in relation to the Indian National Health Plan's child survival goals.

The COVID-19 pandemic has highlighted gaps in health surveillance capacity worldwide, including in high-income countries. In Italy, I develop a small-area excess mortality model to estimate the number of misdiagnosed COVID-19 deaths during the first six months of the pandemic. This analysis reveals important information about the mortality dynamics of the pandemic across sub-populations of Italy.

The results, limitations, and conclusions of these case studies are discussed with recommendations for how these findings influence our understanding of health information systems and implications for greater integration between health surveillance data and policy.

Statement of contributions

My supervisors, Catrin Moore, Pete Gething, and Simon Hay, provided methodological and scientific guidance throughout all stages of this work. Additional contributors for each chapter are listed below.

For Chapter Two, the Mexican National Institute for Geography and Statistics (INEGI) provided data on births and neonatal deaths by municipality. I developed the research question, extracted all data, and developed the underlying statistical model. Bernardo Hernandez provided guidance on factors influencing birth and death registration completeness across Mexico. Michael Chipeta provided guidance for the tuning and validation of the statistical model.

For Chapter Three, data was provided by the Uganda National Tuberculosis and Leprosy Control Programme (NTLP). Tim Lucas, Ewan Cameron, and Dan Weiss provided feedback on early iterations of the joint statistical model. Jennifer Ross at the University of Washington provided guidance on tuberculosis burden estimation methods. Michael Chipeta reviewed this chapter and provided guidance for the sensitivity analyses presented in Appendix B.

The research questions addressed in Chapter Four emerged from a study led by Rakhi Dandona at the Public Health Foundation of India (PHFI), for which I developed the spatial modelling approach: Dandona, R., Kumar, G. A., Henry, N., *et al.* (2020). Subnational mapping of under-5 and neonatal mortality trends in India: the Global Burden of Disease Study 2000–17. *The Lancet*, 395(10237), 1640–1658. I developed the research questions and analysis strategy for Chapter Four in conversation with Catrin Moore and Simon Hay. For the analysis in this chapter, I extended code that was developed by the Local Burden of Disease team at the Institute for Health Metrics and Evaluation (IHME) with assistance from Michael Collison. Stefanie Watson at IHME extracted survey data that was used in this work. Lalit Dandona at PHFI provided guidance on the relationship between health data and policy in India. I produced the model estimates, projections, all figures, and comparisons between data sources presented in this chapter.

For Chapter Five, data was provided by the Italian National Institute of Statistics (Istat). Michael Chipeta, Ahmed Elagali, and Michele Nguyen provided guidance for the tuning and validation of the excess mortality statistical model. The work for Chapter Five is currently under review as: Henry, N. *et al.* Variation in COVID-19 excess mortality by age, sex, and province within Italy [submitted to *Scientific Reports*, June 2020]. I would like to thank all co-authors for their contributions to the manuscript and can confirm that the work produced in this thesis is my own. Other paper contributions alongside this thesis are detailed in Appendix E.

Acknowledgments

Every thesis is the culmination of a journey—mine was marked by detours, late nights, unexpected joys, and the greatest health crisis in a century. I am extremely grateful to Simon Hay for showing me the potential and the purpose of disease mapping. Thank you also to Pete Gething, who helped me refine a passion for seemingly intractable data into a concrete research project. To Catrin Moore, who offered me her unwavering support through the hardest days of my thesis: thank you for guiding the way.

I would also like to thank the many colleagues who have generously shared their time and knowledge with me. I have been enormously lucky to work with and learn from members of the Malaria Atlas Project, the Oxford AMR group, and the Local Burden of Disease team over the past three years. A special thanks to Michael Chipeta, who went above and beyond to help me grow as a spatial statistician and researcher. I am also enormously grateful to have worked with Jennifer Ross, whose deep knowledge of and passion for TB research continues to inspire me.

I am hugely appreciative of my office mates and friends who have been by my side, first in person and then in spirit, throughout my graduate studies. To my VR and U5M teams, thank you for being the salt of the earth; to my Torch Club friends, thank you for adding spice to life.

None of this would have been possible without the love and support of my family. Mom and Dad, thank you for always believing in my adventures. Lizzie, thank you for keeping it real.

Finally, I dedicate this thesis to my fiancée Ellen. Four years ago, at a pivotal moment, she asked me what I wanted to work towards for the rest of my life—this is the start of my answer. Thank you for everything since and for everything to come.

Contents

Abstract ii

Statement of contributions iii

Acknowledgments iv

Table of contents viii

List of figures xiii

List of tables xv

List of abbreviations xvi

1 Introduction 1

1.1 Health system performance: the urgent need for better data . . . 1

1.2 National health data sources: history and uses . . . 3

1.2.1 Civil registration and vital statistics (CRVS) systems . . . 3

1.2.2 Infectious disease surveillance . . . 5

1.2.3 Household surveys . . . 7

1.2.4 Special cases . . . 8

1.2.5 Standards for quality and usability . . . 9

1.3 Relationship between health data availability and health system capacity . . . 10

1.4 Spatial modeling as a tool for responsive health interventions . . . 11

1.4.1 Disease mapping and small area estimation . . . 12

1.4.2 Applications of spatial modelling . . . 13

1.4.3 Limitations . . . 15

1.5 New hierarchical modelling techniques for mapping health outcomes using incomplete health surveillance data . . . 17

1.6 Thesis structure . . . 20

1.7 References . . . 21

2 Joint estimation of neonatal mortality and vital registration completeness across Mexico 25

2.1 Introduction . . . 25

2.1.1 Estimating completeness of birth and death registration across Latin America . 26
2.2 Methods . . . 29
2.2.1 Data preparation . . . 30
2.2.1.1 Grouping of municipalities by social exclusion . . . 31
2.2.2 Joint estimation of neonatal mortality and CRVS data completeness . . . 32
2.2.3 Simulation model . . . 36
2.2.4 Application to neonatal mortality data in Mexico . . . 37
2.3 Results . . . 38
2.3.1 Predictive validity from simulation . . . 38
2.3.2 Neonatal mortality across Mexico . . . 38
2.3.2.1 Relationship between social marginalisation and covariates predictive of mortality . . . 38
2.3.2.2 Joint estimation model of neonatal mortality and CRVS bias . . . 39
2.4 Discussion . . . 41
2.4.1 Performance of the simulation model . . . 41
2.4.2 Relationship between neonatal mortality and vital statistics performance . . . 45
2.4.3 Model limitations . . . 46
2.4.4 Conclusions . . . 46
2.5 References . . . 47

3 Mapping the relationship between tuberculosis burden and case notifications in Uganda 51
3.1 Introduction . . . 51
3.1.1 TB mapping in high-burden settings . . . 53
3.2 Methods . . . 55
3.2.1 Data preparation . . . 56
3.2.2 Estimated duration and incidence-prevalence ratio of TB . . . 59
3.2.3 Small area model incorporating prevalence data and incomplete notifications . 63
3.3 Results . . . 65
3.3.1 Spatial variation in tuberculosis prevalence across Uganda . . . 65
3.3.2 Completeness of case notifications increases over time . . . 67
3.3.3 Effect of including notifications on estimates of TB burden . . . 69
3.4 Discussion . . . 71
3.4.1 Extension to other contexts . . . 73
3.4.2 Limitations . . . 74
3.4.3 Conclusions . . . 75
3.5 References . . . 75

4 A space-time-age model for subnational child mortality estimation in India 79
4.1 Introduction . . . 79
4.1.1 Mortality across Indian states: progress, transition, and inequality . . . 80
4.1.2 Tracking local variation in mortality across India . . . 81
4.2 Methods . . . 84
4.2.1 Geo-locating complete birth history data . . . 84
4.2.2 Space-time-age mortality estimation model . . . 85

4.2.3 Mortality forecasting . . . 88
4.2.4 Comparisons to SRS data at the most detailed spatial level available . . . 88
4.3 Results . . . 88
4.3.1 Progress and local disparities in child survival . . . 89
4.3.2 Comparison to 2025 and 2030 mortality targets . . . 94
4.3.3 Comparing local estimates from SRS and survey-based estimates . . . 95
4.4 Discussion . . . 98
4.4.1 Scale and inequality in mortality . . . 100
4.4.2 Limitations of survey-based spatial mortality mapping . . . 101
4.4.3 Opportunities offered by spatial reporting of CRS and SRS data . . . 103
4.4.4 Conclusions . . . 105
4.5 References . . . 105

5 Variation in COVID-19 excess mortality by age, sex, and province within Italy 108
5.1 Introduction . . . 108
5.2 Methods . . . 111
5.2.1 Overview . . . 111
5.2.2 Data . . . 111
5.2.3 Space-time model . . . 113
5.2.4 Compiling and interpreting results . . . 115
5.3 Model validation . . . 116
5.3.1 Data availability . . . 116
5.4 Results . . . 117
5.4.1 Excess mortality exceeds registered COVID-19 deaths . . . 117
5.4.2 Excess deaths and sex . . . 117
5.4.3 Spatial concentration . . . 121
5.4.4 Diverse experiences of the first wave . . . 122
5.5 Discussion . . . 124
5.6 References . . . 127

6 Discussion 130
6.1 Chapter summary . . . 131
6.2 Methodological strengths . . . 135
6.3 Methodological limitations . . . 136
6.4 Future directions . . . 138
6.5 Conclusions . . . 139
6.6 References . . . 140

A Appendix to Chapter Two 144

B Appendix to Chapter Three 157

C Appendix to Chapter Four 165

D Appendix to Chapter Five 170

E Publications and conference proceedings **175**

F Code repositories **178**

Bibliography **178**

List of Figures

1.1 Estimated completeness of government mortality records in 2015 or the most recent year when surveillance data was available. No mortality records were available for countries symbolised in white. 5

2.1 A visual demonstration of capture-recapture analysis using linked records from two surveys. 28

2.2 Summary of data sources, analyses, and results described in Chapter 2. NMR: neonatal mortality rate, CRVS: civil registration and vital statistics. 30

2.3 Grouping of municipalities into exclusion categories based on indigenous status, proximity to clinics, adult employment in the formal economy, and literacy rates among adult women. 33

2.4 $Left$: Simulated neonatal mortality rates across Mexican municipalities compared to neonatal mortality rates recovered from the spatial statistical model. Points indicate the mean NMR estimated by the model, while vertical line spans represent the 95% uncertainty intervals for NMR estimated by the model for each municipality. $Right$: CRVS bias terms simulated for model input data, on the x-axis, compared to the mean and 95% uncertainty intervals for these bias parameters recovered by the spatial statistical model, on the y-axis. 39

2.5 Distribution of socio-economic indicators across municipalities within each exclusion grouping. The centre line of each bar represents the median value for each indicator across less marginalised (N=1852), moderately marginalised (N=437), and severely marginalised (N=152) municipalities across Mexico, while the vertical range of each bar represents the inter-quartile range of indicator values for each grouping. 40

2.6 Neonatal mortality rate per 1,000 live births by Mexican municipality, 2009-2010, estimated by a joint model that incorporates both census and CRVS data. 42

2.7 Mean estimates for CRVS bias terms predicted by the joint model for neonatal mortality. Municipalities shaded in green indicate under-registration of births relative to neonatal deaths, leading to inflated estimates of mortality based on raw CRVS data. Municipalities shaded in brown indicate under-registration of neonatal deaths relative to births, leading to under-estimation of mortality from raw CRVS data. 43

2.8 Comparison between a joint estimation model with CRVS bias terms and a baseline model where CRVS bias is set to zero. *Panel A* : The two models estimate minor differences in the neonatal mortality rate by Mexican state, with overlapping 95% uncertainty intervals for all states. *Panel B* : The two models generate substantially different and non-overlapping NMR estimates for municipalities in Guerrero, Oaxaca, and Chiapas states in the far south as well as Yucatan state in the far east. 44
3.1 Regions and population centers of Uganda. 55
3.2 Summary of data sources, analyses, and results described in Chapter 3. 56
3.3 Unadjusted prevalence data from the 2014-2015 Uganda National TB Prevalence Survey, aggregated by district. Districts not containing any sampled clusters are symbolised in grey. 57
3.4 Reported case notification rates for pulmonary TB among adults aged 15 and above in 2017 and 2019, the two years for which notification data was available at the district level. 59
3.5 Results from the joint model compared to underlying data sources. *Panel A* : TB prevalence estimated from the joint data model. *Panel B* : scatter plot comparing modeled estimates, with uncertainty, against raw TB prevalence survey data in 56 districts. Panel C: scatter plot comparing modeled estimates by district, with uncertainty, against TB case notification rates among adults averaged across the years 2015-2019. . 66
3.6 Mean estimated completeness for district-level TB case notifications in 2017 and 2019, as predicted by the joint data model. Completeness is estimated for pulmonary TB cases in people aged 15 and above. 68
3.7 Average duration of active pulmonary TB among adults by Ugandan district, estimated as a function of case notification completeness. 69
3.8 Effect of adding notifications to the joint estimation model. Panel A: Mean estimated TB prevalence based on a spatial model incorporating only data from the Uganda National TB Prevalence Survey. Panel B: Difference in mean estimated TB prevalence between the joint data model and a model that only incorporated prevalence data. Panel C: Effect of adding case notifications, taking uncertainty into account. 70
3.9 Comparison of estimates from the joint data model (left) and a model using only prevalence survey data (right) against a low prevalence threshold of 253/100,000 and a high threshold of 506/100,000. A high-confidence estimate indicates that the 95% uncertainty bounds uniformly pass a threshold, while a low-confidence estimate indicates that the model's mean estimate in a district passes a threshold but the 95% uncertainty bounds include both sides of that threshold. 72
4.1 Summary of data sources, analyses, and results described in Chapter 4. 85
4.2 Sample size for under-5 mortality modeling based on survey data between 2000 and 2017 across India. The sample size for each year is approximated here as the number of individuals under age 5 entering each year according to retrospective complete birth history data. At the time of analysis, no complete birth history data was available after 2016, the year of the most recent included household survey. 86

4.3 NMR, IMR, and U5MR per 1,000 live births in 2000 and 2017, estimated from survey data. The Indian National Health Plan 2017 calls for reducing neonatal mortality to less than 16 deaths per 1,000 live births, and under-5 mortality to less than 23 deaths per 1,000 live births, nationwide by 2025. The color scale for mortality varies across age groups, with green centered around the NHP mortality targets for each age group. 90
4.4 Annualised rate of decline (ARoD) in U5MR between 2000 and 2017. An ARoD of 5% is equivalent to a cumulative decline of 58% over 17 years, while an ARoD of 1% is equivalent to a cumulative decline of 16% over 17 years. ARoD rankings by district were stable across age-specific indicators. 92
4.5 Absolute inequalities in U5MR, IMR, and NMR across districts within each Indian state and union territory in 2000 (grey) and 2017 (blue). Each dot represents a district: the lower bound of each vertical spread represents the district with the lowest U5MR in each state and year, while the upper bound of each vertical spread represents the district with the highest U5MR in the same district and year. The large diamond in each vertical spread shows the overall mortality rates across the state as a whole. A blue bar that is shorter than its grey counterpart, as evident across all age groupings in Madhya Pradesh and West Bengal, indicates that absolute between-district inequality has narrowed between 2000 and 2017. The horizontal dashed lines identify NHP targets for each indicator. 93
4.6 *Top* : Mean estimated U5MR by district in 2025, projected forward by extending the annualised rate of decline from 2000-2017. *Bottom* : Comparison between projected 2025 U5MR estimates and the Indian National Health Policy 2017 target of 23 under-5 deaths per 1,000 live births, accounting for uncertainty. Districts shaded in white have a 95% uncertainty interval (UI) that overlaps with the NHP 2025 target; districts shaded in red have a non-overlapping UI that falls above the target; districts shaded in blue have a non-overlapping UI that falls below the target. 96
4.7 Estimated U5MR by district in 2000, 2017, and projected to 2025 across three exemplar Indian states. Each vertical line displays the bounds of the 95% uncertainty interval in a given district and year, and the point within the vertical line displays the corresponding mean U5MR estimate for that district and year. Districts are sorted in decreasing order of their estimates U5MR in 2000. The dotted horizontal line displays the Indian NHP target of fewer than 23 under-5 deaths per 1,000 live births by 2025. 97
4.8 *A* : IMR per 1,000 live births in 2017 across 68 natural division in large Indian states, as estimated by the Indian SRS 2017 report. *B* : By aggregating survey-based estimates of IMR to the natural division level, estimates of IMR from SRS and survey sources can be compared. The ratio between SRS-based IMR estimates and survey-based mean IMR estimates are shown at the natural division level. *C* : Comparison between point estimates of IMR provided by the SRS with the 95% uncertainty interval for IMR estimated by survey data. Compared to survey-based estimates in 2017, the SRS estimated significantly lower IMR for 11 of 68 natural divisions, and significantly higher IMR for 3 of 68 natural divisions. 99
5.1 Registered COVID-19 deaths per 100,000 people by Italian region prior to August 31, 2020. Nationally, 35,500 deaths were registered during this time, or approximately 59 deaths per 100,000 population. 109

5.2 Summary of data sources, analyses, and results described in Chapter 5. 112
5.3 *Top* : Predicted baseline and observed deaths aggregated nationwide across Italy, March through August 2020. Bounds of the 95% uncertainty interval for predicted baseline deaths are shaded in light brown. *Bottom* : Estimated weekly excess deaths across Italy from March through May 2020, in red, compared with medically-certified deaths from COVID-19, in green. 118
5.4 *Left* : Map showing the ratio between the number of registered COVID-19 deaths in the northern regions of Italy and the total excess deaths in these regions, 26 February through 26 May 2020. *Right* : Line charts showing the total estimated excess deaths, in red, and the number of registered COVID-19 deaths, in green, by region and week. 119
5.5 Mean estimates for weekly excess deaths across Italy by age group, March through May 2020. Two nationwide age group breakdowns are shown to the right for comparison: the national age distribution as of 1 January 2020, and the proportion of nationwide deaths that occurred in each age group during the years 2015 through 2019. 119
5.6 Map of standardised mortality ratios by age group and province over the entire period of March through May 2020. In these figures, significance is defined as falling outside of the 95% uncertainty interval for baseline mortality for a given age group and province during this period. 121
5.7 Concentration of excess mortality burden among the 3, 7, and 22 provinces with the greatest number of excess deaths, out of 107 provinces total. These province groupings respectively accounted for over 25%, 50%, and 75% of the national excess mortality burden between March and May 2020. The graph on the bottom left displays mean estimated excess deaths by province grouping and week during this time period. . . . 123
5.8 Excess deaths vary across provinces within the Lombardy region. *Top left* : Map of Northern Italy showing the week when province's excess mortality was estimated to peak during the first wave of COVID-19 in March through May 2020. *Top right* : Map of the peak weekly standardised mortality ratio (SMR) for each province across Northern Italy across all weeks in March through May 2020. *Bottom left* : Estimated excess deaths by week for Lodi province. *Bottom right* : Estimated excess deaths by week for Varese province. 124
6.1 World map summarising the settings and themes of the four case studies presented in this thesis. 131
B.1 Reproduction of Figure 3.4 for a model with a fixed disease duration of 1 year. 161
B.2 Reproduction of Figure 3.5 for a model with a fixed disease duration of 1 year. 162
B.3 Reproduction of Figure 3.4 for a model with a fixed disease duration of 2.15 years. . . 163
B.4 Reproduction of Figure 3.5 for a model with a fixed disease duration of 2.15 years. . . 164
C.1 Map of gridded covariate showing travel time by motorised vehicle to the nearest settlement of 50,000 or greater in 2000 and 2017. Source: Malaria Atlas Project 166
C.2 Map of gridded covariate showing relative nighttime light intensity in 2000 and 2017. Source: Defense Meteorological Satellite Program 166
C.3 Map of gridded covariate showing the estiamted proportion of children under age 5 who received the DPT3 vaccine in 2000 and 2017. Source: IHME 167
C.4 Map of gridded covariate showing the estimated mean years of education attained for women aged 15-45 in 2000 and 2017. Source: IHME 167

C.5 Map of gridded covariate showing the estimated ratio between children below age 5 and women of childbearing age (15-45) in 2000 and 2017, used as an approximation of fertility. Source: Derived from WorldPop age-specific population estimates. 167
C.6 Map of gridded covariate showing urban land cover extent in 2000 and 2017. Source: Global Human Settlement Layer 168
C.7 Map of gridded covariate for estimated PM 2.5 air pollution concentration in 2000 and 2017. Source: IHME 168
C.8 Map of gridded covariate showing estimated *Plasmodium falciparum* incidence in 2000 and 2017. Source: Malaria Atlas Project 168
C.9 Map of gridded covariate showing the estimated proportion of children under age 5 who are moderately or severely stunted (low height for age) in 2000 and 2017. Source: IHME 169
C.10 Map of gridded covariate showing estimated population density in 2000 and 2017. Source: WorldPop 169

List of Tables

2.1 Penalised complexity prior specifications for the standard deviations of the CRVS bias terms in each of three municipality groupings. Column 3 specifies the range which 95% of municipality bias terms are likely to fall within if the standard deviation is equal to the threshold, U. Across all groupings, the true grouped standard deviations are estimated to exceed their associated thresholds with 5% probability. 36

2.2 Comparison between true underlying parameter terms and estimated values, with mean and 95% uncertainty intervals, from the simulation model. 38

3.1 Plausible duration of clinically active tuberculosis by HIV coinfection and treatment status, as published alongside the WHO Global Tuberculosis Report 2020. 60

4.1 Key features of mortality data sources in India. 82

5.1 Covariates used to estimate baseline mortality by age group, sex, province, and week across Italy from January 2015 through August 2020. The source and space-time resolution of each covariate is listed. 113

5.2 Estimated excess mortality by age and sex grouping, aggregated to the national level for the weeks of 26 February through 26 May 2020. Nationally, all age-sex groupings experienced significantly elevated mortality compared to the expected baseline during this period. 120

A.1 List of municipalities included in the moderately marginalised category, along with relevant indicators. 144

A.2 List of municipalities included in the severely marginalised category, along with relevant indicators. 153

B.1 Estimates of active TB duration as a function of case notification completeness, contingent on various assumptions about other factors influencing disease duration. The default scenario corresponds to the duration function applied in Chapter 3. All following scenarios change one parameter other than completeness. The duration equation is listed for each scenario, along with estimated values of duration at varying levels of case notification completeness (denoted as PI). 158

B.2 Results of sensitivity testing based on duration. Summaries for estimated TB prevalence and notification completeness are provided for each of the three model specifications. 159

D.1 In-sample predictive validity metrics comparing model estimates to individual data observations. RMSE: root mean squared error. RSE: relative squared error. 172

D.2 In-sample predictive validity metrics comparing model estimates to observed data by province-year grouping. RMSE: root mean squared error. RSE: relative squared error. 172

D.3 Out-of-sample predictive validity metrics comparing model estimates to individual data observations. RMSE: root mean squared error. RSE: relative squared error. 173
D.4 Out-of-sample predictive validity metrics comparing model estimates to observed data by province-year grouping. RMSE: root mean squared error. RSE: relative squared error. 173

List of abbreviations

AHS Annual Health Surveys, India

AIS AIDS Indicator Surveys

ART Antiretroviral therapy

BH Birth history

CAR Conditional autoregressive

CRS Civil Registration System, India

CRVS Civil registration and vital statistics

DHS Demographic and Health Surveys

DLHS District Level Household Surveys, India

DPT3 Diptheria, tetanus toxoid, and pertussis vaccine

EuroMOMO European Mortality Monitoring Activity

GLM Generalised linear modelling

HAQ Health Access and Quality Index

HDI Human Development Index

IHME Institute for Health Metrics and Evaluation

IID Independently and identically distributed

IMR Infant mortality rate

INEGI National Institute for Geography and Statistics, Mexico

INLA Integrated Nested Laplace Approximation

IQR Inter-quartile range

Istat Italian National Institute of Statistics

KEMRI Kenya Medical Research Institute

LMICs Low- and middle-income countries

MAP Malaria Atlas Project

MICS Multiple Indicator Cluster Surveys

MIS Malaria Indicator Surveys

NFHS National Family Health Surveys, India

NHP National Health Policy, India

NMR Neonatal mortality rate

NTLP National Tuberculosis and Leprosy Programme, Uganda

SCORE Survey, Count, Optimise, Review, Enable

SMR Standardised mortality ratio

SRS Sample Registration System, India

TB Tuberculosis

U5MR Under-5 mortality rate

UI Uncertainty Interval

UN United Nations

UN IGME United Nations Inter-agency Group on Mortality Estimation

UNICEF United Nations Children's Fund

USAID United States Agency for International Development

WHO World Health Organization

Chapter 1

Introduction

1.1 Health system performance: the urgent need for better data

Nearly 75 years after the United Nations Universal Declaration on Human Rights asserted the fundamental and universal right of all people to "a standard of living adequate to the health and well-being of himself and his family, including [...] medical care and necessary social services,"[1] the opportunity to live a full and healthy life varies vastly across the globe. Gaps in health between countries and regions of the world are well-documented. Health inequalities also manifest themselves within countries and even within the same town. These inequalities appear in the earliest years of life: in Kano state in northern Nigeria, children are 2.5 times more likely to die before their fifth birthday compared to children born in the capital, Lagos.[2] Inequalities also appear in high-resource settings, such as urban counties in the United States, where life expectancy for men can vary by nearly 20 years across neighbourhoods in a single U.S. county.[3] Inequalities in human flourishing from the local to the international level are deeply troubling; addressing these barriers is arguably the primary challenge of modern global health.[4]

Addressing unequal barriers to health naturally raises questions about which institutions have the responsibility and capability to develop interventions that can alleviate health burden. The discourse around health on the international stage has shifted over the past 20 years to a conception of "global health," where multilateral institutions are best positioned to coordinate health interventions.[5] While international agencies and funders have coordinated impressive responses to acute mortality

and disease threats over the past 20 years, sustainable health services must ultimately be delivered by health systems that are led by national stakeholders and operated locally.[6,7]

Health systems can be defined in terms of their human resources and material components: they are driven by health care workers who rely on a financial and material infrastructure that is, in turn, managed by governing institutions.[8] They can also be defined in terms of their key operations: the World Health Organization (WHO) lists service delivery, health workforce, information, medical products, vaccines and technologies, financing, and leadership and governance as the six core building blocks that constitute a functioning health system.[6] Within a health system, health policies may be negotiated and implemented at many scales, from individual healthcare workers up to national ministries of health.

While health systems necessarily develop in the context of local conditions and priorities, they share the unifying aim of improving the health of the people they serve. Therefore, any attempt to manage or improve healthcare must be measured against its potential impact on health outcomes.[8] Health policy-makers need to make decisions about efficiently allocating funding, prioritising at-risk groups, identifying and responding to health crises, and implementing long-term policy development and reform. All of these decisions require a consensual understanding among parties of conditions on the ground—that is, they require data.[9] Without data on health outcomes, other sociological and economic analyses can only describe, not drive, health policy.[8]

Complementary to their service provision activities, both international agencies and national bodies operate data collection and statistics systems that are designed to reveal actionable information about the state of health in a country. Many national ministries of health and statistics maintain civil registration and vital statistics (CRVS) operations to systematically register vital events such as birth and deaths, as well as surveillance systems for notifiable infectious diseases such as HIV, tuberculosis (TB), and measles. In high-resource countries, largely complete and high-quality health information systems facilitate epidemiological investigation and decision-making. In many lower-resource settings, where most childhood deaths and disease burden are concentrated, CRVS systems may be absent or incomplete, while infectious disease surveillance may be hindered by low completeness and reporting lags. Nearly all health surveillance systems in high-resource settings are operated electronically; in

lower-resource settings, health surveillance may be collected on paper records or through hybrid paper and electronic records, although many countries are gradually shifting towards greater digital integration in record-keeping.[10] To alleviate data gaps in countries with deficient health surveillance systems, international health institutions often fund household surveys that systematically collect key demographic and health data. These surveys can be further supplemented by health modelling approaches: notably, modern spatial statistical modelling can reveal local inequalities that may fall below the sampling frame of the original survey.[11]

This thesis asserts that high-quality CRVS and infectious disease surveillance are irreplaceable as a foundation for responsive health decision-making, and that these systems are therefore essential prerequisites for delivering sustainable and equitable health services to all people. The following sections of this introduction describe the operation of health data systems as well as the spatial modelling approaches that have been designed to supplement them in contexts where data is limited. Past scholarship on data governance in global health has expressed the concern that modelling approaches have crowded out demand for high-quality national health surveillance without offering the same benefits. In conclusion to this chapter, and more expansively throughout this thesis, I offer a partial solution to the problem of data governance: a statistical modelling framework that robustly incorporates deficient health surveillance records to measure health outcomes, estimates bias in the health surveillance system, and provides an incentive for its improvement.

1.2 National health data sources: history and uses

1.2.1 Civil registration and vital statistics (CRVS) systems

Civil registration and vital statistics (CRVS) systems facilitate the legal registration, compilation, and standardised dissemination of vital events records. Vital events comprise a wide variety of activities that change a person's legal status, including birth, marriage, separation, adoption, emancipation, legitimation, and death, among others. Of these, accurate registration of birth and death are both crucially important for the individual and for understanding population health.

In areas where vital registration is collected, birth and death records are often legally mandated to be recorded within a certain time window of the vital event. For births and deaths that occur in

a health facility, the event can often be registered with an on-site functionary; if this service is not available, the family members of the new-born or deceased individual may be required to register on-line or at a government office afterwards. If a country legally mandates cause-of-death reporting, an underlying cause of death must be medically registered at the time of death or verified afterwards through a process of interviews called verbal autopsy.[12] Local registration offices then report key aspects of the registration upwards to their regional and national counterparts, while private and religious health care facilities may compile and share their own records through separate channels. At the national level, these records are then checked for quality and compiled into regular statistical reports, which may tabulate detailed statistics by location, time, and age or cause grouping in the case of death.[13,14] These systems often rely on legal mandates to report rather than actively seeking out new births and deaths: this passive surveillance approach can present a problem in countries where vulnerable groups face greater barriers to vital event reporting.[15,16]

At the population level, detailed CRVS data can provide crucial information to health policy-makers; at the individual level, civil and death registration can provide rights and privileges to the registered. In many countries, valid birth certification is the key to accessing school, social services, and health insurance. Some countries have embraced the importance of universal CRVS: for example, the government of Mexico has declared that free and universal birth registration is a constitutional right of all Mexicans. In the case of death, family members of the deceased may be legally entitled to social and financial support once the death is registered.[13]

Despite its importance to governance and the individual, birth and death registration is often least functional in the countries where health burden is highest. As of 2004, fewer than 1 in 100 residents of South-east Asia and fewer than 1 in 10 Africans were covered by any birth or death registration.[17] Today, enormous gaps in service still remain. Figure 1.1, below, shows the estimated coverage of death registration among children under 5 in 2015 or the most recent year of data available. While almost all high-income countries experienced death registration completeness of over 90%, death registration coverage remained below 60% in all states of India. Peru, the Dominican Republic, and the north-western states of Brazil also had estimated coverage levels below 50%. In all sub-Saharan African countries besides Botswana and South Africa, no mortality estimates based on vital

registration were available after 2010.[18] A similar geographical pattern is evident when evaluating the quality of cause-of-death assignment in CRVS data. One review of cause-of-death registration quality found that over 30% of all registered deaths in Egypt, Saudi Arabia, Bolivia, and Iraq could not be assigned with certainty to even a broad cause-of-death grouping, as opposed to high-income countries such as Finland, Australia, and Ireland where fewer than 10% of records had the same coding issues.[19]

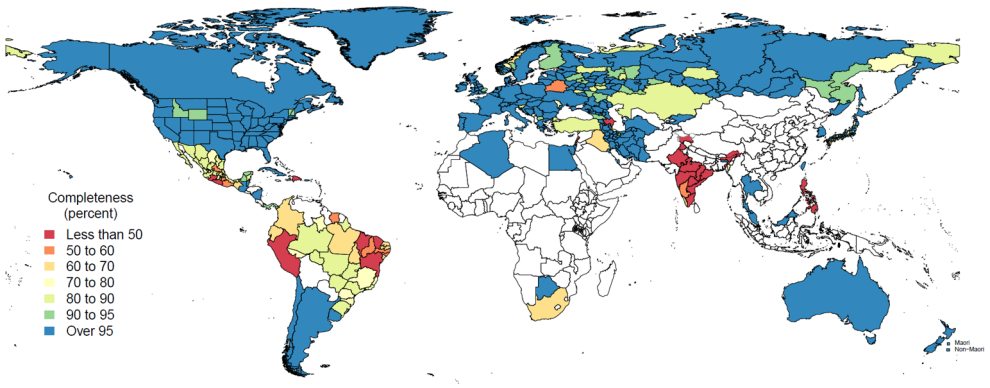


Figure 1.1: Estimated completeness of government mortality records in 2015 or the most recent year when surveillance data was available. No mortality records were available for countries symbolised in white.

These massive discrepancies are partly attributable to different institutional histories: while England and some American colonies have been tabulating death records since the 17th century[20], colonial administrations often offered no vital registration services outside of a limited register for the European colonisers, leaving little basis for an effective CRVS system after independence. Wide differences in health spending per capita across countries are also partly to blame, as is the greater fragility of health systems: an early report has found that COVID-19 disrupted CRVS collection in many low-resource settings, due to birth and death registration not being classified as essential health services.[21] Regardless of cause, the low coverage and variable quality of CRVS in many countries serves as a barrier to outcome-driven health policy.

1.2.2 Infectious disease surveillance

In addition to vital events, many high-resource settings maintain surveillance systems for so-called "notifiable infectious diseases" which are deemed to require health system action, including diseases such as mumps, cholera, hepatitis A, and yellow fever, among others. Under normal operating

conditions, any report of a notifiable infectious disease triggers control efforts such as mandatory contact tracing. Reports may be rapidly shared with a central body to allow for risk assessment and early warning of possible outbreaks.[22] In addition to notifiable infectious diseases, high-income health systems tabulate and publish weekly reports on the incidence of diseases such as influenza, HIV, and tuberculosis, allowing for the rapid identification of trends.[23]

While low-income settings may lack the resources to quickly share information about a variety of notifiable infectious diseases, almost all high-burden countries operate national programs for surveillance and control of priority infectious diseases such as HIV, tuberculosis (TB), and malaria which make up a substantial portion of all disease burden. In low- and middle-income countries (LMICs), these programs are often supplemented by funding from international institutions such as the Global Fund.[24] While these programs typically set up legal reporting requirements for priority infectious diseases, the data collection process can be hampered by a lack of electronic reporting systems; limited access to labs where infections can be confirmed; and missing data from private care providers.[25] A previous investigation of case notifications to a national TB control program in a low-resource setting found that spatial variation in case notifications was driven more by program funding and access to health services than any discernible underlying pattern in disease burden.[26] As with CRVS data in low-resource settings, these data limitations serve as substantial barriers to the use of infectious disease surveillance to inform health policymaking.

Over the past year, the COVID-19 pandemic has highlighted the challenges of maintaining infectious disease surveillance systems that can inform policy responses to health emergencies. As of 2019, the United States was ranked as one of the top countries worldwide for pandemic preparedness,[27] and its surveillance system for reporting communicable diseases was used as a template for other high-income countries.[28] Despite this, throughout the first year of the COVID-19 pandemic, the United States state information systems for COVID-19 reporting suffered from systematic data quality and completeness issues, obscuring important patterns such as the geographical distribution of diagnosed COVID-19 cases[29] and differential risks across racial groupings.[30] Simultaneously, the COVID-19 pandemic has placed pressure on surveillance strategies for other infectious diseases, significantly reducing case notifications compared to previous years.[31] Designs for future pandemic

preparedness must include infectious disease surveillance systems that can transmit timely and accurate information during health emergencies.

1.2.3 Household surveys

Given the limitations endemic to routine national surveillance data in low-resource settings, health decision makers at the national and international levels often turn to household surveys as a primary source for country health information. These types of surveys are perhaps exemplified by the Demographic and Health Surveys (DHS),[32] funded primarily by the U.S. Agency for International Development; as well as the Multiple Indicator Cluster Surveys (MICS),[33] funded primarily by UNICEF. These survey series are mainly conducted in LMICs, and are often considered to be the "gold standard" data source in place of deficient vital records or infectious disease surveillance. A standard DHS or MICS survey is designed to be representative at the national or first administrative level (often called states, regions, or provinces), sometimes split by urban and rural respondent groups. Depending on the size of the country, between 100 and approximately 1500 survey sites are selected based on a cluster sampling design, and members of approximately 30-60 households are surveyed at each site. Survey questions are taken from a standard questionnaire used across a survey round. These surveys focus on maternal and child health, reproductive health, nutrition, education, and health behaviours. After households are surveyed over a matter of months, the questionnaires are tabulated by a central agency, and a report is released with national and broad regional summaries for the survey country and year alongside the de-identified individual-level survey response data. In some cases, spatial identifiers are released for each cluster after random noise is added to de-identify the cluster location using a process called "jittering."

As a tool for health decision-making at the national level, household surveys offer several advantages over incomplete health surveillance data. They are designed to be systematic, detailed in certain topics, and representative of the national population. Because many identifying questions are asked by individual household, follow-up research can identify links between risk factors and outcomes for a surveyed country. However, even well-designed and executed surveys must be interpreted with a degree of caution. The topics are deliberately limited, providing little information about diseases

that cause high mortality among adults. Time gaps between surveys, greater than a decade in some countries, make inference about time trends in health difficult without simplifying assumptions. Additionally, because the surveys require respondents to recall past events, the responses may be biased in important and non-random ways, which can be exacerbated further based on the survey team or question wording in a particular survey round. Perhaps most relevant to the contents of this thesis, the design of these household surveys is not intended to be representative below the level of the country or its top-level administrative units. Researchers have proposed model-based solutions to each of these shortcomings, which will be discussed further in the sections below as well as in later chapters.

1.2.4 Special cases

While the previous sections broadly describe the health data context in most LMICs, China and India deserve separate description due to their large populations and individualised approaches to health surveillance. Due to the cost of directly tracking vital events and infectious diseases across populations of over 1 billion people, both China and India have developed strategies for regular surveillance of representative sub-populations and priority groups.

India is notable for conducting regular household surveys on a scale comparable to CRVS coverage in many other countries: these include the Annual Health Surveys, the District Level Health Surveys, and the National Family Health Surveys. For all of these survey series, Indian government institutions coordinate survey administration and maintain the resulting datasets.[34] India has also developed a mortality registration system designed to cover select areas of the country: this system was estimated to cover approximately 75% of its target population as of 2015.[35]

China has a long history of census-based estimation of population and health status stretching back to the 1940s.[36] More recently, China has estimated national demographic trends using a combination of censuses, household surveys, and a CRVS system that is rapidly increasing in completeness.[37] These data sources are supplemented by sentinel surveillance of maternal and child mortality, as well as a nationwide notifiable infectious diseases reporting program.[22,38]

1.2.5 Standards for quality and usability

Ultimately, health information systems must be evaluated based on their capacity to improve health and well-being both directly, for the recorded individuals, as well in the aggregate, as a tool for developing sound health policy. In line with the effort to improve the quality of health information systems worldwide, multilateral institutions have produced standards by which the quality of these systems can be measured. In its *Principles and Recommendations for a Vital Statistics System (Revision 3)*, the United Nations Statistics Division recommends that CRVS systems should be designed in line with five guiding principles. A CRVS system should be compulsory and universal, meaning that it is legally mandated, accessible, and using standardised definitions nationwide. Compiled and vetted tabulations of vital records should be prepared in a timely manner. CRVS should contain accurate information about individuals and vital events. The CRVS system should be complete, with high coverage across all sub-populations. Finally, the vital statistics system should keep individually-identifying information confidential in all publicly-available data releases.[14] These principles recognise three aspects of vital records which apply to all health statistics data: they support the legal and human rights of individuals, enable informed health policy-making in the aggregate, and yet may contain potentially compromising information that must be handled sensitively. Indeed, these principles can be extended to evaluate any system or statistical tool that purports to measure the health of a population.

Another perspective was developed in the World Health Organization's Survey-Count-Optimise-Review-Enable (SCORE) report, which rated countries worldwide based on five necessary capacities for any health information system. These follow the SCORE criteria: A successful health information system should *S*urvey populations and health risks; *C*ount births, deaths, and causes of death; *O*ptimise health service data; *R*eview progress and performance; and *E*nable data use for policy and action.[39] While these criteria emphasise the combined effect of all available data sources on a country's health surveillance and policymaking capacity, one can also evaluate individual CRVS and infectious disease surveillance programs against these criteria.

Throughout this thesis, I apply these evaluation criteria to assess CRVS systems, health surveillance systems, household survey and census data sources, and modelled estimates across a diverse set of country data contexts.

1.3 Relationship between health data availability and health system capacity

It is uncontested among global health actors that, over the long term, high quality health surveillance systems convey a set of unique health and social benefits to included individuals and the health system as a whole. However, among LMICs currently lacking high-quality health surveillance systems, the required spending, training, and infrastructure building required to develop these systems can be extremely expensive. In contexts where per-capita health spending can be as low as a few dollars per year, it can be difficult to justify additional expenses whose full benefits might not become apparent for years or decades to come. A political tendency to focus on the next election cycle can stymie the incentives to strengthen health surveillance mechanisms over the long term. Building a national health surveillance system can also entail bridging existing reporting structures across multiple institutions and levels of government. This synthesis can be a challenging obstacle even for highly motivated governments: for example, although Nigeria has long-standing laws mandating death registration nationwide, institutional and administrative challenges led to only 1 in 10 deaths being registered in 2017.[40]

Because robust health information systems can inform interventions to reduce health burden, strengthening health surveillance benefits the missions of global health funders and aid organisations as well as national ministries of health. However, AbouZahr *et al* argue that global health actors have traditionally viewed CRVS and infectious disease surveillance purely as sources of health statistics, not as goods in their own right.[9] This view complicates the argument for health surveillance strengthening in a funding environment where lives saved per dollar is an important metric. This is reflected in the early history of international support for CRVS systems: in the early years of the WHO, an office was assigned the task of CRVS strengthening, but was given neither the authority nor the funding to act on its mandate.[9] Furthermore, multilateral institutions may have well-developed feedback systems rooted in non-government data sources such as international household survey systems, combined with modelled estimates. If country health surveillance systems are ultimately intended to supplant household surveys as the primary source of health information in a country, the resulting system may be more informative for country health decision making but less informative

for funders: even highly-developed health information systems are likely to restrict external data sharing and be less comparable across borders than published results from international household survey series. As the locus of global health governance shifts towards national governments, standards for data transparency should be maintained so that multilateral institutions can successfully carry out their convening and coordinating functions.

1.4 Spatial modeling as a tool for responsive health interventions

Health researchers have relied on models to interpret data and inform action from the founding of epidemiology, when Dr. John Snow mapped cholera cases across London. While health data can be interpreted through a wide range of informal or theoretical models, modern epidemiological parlance uses "modelling" as a shorthand to refer to formal mathematical and statistical methods that attempt to approximate the "data generating process" for a measurable health phenomenon. For many applications, Bayesian data analysis methods are considered to be state-of-the-art, as they offer appealing interpretations of randomness and uncertainty as products of incomplete information.[41] This class of methods uses observed data to fit a set of underlying parameters relevant to the health condition of interest. For spatial analyses using Bayesian methods, the parameters of interest estimate the variation of a health condition across spatial units; after fitting the parameters, they can then be projected to estimate outcomes and uncertainty in them across a wider set of locations.

Health researchers are increasingly using spatial predictive modelling to estimate variation in health burden. These methods can be applied across a wide range of contexts to fill in the gaps in existing data sources, identify focal areas of disease burden, and reveal inequalities between subpopulations. The rise of Bayesian spatial analyses in health can be partly explained by their increasing ease of use and a history of high-profile applications to health policy. However, any successful application of spatial analyses to health policymaking must also deal with the disadvantages inherent to this set of methods, including the difficulty of communicating uncertainty across space, potential model misspecification, data limitations, and model interpretability.

This section introduces two approaches to spatial analyses of health and their past applications to assess health burden in both low-resource and high-resource data settings. It highlights both the inherent strengths and potential pitfalls of applying spatial modelling techniques to questions of health policy, which is expanded in later chapters as a key theme of this thesis.

1.4.1 Disease mapping and small area estimation

Many modern spatial analyses of health trace their roots to one of two historically-distinct but convergent approaches to spatial modelling: 1) small area estimation and 2) geostatistics. As I highlight in later chapters, both approaches extend traditional generalised linear mixed-effects regressions by adding model effects that capture spatial variation in the data. While the former methodological tradition emphasises neighbourhood variation across discrete areal units and the latter estimates variation across a continuous spatial surface, both offer statistical formulations of Walter Tobler's famed First Law of Geography: "everything is related to everything else, but near things are more related than distant things."[42] At their core, both approaches work with small, noisy samples of data that are assumed to be drawn from a predictably-structured underlying risk surface. Both approaches also decompose variation in the underlying risk surface into covariate-predicted, spatially-correlated, and independently distributed variation components.[43] Both frameworks were developed throughout the mid-20th century, but exploded in popularity in the 1990s thanks to the greater availability of statistical software and the development of new techniques that increased the computational efficiency of these models.[44,45]

Small area estimation describes a modelling approach used to estimate an outcome across a number of discrete areal units where the number of individuals sampled may be small, requiring the analyst to account for stochastic (random) variation. Early model formulations developed to estimate outcomes in this setting did not explicitly account for spatial autocorrelation, although they were fit across spatial units. For example, the Fay-Herriot model formulation, published in 1979 to estimate per-capita income across U.S. counties and widely used thereafter, deployed the formal assumption that all county-level observations were sampled independently and shared the same variance function, with no neighbourhood dependence.[46] Later models added a term for local autocorrelation across

space, allowing modellers of health to draw predictive information from the local neighbourhood structures across a study area.[44] Disentangling the contributions of covariate-associated, spatial, and non-spatial random variation remains an active area of research in this field.[43,47]

Geostatistical modelling aims to predict an outcome across a continuous surface where the variance-covariance relationship between any two points depends on their relative positions in space. This class of modelling, under a class of methods known as "kriging," was first widely applied in the 1960's as a method for finding mineral deposits.[48] However, because the methods to find the most likely latent surface initially required that the analyst repeatedly invert a square matrix that scaled with the number of observed points, geostatistical models remained computationally impractical to fit to large datasets even with the advent of modern statistical computing. In 2011, a methodological innovation allowed for the approximation of a continuous surface across a mesh, enabling for far larger models to be fit across continuous space.[45,49] The Integrated Nested Laplace Approximation (INLA) software, which combines this innovation with a fast approximation to a Markov Chain Monte Carlo sampler, has become a widely-used tool for implementing geostatistical models.[50,51]

While these two methods sit at the core of many modern spatial statistical analyses of health, extensions and alternatives to these two approaches remain an active area of research in health statistics. Extensions to disease mapping have added an additional dimension for space-time analysis,[52] incorporated principles from species distribution mapping,[53] and used machine learning methods to generate spatial covariates that improve the final model fit.[54] Beyond the framework of mixed effects modelling, other teams have modernised concepts from cartography and Geographic Information Science to estimate outcomes over a large spatial field: notably, the WorldPop project has extended the concept of dasymetric mapping to estimate human population at a fine geographic scale worldwide.[55] Beyond spatial predictive modelling, another widely applied set of methods identifies "hot spots" by estimating the probability that local deviations from the mean occurred from random chance alone.[56,57]

1.4.2 Applications of spatial modelling

As spatial predictive modelling methods have grown in popularity, health researchers have quickly seen their utility for mapping disease indicators and risk factors based on household survey data.

Under the provenance of "disease mapping", geostatistical and small area analyses have been applied to estimate underlying disease burden based on geolocated data systematically sampled from the population.[11] These models can be applied to understand local variation in a wide variety of health phenomena, from infectious diseases[58] to maternal and child health.[59] A number of working groups have extended the principles of spatial predictive modelling to map health burden and interventions across many LMICs: these include the Malaria Atlas Project (MAP), which primarily maps malaria and related interventions;[60,61] the Institute for Health Metrics and Evaluation (IHME), which has mapped a variety of indicators related to maternal and child health across low- and middle-income countries;[62] the Kenya Medical Research Institute (KEMRI), which has mapped many health interventions across Kenya and East Africa;[63] and networks of epidemiologists and demographers who have developed new methods for estimating disease prevalence and mortality across space using household survey data.[11,64]

These teams benefit not only from the continual development of modelling tools that can be run on a personal computer, but also from an increasingly diverse range of spatial data sources. Both small area and geostatistical models can draw predictive power from covariates, spatially-varying surfaces that may track variation in the outcome. The list of high-quality spatial covariates is growing rapidly: some of the covariates related to health include remotely-sensed estimates of short-term weather patterns, as well as longer-term land use and climate change trends, which may be disease risk factors.[65] Health datasets based on volunteered geographic information can achieve high coverage in remote areas: for example, the Humanitarian OpenStreetMaps Team has rapidly compiled maps of road and building features in rural areas struck by natural disasters.[66] Modelled estimates of health outcomes and populations can also be used as a model input: the WorldPop population estimates, for instance, are an input to many spatial models of health.[55]

In low-income countries, spatial models of health outcomes mapped by these have laid the groundwork for targeted health interventions led by UNICEF, USAID, and the Bill and Melinda Gates Foundation, among others. In lieu of mapping high-quality health surveillance data, these modelled estimates of health outcomes allow multilateral institutions to target and design new

surveys; track the effect and scaling up of interventions; and to develop programs that target the areas in greatest need.[11,58]

In high-income countries, the greater fidelity of surveillance-derived health data allows for a wider range of modelling approaches. Some researchers have taken advantage of this greater fidelity to estimate health outcomes among very small areal units, such as United States census tracts with an average population size of 4,000.[3,67] Others have split the population further, developing age-period-cohort models for cancer incidence[68] as well as mortality modelling over the dimensions of space, age group, year, and cause of death.[69] Mortality reports tabulated by week and year have allowed for estimation of seasonality in causes of death as well as the estimation of weekly "excess mortality" above an expected baseline, described in Chapter 5 of this thesis.[70] Individually- and spatially-linked records from health surveillance, when confidentially shared with researchers, can also serve as the basis for spatial risk factor detection using observational data.

Health analyses performed in high-income countries often rely on the assumption that the underlying reported data covers the full population, with no missing data or delays relevant to the analysis: this assumption allows for an analysis where all noise is attributable to stochastic variation that comes from sampling a small population. This assumption is justified by select data validity and completeness audits performed in high-resource health surveillance systems. However, the COVID-19 pandemic has revealed previously untested weaknesses even in high-resource health surveillance systems, including long reporting lags, misreported death totals showing up as negative deaths on COVID-19 dashboards, and under-reporting of COVID-19 as an underlying cause of death.[70]

1.4.3 Limitations

In low-resource settings, a paucity of data may leave health policy-makers with no choice but to respond to data collected from household surveys, possibly in combination with modelled estimates if available. While modelled household survey estimates may be presented to policy-makers as the best possible health information available in a country, they may actually obscure major gaps in knowledge and can possibly distort the consensus understanding of the state of health on the ground, leading to sub-optimal decision making. To understand these limitations, we must compare these

estimates to the principles that should be embodied by a successful health surveillance system and briefly explore the political economy of health data production.

One limitation common to spatial analyses of health is the uncertainty associated with the interpretation of data observations with small sample sizes, along with the difficulty of communicating that uncertainty across space using a map. In Bayesian statistics, uncertainty is often expressed in terms of 95% Uncertainty Intervals (UIs), which represent the range where the output parameters would fall 95% of the time if the model was properly specified.[41] Translated into a spatial context, the output of Bayesian spatial modelling is a large number of "candidate maps" that represent possible explanations of the underlying data, where better explanations are more likely to be included.[71] Summaries of these candidate maps, including the mean and 95% UIs for each pixel-level posterior estimate, can be produced. However, this uncertainty is particularly difficult to communicate across space, where it cannot be intuitively displayed along with a central summary estimate in a single map. Past research on cartography in policymaking provides additional evidence that observers tend to understand summary maps as settled truth: in this way, an astute advocate can convey a sense of finality to unsettled questions in space simply by declaring that "it's there on the map."[72] This tendency, combined with policy processes with no built-in methods for incorporating uncertainty, can lend interpretations to summary maps that may not be supported by the underlying model results. For this reason, spatial modellers of health must be extremely careful to place uncertainty in the foreground of their analyses and reported results.

However, if a spatial model is incorrectly specified, even cautious interpretations of mapped results can mislead policy-makers. All models are built on assumptions, with spatial models being particularly sensitive to the selection of predictive covariates and prior assumptions about the strength of spatial neighbourhood effects. If modelled spatial estimates are themselves used as covariates in a predictive spatial model, this can lead to a circularity problem of "models on top of models," where the effect of a raw covariate used to predict an intermediate outcome can become overstated in areas where no data was sampled. Because household surveys may offer little to no information about particular causes of death, modelling assumptions can have a disproportionate role in the results, with no immediate method to check the validity of those results. For example, past differences in

global malaria mortality estimates reported by IHME and the WHO were due largely to differences in model assumptions that had a large effect in countries where little to no death registration data was available.[73] In this way, seemingly innocuous choices about modelling approaches, prior specifications, and processing of input data in data-sparse areas can have large implications for global health policymaking and financing. When the number of areal units to be estimated is large, these same data sparsity issues can manifest even for outcomes that are regularly measured by household surveys. In the past year, varying estimates of COVID-19 death total across Africa underscore that even modern statistical techniques cannot overcome a lack of reliable data.

If we are to measure the success of a health data system by its long-term impact on health outcomes, we must also interrogate a health information system that sidelines local and national governments in the health data governance equation. Although the computational costs to spatial modelling have dropped dramatically over the past decade, learning spatial statistics without instruction from a current practitioner requires a steep learning curve that serves as a barrier to entry for many health statisticians and epidemiologists worldwide. When high-income institutions collect, manage, and store survey data which other high-income institutions then digest and share with multi-laterals as authoritative health estimates, it leaves countries with fewer levers to insert their perspectives into decision-making around global health financing and action,[74] sparking fears of extractive data imperialism.[75] Spatial statisticians mapping health outcomes in low-income countries must therefore navigate the potential effects of their actions in a shifting global health funding landscape while also striving for objectivity and reproducibility in their statistical practice—a challenging proposition.

1.5 New hierarchical modelling techniques for mapping health outcomes using incomplete health surveillance data

This thesis attempts to bridge the divide between the statistical and political economy challenges that impede the development of strong health surveillance systems in low-resource settings. In the following chapters, I introduce a class of spatial statistical models that incorporate deficient vital and health surveillance records to offer new insights into the health of a country that are impossible to

derive from other sources. By bringing together data from both household surveys and routine health surveillance, these models simultaneously estimate outcomes of interest and the current completeness of health surveillance data sources. At the same time, thanks to the large sample sizes captured by health surveillance data, outcomes for this class of models will become more certain as the fidelity of the underlying health surveillance data source increases. This creates an incentive for countries to improve the quality of health surveillance data. Ultimately, the goal of this class of models is to enable a virtuous cycle where greater reliance on health surveillance data leads to a greater emphasis on funding for health surveillance and health system strengthening activities, ultimately leading to health outcomes that are both improved and better-documented.

This approach recognises that surveillance systems for infectious diseases and vital events are irreplaceable building blocks of a functioning health system.[8] By emphasising both the statistical and political economy challenges inherent to health surveillance strengthening, it aligns with past literature on the "critical cartographic" project that recognises the inherently value-laden and political implications of maps.[76] The spatial aspect of these models is important, since variation in observed data across a country can provide important clues about the overall quality of a health surveillance data source. The principles of universality, timeliness, accuracy, completeness, and confidentiality in a health surveillance system—as well as barriers to these—often correspond to processes in the local health system that can be assessed using a spatial approach. Furthermore, barriers to high-quality health data overlap with health service capacity issues at the local level, and targeting improved data quality may ultimately overlap with efforts to increase access to health care. This approach also recognises that new statistical methods are necessary, but that methods alone are not sufficient to work towards adequate health data governance in low- and middle-income countries.[77] Complementary innovations in national and international institutions are needed to develop sustainable, high-quality health surveillance systems that can address national health inequalities.

The joint survey-surveillance approaches described in this thesis apply recent innovations in the spatial statistics to maximise the information derived from each data type. While simple "cross-walks" with a single correction term have long been deployed to translate between different data types, I describe a more nuanced approach to joint estimation, made possible by a new software

package that can be adapted for custom spatial estimation.[78] These methods all rely on a Bayesian hierarchical modelling approach where the data-generating process also includes the generation of CRVS and health surveillance that is variably incomplete over space.[79] Using these modelling innovations, countries with incomplete health surveillance data can estimate the biases in that data source by comparing it to the "gold standard" of household survey data. Conversely, spatial and temporal gaps in household survey data can be filled in based on space-time trends identified in the locations where health surveillance is strong.

The programmatic benefits to a space-time modelling approach for health surveillance completeness are readily apparent. As with health system strengthening, health surveillance strengthening should be outcome-driven: by explicitly tracking changes in health surveillance bias over space, these models identify districts with high data completeness as well as targets for improvement within a country. Using a spatial statistical approach, health researchers can target different levels of policy and funding impact ranging from the local and state to the national level. Modelling changes over time also allows for the identification of possible facilitators of, and barriers to, improvements over time. This approach can even be applied in a high-resource context to check the completeness assumption deployed in most models of health.

Embracing both the statistical and programmatic challenges of health surveillance strengthening also requires adapting models to the data contexts and policy needs of particular countries. To understand the data generating process underlying particular country contexts, this thesis is divided into four country collaborations, emphasising differences in their data environments and the institutional history of their health surveillance systems. Although the resulting models are tailored to their country contexts, they can be generalised into templates that can be modified and then deployed across a wide variety of contexts. For example, a framework developed to jointly model rare health events and surveillance completeness can be used both to model neonatal mortality in Mexico (Chapter 3) as well as tuberculosis incidence in Uganda (Chapter 4). Productive engagement between modelling and policy also entails a transparent approach to communication that brings model limitations and potential for harm to the foreground of the conversation.[74] The analysis code and

data for this thesis has been developed using open-source frameworks whenever possible and released publicly on-line (Appendix F) as a starting point for greater adoption.[80]

1.6 Thesis structure

This thesis is structured as a series of investigations of surveillance-based spatial models of health, contextualised into four case studies with particular policy and disease contexts. In Mexico, I demonstrate how a country with high-quality CRVS data may still face data quality challenges at the subnational level; I then develop a model to jointly estimate neonatal mortality and CRVS completeness that incorporates prior knowledge of locations with variable CRVS quality. This class of joint spatial model has a wide set of applications in health: in the next chapter, I demonstrate how a similar model can be used to estimate TB prevalence as well as case notification completeness across Uganda. I take a more holistic approach to assess the capacity of India's three surveillance systems for child mortality; using a spatial model, I then demonstrate how more granular reporting of two surveillance systems could accelerate progress towards goals set in the Indian National Health Plan. Finally, in Italy, I estimate all-cause mortality across the dimensions of space, time, age, and week to predict excess mortality from the COVID-19 pandemic, revealing new insights about the interpretation of cause-specific death tabulations by Italian region. This last analysis demonstrates weaknesses in Italy's cause-specific mortality reporting system during the first months of the COVID-19 pandemic, challenging the dichotomy between complete and incomplete CRVS analysis strategies.

Finally, my concluding chapter reflects on the methods and themes common across these analyses. The most important of these reveals how the coverage of health surveillance data enables small-population estimation that can inform responsive and equitable health policy. This trend can be explored by jointly estimating local variation in disease burden and input data quality, a novel approach with wider applications in health modelling and elsewhere. Reflecting on the role of spatial statistical modelling in health, this chapter reiterates the immediate opportunities for spatial models to supplement developing health surveillance and household surveys, and concludes that spatial models of health should strengthen the development of complete, high-quality national health surveillance systems.

1.7 References

1. UN General Assembly. *Universal declaration of human rights* 1948.
2. Burstein, R. *et al.* Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature* **574,** 353–358. http://www.nature.com/articles/s41586-019-1545-0 (Oct. 2019).
3. Dwyer-Lindgren, L. *et al.* Variation in life expectancy and mortality by cause among neighbourhoods in King County, WA, USA, 1990-2014: a census tract-level analysis for the Global Burden of Disease Study 2015. *The Lancet Public Health* **2,** e400–e410. http://dx.doi.org/10.1016/S2468-2667(17)30165-2 (2017).
4. Ruger, J. P. Ethics and governance of global health inequalities. *Journal of Epidemiology and Community Health* **60,** 998–1003 (2006).
5. Brown, T. M., Cueto, M. & Fee, E. The World Health Organization and the transition from international to global public health. *American Journal of Public Health* **96,** 62–72 (2006).
6. World Health Organization. *Everybody's business: strengthening health systems to improve health outcomes: WHO's framework for action.* tech. rep. (World Health Organization, Geneva, 2007), 1–44.
7. World Health Organization. *Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies* tech. rep. 1 (2010), 1–92. http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.35.021103.105711.
8. Roberts, M. J., Hsiao, W., Berman, P. & Reich, M. R. *Getting health reform right: a guide to improving performance and equity* (2008).
9. AbouZahr, C. *et al.* Towards universal civil registration and vital statistics systems: The time is now. *The Lancet* **386,** 1407–1418. http://dx.doi.org/10.1016/S0140-6736(15)60170-2 (2015).
10. Rao, C. Elements of a strategic approach for strengthening national mortality statistics programmes. *BMJ Global Health* **4,** 1–10 (2019).
11. Diggle, P. J. & Giorgi, E. Model-based geostatistics for prevalence mapping in low-resource settings. *Journal of the American Statistical Association* **111,** 1096–1120. arXiv: 1505.06891. http://dx.doi.org/10.1080/01621459.2015.1123158 (2016).
12. Mikkelsen, L. & Lopez, A. D. *Improving the quality and use of birth, death and cause-of-death information* tech. rep. (2010), 1–92. http://www.who.int/healthinfo/tool_cod_2010.pdf.
13. Setel, P. W. *et al.* A scandal of invisibility: making everyone count by counting everyone. *Lancet* **370,** 1569–1577 (2007).
14. United Nations Statistics Division. *Principles and Recommendations for a Vital Statistics System Revision 3* **19.** http://unstats.un.org/unsd/Demographic/standmeth/principles/M19Rev3en.pdf (2014).
15. Fisker, A. B., Rodrigues, A. & Helleringer, S. Differences in barriers to birth and death registration in Guinea-Bissau: implications for monitoring national and global health objectives. *Tropical Medicine and International Health* **24,** 166–174 (2019).
16. Hernández, B. *et al.* Subregistro de defunciones de menores y certificación de nacimiento en una muestra representativa de los 101 municipios con más bajo índice de desarrollo humano en México. *Salud Pública de México* **54,** 393–400 (2012).
17. Mahapatra, P. *et al.* Civil registration systems and vital statistics: successes and missed opportunities. *Lancet* **370,** 1653–1663 (2007).
18. Roth, G. A. *et al.* Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **392,** 1736–1788. https://linkinghub.elsevier.com/retrieve/pii/S0140673618322037 (Nov. 2018).
19. Johnson, S. C. *et al.* Public health utility of cause of death data: applying empirical algorithms to improve data quality. *BMC Medical Informatics and Decision Making* **21,** 1–20 (2021).
20. Blake, J. B. The Early History of Vital Statistics in Massachusetts. *Bulletin of the History of Medicine* **29,** 46–68. https://www.jstor.org/stable/44449136 (1955).

21. AbouZahr, C. *et al.* The COVID-19 pandemic: Effects on civil registration of births and deaths and on availability and utility of vital events data. *American Journal of Public Health* **111,** 1123–1131 (2021).
22. Vlieg, W. L. *et al.* Comparing national infectious disease surveillance systems: China and the Netherlands. *BMC Public Health* **17,** 1–9 (2017).
23. Thacker, S. B., Berkelman, R. L. & Stroup, D. F. The Science of Public Health Surveillance. *Journal of Public Health Policy* **10,** 187–203. http://www.jstor.com/stable/3342679 (1989).
24. Mauch, V. *et al.* Structure and management of tuberculosis control programs in fragile states-Afghanistan, DR Congo, Haiti, Somalia. *Health Policy* **96,** 118–127. http://dx.doi.org/10.1016/j.healthpol.2010.01.003 (2010).
25. Uplekar, M. *et al.* Mandatory tuberculosis case notification in high tuberculosis-incidence countries: Policy and practice. *European Respiratory Journal* **48,** 1571–1581. http://dx.doi.org/10.1183/13993003.00956-2016 (2016).
26. Rood, E. *et al.* A spatial analysis framework to monitor and accelerate progress towards SDG 3 to end TB in Bangladesh. *ISPRS International Journal of Geo-Information* **8,** 1–11 (2019).
27. Johns Hopkins Center for Health Security. *Global Health Security Index: 2019 Report* tech. rep. (Johns Hopkins University, 2019), 270. www.ghsindex.org.
28. Bagherian, H., Farahbakhsh, M., Rabiei, R., Moghaddasi, H. & Asadi, F. National communicable disease surveillance system: A review on information and organizational structures in developed countries. *Acta Informatica Medica* **25,** 271–276 (2017).
29. Gold, J. A. *et al.* COVID-19 Case Surveillance: Trends in Person-Level Case Data Completeness, United States, April 5-September 30, 2020. *Public Health Reports* **136,** 466–474 (2021).
30. Krieger, N., Testa, C., Hanage, W. P. & Chen, J. T. US racial and ethnic data for COVID-19 cases: still missing in action. *The Lancet* **396,** e81. http://dx.doi.org/10.1016/S0140-6736(20)32220-0 (2020).
31. Kwak, N., Hwang, S. S. & Yima, A. J. Effect of COVID-19 on Tuberculosis Notification, South Korea. *Emerging Infectious Diseases* **26,** 2506–2508 (2020).
32. Corsi, D. J., Neuman, M., Finlay, J. E. & Subramanian, S. V. Demographic and health surveys: A profile. *International Journal of Epidemiology* **41,** 1602–1613 (2012).
33. Khan, S. & Hancioglu, A. Multiple Indicator Cluster Surveys: Delivering Robust Data on Children and Women across the Globe. *Studies in Family Planning* **50,** 279–286 (2019).
34. Dandona, R., Pandey, A. & Dandona, L. A review of national health surveys in India. *Bulletin of the World Health Organization* **94,** 286–296A (2016).
35. Kumar, G. A., Dandona, L. & Dandona, R. Completeness of death registration in the Civil Registration System, India (2005 to 2015). *Indian Journal of Medical Research* **149,** 740. http://www.ncbi.nlm.nih.gov/pubmed/23144490%20http://www.ijmr.org.in/text.asp?2019/149/6/740/265955 (2019).
36. Banister, J. & Hill, K. Mortality in China 1964-2000. *Population Studies* **58,** 55–75 (2004).
37. Zeng, X. *et al.* Measuring the completeness of death registration in 2844 Chinese counties in 2018. *BMC medicine* **18,** 176 (2020).
38. He, C. *et al.* National and subnational all-cause and cause-specific child mortality in China, 1996-2015: a systematic analysis with implications for the Sustainable Development Goals. *The Lancet Global Health* **5,** e186–e197 (2017).
39. World Health Organization. *SCORE for health data technical package: global report on health data systems and capacity, 2020* tech. rep. (World Health Organization, Geneva, 2021), 1–104. https://www.who.int/publications/global-report-on-health-data-systems-and-capacity-2020.
40. Makinde, O. A. *et al.* Death registration in Nigeria: a systematic literature review of its performance and challenges. *Global Health Action* **13.** https://doi.org/10.1080/16549716.2020.1811476 (2020).
41. McElreath, R. *Statistical Rethinking* (Taylor & Francis, Boca Raton, FL, 2016).

42. Tobler, W. R. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* **46,** 234–240 (1970).
43. Riebler, A. *et al.* An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research* **25,** 1145–1165. arXiv: 1601.01180 (2016).
44. Besag, J., York, J. & Mollié, A. A Bayesian image restoration with two applications in spatial statistics Ann Inst Statist Math 43: 1-59. *Find this article online* **43,** 1–20 (1991).
45. Lindgren, F. & Rue, H. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society. Series B* **73,** 423–498 (2011).
46. Fay, R. E. & Herriot, R. A. Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association* **74,** 269 (1979).
47. MacNab, Y. C. On Gaussian Markov random fields and Bayesian disease mapping. *Statistical Methods in Medical Research* **20,** 49–68 (2011).
48. Oliver, M. A. in *Handbook of Applied Spatial Analysis: Software Tools, Methods, and Applications* (eds Fischer, M. M. & Getis, A.) 319–352 (Springer, 2010). https://link.springer.com/book/10.1007/978-3-642-03647-7.
49. Miller, D. L., Glennie, R. & Seaton, A. E. Understanding the Stochastic Partial Differential Equation Approach to Smoothing. *Journal of Agricultural, Biological, and Environmental Statistics* **25,** 1–16. arXiv: 2001.07623. https://doi.org/10.1007/s13253-019-00377-z (2020).
50. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **71,** 319–392 (2009).
51. Krainski, E. *et al. Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA* **September** (2018).
52. Mercer, L. D. *et al.* Space-time smoothing of complex survey data: Small area estimation for child mortality. *Annals of Applied Statistics* **9,** 1889–1905 (2015).
53. Hay, S. I. *et al.* Global mapping of infectious disease. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368** (2013).
54. Bhatt, S. *et al.* Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *Journal of the Royal Society Interface* **14.** arXiv: 1612.03278 (2017).
55. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* **4,** 2–5 (2017).
56. Kulldorff, M. A spatial scan statistic. *Communications in Statistics - Theory and Methods* **26,** 1481–1496 (1997).
57. Banerjee, S., Carlin, B. P. & Gelfand, A. E. *Hierarchical modeling and analysis for spatial data* 2nd Editio (CRC Press, 2014).
58. Pigott, D. M. *et al.* Prioritising infectious disease mapping. *PLoS Neglected Tropical Diseases* **9,** 1–21 (2015).
59. Liang, J. *et al.* Maternal mortality ratios in 2852 Chinese counties, 1996-2015, and achievement of Millennium Development Goal 5 in China: a subnational analysis of the Global Burden of Disease Study 2016. *The Lancet* **393,** 241–252. http://dx.doi.org/10.1016/S0140-6736(18)31712-4 (2019).
60. Weiss, D. J. *et al.* Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000-17: a spatial and temporal modelling study. *The Lancet* **394,** 322–331. http://dx.doi.org/10.1016/S0140-6736(19)31097-9 (2019).
61. Nguyen, M. *et al. Mapping malaria seasonality: a case study from Madagascar* 2019. arXiv: 1901.10782. http://arxiv.org/abs/1901.10782.
62. Osgood-Zimmerman, A. *et al.* Mapping child growth failure in Africa between 2000 and 2015. *Nature* **555,** 41–47. http://www.nature.com/articles/nature25760 (Mar. 2018).
63. Maina, J. *et al.* A spatial database of health facilities managed by the public health sector in sub Saharan Africa. *Scientific Data* **6,** 1–8 (2019).

64. Wakefield, J., Okonek, T. & Pedersen, J. Small Area Estimation for Disease Prevalence Mapping. *International Statistical Review,* insr.12400. https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12400 (July 2020).
65. Ericksen, P. *et al. Mapping hotspots of climate change and food insecurity in the global tropics* tech. rep. June (CGIAR Research Program on Climate Change, 2011). http://www.dfid.gov.uk/r4d/PDF/Outputs/CCAFS/ccafsreport5-climate_hotspots_final.pdf%5C%0Ahttp://cgspace.cgiar.org/bitstream/handle/10568/3826/ccafsreport5-climate_hotspots_final.pdf?sequence=13.
66. Thomson, D. R. *et al.* Extending Data for Urban Health Decision-Making: a Menu of New and Potential Neighborhood-Level Health Determinants Datasets in LMICs. *Journal of Urban Health* **96,** 514–536 (2019).
67. Zhang, X. *et al.* Multilevel regression and poststratification for small-area estimation of population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. *American Journal of Epidemiology* **179,** 1025–1033 (2014).
68. Papoila, A. L. *et al.* Stomach cancer incidence in Southern Portugal 1998-2006: A spatio-temporal analysis. *Biometrical Journal* **56,** 403–415 (2014).
69. Dwyer-Lindgren, L. *et al.* US county-level trends in mortality rates for major causes of death, 1980-2014. *JAMA - Journal of the American Medical Association* **316,** 2385–2401 (2016).
70. Weinberger, D. M. *et al.* Estimation of Excess Deaths Associated with the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Internal Medicine* **06520,** E1–E9 (2020).
71. Patil, A. P., Gething, P. W., Piel, F. B. & Hay, S. I. Bayesian geostatistics in health cartography: The perspective of malaria. *Trends in Parasitology* **27,** 246–253 (2011).
72. Elwood, S. Beyond cooptation or resistance: Urban spatial politics, community organizations, and GIS-based spatial narratives. *Annals of the Association of American Geographers* **96,** 323–341 (2006).
73. Tichenor, M. & Sridhar, D. Metric partnerships: Global burden of disease estimates within the World Bank, the World Health Organisation and the Institute for Health Metrics and Evaluation. *Wellcome Open Research* **4** (2020).
74. Cinnamon, J. Data inequalities and why they matter for development. *Information Technology for Development* **26,** 214–233 (2020).
75. Marchais, G., Bazuzi, P. & Amani Lameke, A. The data is gold, and we are the gold-diggers': whiteness, race and contemporary academic research in eastern DRC. *Critical African Studies* **12,** 372–394. https://doi.org/10.1080/21681392.2020.1724806 (2020).
76. Crampton, J. W. & Krygier, J. An introduction to critical cartography. *Acme* **4,** 11–33 (2006).
77. Tiffin, N., George, A. & Lefevre, A. E. How to use relevant data for maximal benefit with minimal risk: Digital health data governance to protect vulnerable populations in low-income and middle-income countries. *BMJ Global Health* **4,** 1–9 (2019).
78. Osgood-Zimmerman, A. & Wakefield, J. *A Statistical Introduction to Template Model Builder: A Flexible Tool for Spatial Modeling* 2021. arXiv: 2103.09929. http://arxiv.org/abs/2103.09929.
79. Schmertmann, C. P. & Gonzaga, M. R. Bayesian Estimation of Age-Specific Mortality and Life Expectancy for Small Areas With Defective Vital Records. *Demography* **55,** 1363–1388 (2018).
80. Shannon, J. & Walker, K. Opening GIScience: A process-based approach. *International Journal of Geographical Information Science* **32,** 1911–1926. https://doi.org/10.1080/13658816.2018.1464167 (2018).

Chapter 2

Joint estimation of neonatal mortality and vital registration completeness across Mexico

2.1 Introduction

Comprehensive surveillance of vital events and disease burden across a country is critical for national health planning, making it an important component of the right to high-quality medical care enshrined in the Universal Declaration of Human Rights.[1,13] Many countries implement a system of civil registration that requires medical professionals to produce legal documents for vital events such as births and deaths; these records are compiled in nationwide vital statistics that provide the foundation for national strategic planning.[81] However, despite their importance in setting national priorities for health care, local variation in the quality and completeness of these Civil Registration and Vital Statistics systems (CRVS) remains poorly understood.

Many geospatial studies investigating local variation in disease burden derive their estimates from cluster-level observations in household surveys and censuses.[11,64] Routine health surveillance includes features that make it an appealing alternative or supplement to traditional geospatial data sources: most notably, the sample sizes associated with CRVS datasets are typically orders of magnitude larger than those collected in any household survey. While years may pass between two household surveys, functioning health surveillance systems provide an unbroken series of observations over time. Many surveillance systems already report health status at the administrative level

that is most relevant to country-level financing and planning. More broadly, global health researchers have the opportunity to invest in, and advocate for, data sources that are fundamentally tied to the success of national health systems in the countries where our research is focused.[9]

However, critical issues must be resolved before CRVS data can be incorporated into geospatial analyses of health. The most pressing of these is the question of varying incompleteness in health surveillance in space and time. Previous analyses have shown that the completeness of CRVS data varies across countries, across states or provinces within countries, and over time;[82] completeness is also generally lower for the registration of deaths among children than among adults.[83] While completeness of death registration has improved in many countries since 2000, it remains low in many low- and middle- income countries where the global burden of child mortality is concentrated.[84] While methods have been developed to account for incompleteness in health surveillance data at the national and state levels,[85–87] these methods cannot be directly applied at more local levels due to the general geospatial problem of small sample sizes.

In this chapter, I demonstrate how household survey data and CRVS data can be incorporated into a novel small area model that simultaneously estimates neonatal mortality and CRVS incompleteness at a local level. I present results for this model in the context of Mexico, which has a CRVS system that is considered near-complete at the national level[88,89] but may be incomplete in marginalised municipalities.

2.1.1 Estimating completeness of birth and death registration across Latin America

Many Latin American countries are now facing challenges and opportunities associated with estimating mortality based on CRVS records that are increasing in quality each year. Most countries across the region first developed vital statistics programs during an international push for civil registration in the 1960s;[10] however, by the early 2000s, many of these systems remained highly incomplete.[84] In Mexico, a health reform policy called the Seguro Popular dramatically increased health system coverage and registration during the early years of the 21st century,[90] while in Brazil, increased interest in health statistics performance produced a series of studies on CRVS completeness.[91–94] Today, international statistical groups such as the UN Inter-agency Group on Mortality Estimation

(UN IGME) and the Institute for Health Metrics and Evaluation (IHME) rate Argentina, Brazil, Chile, and Mexico as having some of the most complete CRVS systems in the world,[88,89] while other countries in the region such as Ecuador and Colombia are transitioning towards a health surveillance system based on primarily CRVS data.[95] Data sources for health are relatively plentiful across many Latin American countries, including both household surveys and vital records.

While the completeness of CRVS data is of great interest to ministries of health in these countries, estimating completeness of vital records for children at the subnational level remains challenging. For decades, capture-recapture analysis was the preferred method for estimating the completeness of vital records. This method was derived from ecology, where it is still used to estimate wildlife populations based on multiple surveys.[96] The method relies on an assumption, shown graphically in Figure 2.1, that each event of interest has an independent and equal probability of being recorded by a particular data source, and data sources are therefore capturing independent samples of a total population. In 1949, health statisticians Chandra Sekar and Deming applied this method to estimate true underlying birth and death rates based on a combination of a government registry and a household survey in a neighbourhood in Calcutta.[97] Since then, the same basic approach has been extended to estimate disease incidence,[98] cause-specific mortality, and prevalence of genetic disorders, to name a few examples.[99] From the 1950s through the 1990s, the capture-recapture method was widely used to estimate incompleteness of mortality data in middle-income countries.[100,101] However, this method fell out of favour in the early 2000s given the weakness of its assumptions: under realistic circumstances, events captured by one survey are typically more likely to be captured by other surveys as well, meaning that the results from capture-recapture analysis are almost always under-estimates.[102,103]

More recently, some countries have begun to implement small-scale audits to check the completeness of CRVS records for child mortality. These audits operate in a relatively small number of districts: within each district, a trained team of investigators develops an exhaustive record of mortality based on data from hospitals, churches, graveyards, and sometimes a household survey. These audits have been conducted in select regions of Brazil,[92,93,104] Colombia,[105] and Mexico.[16] In Mexico, a 2009 audit of 101 municipalities with a low Human Development Index identified that 68%

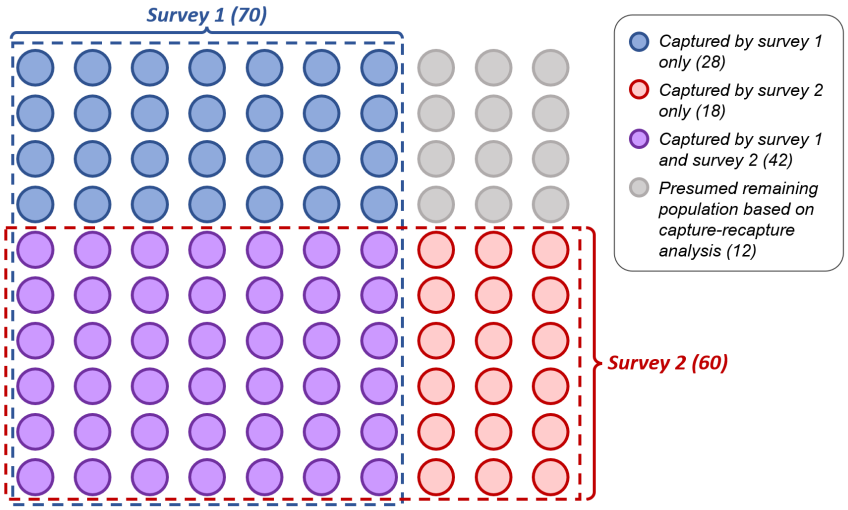


Figure 2.1: A visual demonstration of capture-recapture analysis using linked records from two surveys.

of births and 22% of deaths among children under 5 had not been entered into government registries, in contrast to relatively high coverage of birth and death registration across the rest of the country.[16]

A previous modelling study in Brazil combined CRVS records with completeness estimates from a series of local CRVS audits to estimate child mortality by Brazilian mesoregion.[106] This study provided a novel Bayesian approach for estimating child mortality from incomplete CRVS data: the true under-5 mortality rate, m_i, was fit indirectly to the data $D_i \sim Poisson(N_i m_i \pi_i)$, where N_i and D_i are the under-5 population and deaths recorded by CRVS in a mesoregion, and π_i is an incompleteness parameter with a strong prior based on previous CRVS audits. However, because the mortality rate m_i was only fit indirectly using small-counts data, child mortality estimates were subject to wide uncertainty at the mesoregion level. Additionally, while the study was able to draw on a series of existing audits estimating the completeness of under-5 mortality records across the mesoregions of Brazil,[92,93] the relative expense of these audits reduces their capacity to enumerate child mortality completeness across other Latin American countries.

A third approach to mortality estimation across Latin America has prioritised collating data from multiple sources to produce a unified estimate of mortality. Two classes of data have historically informed estimates of child mortality: birth history (BH) data, which retrospectively lists the life histories of all children born to a mother, and CRVS data, which attempts to enumerate all births and

deaths in a given time period. At the national level, multi-source estimation methods have combined estimates from these two data types to estimate both mortality trends and CRVS completeness,[15,89] but the methods deployed in these studies are problematic for subnational estimation due to the smaller sample sizes available at each observation. Most modelled subnational estimates of child mortality have relied solely on birth history data from household surveys, informed by spatial covariates.[2,107,108] However, in countries with high-quality CRVS systems such as Mexico, vital records are a preferable data source for mortality estimation under most circumstances due to their greater spatial coverage, larger sample sizes, and integration into existing processes for health decision-making.[14]

In the following section, I describe a small-area model to estimate neonatal mortality across Mexican municipalities by combining BH and CRVS data, with an emphasis on the interpretation of bias in CRVS data.

2.2 Methods

I developed a small area model that simultaneously estimates neonatal mortality rates and CRVS bias by municipality. This model incorporates two sources of data for neonatal mortality, both of which collect data about births and age-specific mortality: (1) birth histories collected from household surveys and (2) birth and death records from a civil registration system. To test the predictive validity of this joint estimation framework, I first fit a model using observations generated from simulated surfaces of neonatal mortality and CRVS bias, measuring the correspondence between the simulated underlying surfaces and the recovered model parameters. I then applied this model to estimate neonatal mortality and CRVS bias by Mexican municipalities in 2009-2010 using mortality data sources and covariates published by the Mexican Institute for Geography and Statistics (INEGI). By fitting multiple models with different priors on CRVS bias by municipality, I explored the effects of CRVS bias terms on municipal and state-level neonatal mortality estimates during the years 2009-2010. Figure 2.2, below, summarizes the modelling process described in this chapter.

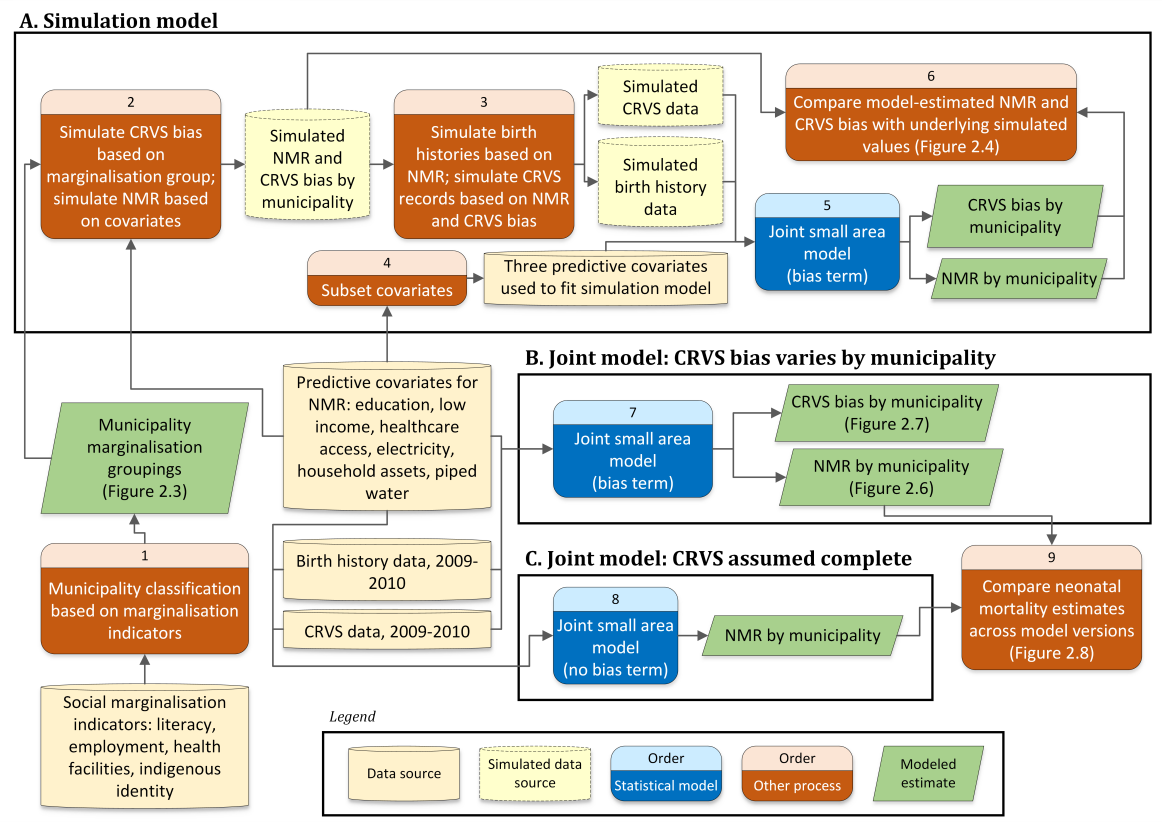


Figure 2.2: Summary of data sources, analyses, and results described in Chapter 2. NMR: neonatal mortality rate, CRVS: civil registration and vital statistics.

2.2.1 Data preparation

The model incorporates birth and death data from civil registration, mortality data from household surveys, and areal-level covariates. I prepared data across two years, 2009 and 2010, in order to increase the sample sizes of observed births and to avoid CRVS reporting delays in single years while still capturing a relatively short time period within which neonatal mortality can be safely assumed to remain relatively stable. I downloaded publicly-available microdata (that is, data available by anonymised individual) on births and deaths by municipality from the INEGI website.[109] While both birth and death records were anonymised, birth records contained information on the date of birth and municipality of residence at birth, while death records contained information on the date of death, municipality of residence at death, and age at death. I summed all deaths that occurred in the years 2009-2010 among neonates under 28 days of age by municipality, and summed all births by municipal-

ity over the same time period: these summed values comprise the CRVS observations that were used as input to the geostatistical model. Less than 0.1% of all birth and deaths were registered to a state of residence, but not a municipality of residence; these observations were distributed across municipalities within a state in proportion to the number of births and deaths registered to each municipality. Neonatal deaths were assumed to occur in the same municipality where a child was registered at birth.

I downloaded and prepared BH data from the 2010 Mexican Population and Housing Census, which was conducted in May through June 2010. This census administered a household survey to a sample of census respondents: anonymised copies of these these survey responses were shared publicly on-line, along with the sample weights and municipality of residence corresponding with each household.[110] The household survey asked all women over 12 about the date of birth of their most recent child, whether the child was still living, and if not, that child's age at death. I censored these birth history observations to exclude all births prior to January 2009, corresponding to the first observed month in 2009-2010 CRVS data, and after April 2010, one month prior to the conclusion of data collection. Among all remaining birth observations, neonatal mortality was coded as any death occurring under 28 days of age. To obtain the denominator for NMR observations from the census, I summed all births observed during this time period by municipality; to calculate the numerator, I multiplied birth denominators by a weighted mean of neonatal mortality observations across households in each municipality, weighting household-level mortality observations inversely to their sampling probabilities. I also estimated social and demographic covariates from census microdata by taking survey-weighted means of household responses by municipality.

All municipalities were matched to a shapefile, published by INEGI, which corresponds to the 2010 Population and Housing Census enumeration boundaries.[111]

2.2.1.1 Grouping of municipalities by social exclusion

Previous studies of CRVS completeness across Mexico suggest that while birth and death registration are close to complete across most municipalities, registration remains incomplete in a subset of municipalities where residents are unable to access government services due to social exclusion. Therefore, an index of social marginalisation based on factors identified in previous studies of birth and death

under-registration can be a useful method for informing prior estimates of CRVS bias in a joint mortality estimation model.[16,112] While a previous study in Mexico used the Human Development Index (HDI) as a measure of social exclusion by municipality,[16] the inclusion of mortality as one component of the HDI presents a circularity problem given that mortality is also a desired output of this model. Here, I group municipalities based on four dimensions of social exclusion measured by the 2010 Population and Housing Census: the literacy rate among adult women, the proportion of adults employed in the formal economy, the number of health clinics *per capita* within a municipality, and the proportion of residents who self-identify as indigenous. Of these indicators, the first two are associated with the "knowledge" and "standard of living" components of the HDI, while the latter two correspond to barriers specifically identified in previous studies of birth registration and maternal care in Mexico.[112–114]

I placed Mexican municipalities into one of three groupings based on these relevant dimensions of social marginalisation. Municipalities with a majority of indigenous residents, no clinics, and *either* a formal employment rate of less than 25% *or* a literacy rate of less than 75% among adult women were assigned to the most severe level marginalisation grouping: this grouping covered 6.2% of municipalities (152/2,441). A second group of municipalities with a majority of indigenous residents, no clinics, and *either* a formal employment rate of less than 50% *or* a literacy rate of less than 90% among adult women were assigned to a moderately marginalised group: this grouping covered 17.9% of municipalities (437/2,441). All other municipalities (75.9%, 1,852/2,441) were assigned to the least marginalised grouping, corresponding to the observation from previous studies that CRVS under-registration is concentrated within a relatively small number of municipalities.[16] As shown in Figure 2.3, marginalised municipalities assigned using this standard are largely concentrated in the southern states of Guerrero, Oaxaca, Chiapas, and Yucatan.

2.2.2 Joint estimation of neonatal mortality and CRVS data completeness

Here, I present a new spatial model that simultaneously estimates neonatal mortality and CRVS completeness across small spatial areas, using data from both BH and CRVS sources in a single country. This spatial modelling approach extends the principles of generalized linear mixed-effects models. Mixed-effect models refer to regression approaches that include both fixed effects, which fit a

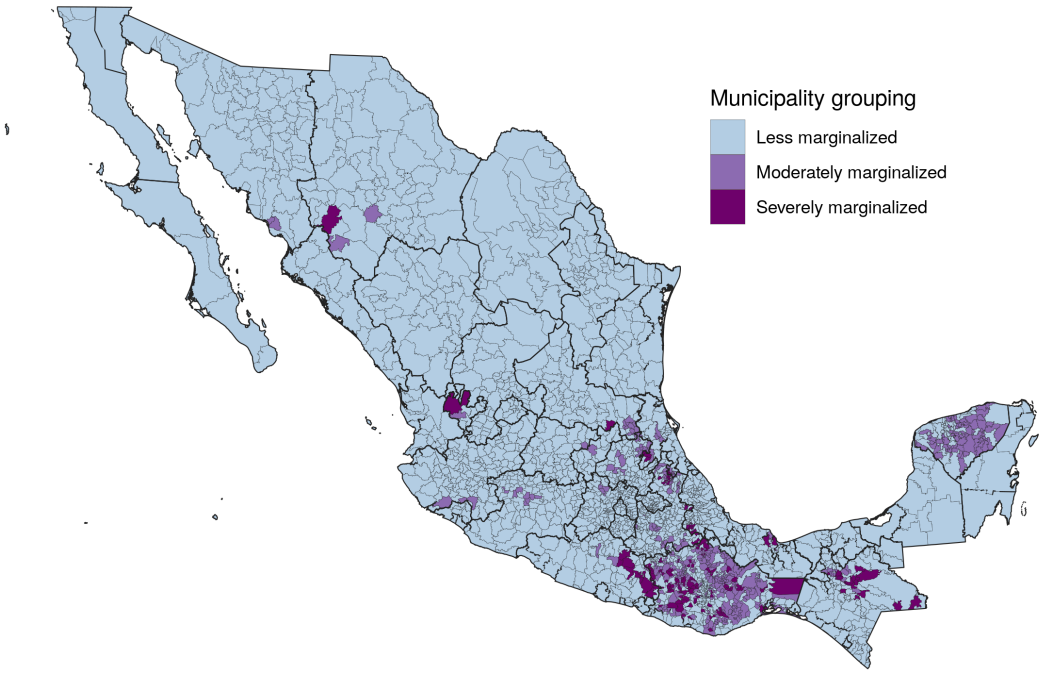


Figure 2.3: Grouping of municipalities into exclusion categories based on indigenous status, proximity to clinics, adult employment in the formal economy, and literacy rates among adult women.

parametric relationship with a predictive covariate; as well as random effects, which are nonparametric terms that account for variation between groupings, such as municipalities or years, that is not explained by the fixed effects. Generalized linear models extend this regression approach further by incorporating a link function, such as the natural logarithm, to better approximate the data generating process.[115] Spatial models are generalized linear mixed-effects models that fit a random effect based on the spatial location of data points. This effect captures the influence of latent predictors that, while not directly observed, are structured predictably in space.[57]

For this spatial model, the two outcomes of interest are the probability of death before reaching one month of age (${}_{1mo}q_0$ in demographic notation, which I will refer to as Q in the following definitions), and the ratio between birth and death under-reporting that expresses itself as bias in CRVS estimates of neonatal mortality (which I will refer to as π in the following definitions). Both of these terms are defined for each municipal unit across Mexico, with municipality i denoted as s_i.

The mortality surface of interest, Q, is defined as follows:

$$log(Q_s) = \alpha + \vec{\beta} X_s + Z_s$$

Here, Q_s is a log-linear surface indexed by spatial unit s. The estimated value for Q_s is centred around an intercept α and varies according to covariate fixed effects $\vec{\beta} X_s$, with known X_s denoting predictive covariates that vary by municipality.

In addition to covariate effects, this model draws predictive power from the neighbourhood structure of municipalities across Mexico. Remaining variation not captured by covariates is fit by a spatially-structured random effect Z_s, which corresponds to the BYM2 spatial model formulation described by Riebler and colleagues.[43] This random effect decomposes the variance not explained by fixed effects into (1) spatial and (2) independently, identically distributed (IID) components, with model parameters identifying the overall variance of Z_s as well as the proportion of excess variation associated with spatial autocorrelation. The spatial autocorrelation term by this model corresponds to the variation in each municipality's random effect term that can be explained by the average effect terms across all neighbouring municipalities.[44] For this spatial model, I defined the neighbourhood of a given municipality as all directly adjacent municipalities, based on a shapefile of municipalities from INEGI.[111]

The number of deaths from aggregated birth histories collected in the 2010 census, D_{BH}, is fit to a Poisson distribution centred around the number of births from these birth histories, N_{BH}, multiplied by the probability of dying before 1 month of age in each municipality, Q_s. Because the census household survey only collected information about the most recent birth for each adult woman, this survey may exclude a small number of children born in 2009 and succeeded by a more recent birth in late 2009 to 2010. To account for the effect of this exclusion as well as recall bias, the estimated mortality $N_{BH}(i) * Q(s_i)$ are adjusted by a bias term, γ_{BH}, which is shared across all birth history observations:

$$D_{BH}(i) \sim Poisson(\ N_{BH}(i)\ Q(s_i)\ \gamma_{BH}\ )$$

Simultaneously, the same mortality surface is fit using birth (N_{CRVS}) and neonatal death (D_{CRVS}) observations from 2009-2010 Mexican CRVS data. A bias term, π, is also used to adjust

the relationship between true mortality and the observed number of deaths in CRVS data. Unlike the constant bias term applied to birth history observations, bias terms for CRVS are fit separately for each municipality and are therefore denoted π_{s_i}:

$$D_{CRVS}(i) \sim Poisson(\ N_{CRVS}(i)\ Q(s_i)\ \pi_{s_i}\ )$$

While the separate completeness rates for birth certification and neonatal death records are not identifiable from the model, the CRVS bias terms, π_{s_i}, indicate the estimated ratio between the completeness of these two systems in a given municipality. A CRVS bias term of $\pi_A = 1$ indicates that the under-reporting rates for births and deaths are identical in municipality A, but cannot distinguish between municipalities with complete birth and death registration as opposed to municipalities where birth and death registration both 80% complete. However, values of π_{s_i} far from 1 suggest that one source is substantially less complete than another. For example, $\pi_B = .5$ indicates that death registration is half as complete as birth certification in municipality B; conversely $\pi_C = 3$ indicates that death registration is three times as complete as birth registration in municipality C.

Previous investigations have identified how Mexico's indigenous communities face greater barriers to accessing delivery and postnatal health care, a form of structural violence leading to greater neonatal mortality as well as incomplete registration of births and neonatal deaths among these communities.[112–114] Based on these findings, I pool information about the relative completeness of birth and death registration across the three social marginalisation groupings, $group(s_i)$. The CRVS bias terms for municipalities within each marginalisation group are assumed to vary according to a log-normal distribution centred around 1, with a pooled variance shared across the group:

$$log(\pi_{s_i}) \sim N(0,\ \sigma^2_{group(s_i)})$$

Priors were set on each grouped standard deviation, reflecting past findings that birth and neonatal death registration are likely to be complete in most municipalities, and that reporting bias is more likely to arise in marginalised communities.[16,89,112,114,116] I employed penalised complexity priors, which have an appealing interpretation taking the form:

$$P(\sigma_{group_j} > U_j) = p_{tail}$$

In this specification, σ_{group_j} is the standard deviation for the natural logarithm of this log-normal distribution centred around 1 (no bias) for a particular marginalisation group; U_j is a researcher-defined upper threshold for likely values of the log standard deviation, and p_{tail} is the researcher-defined estimate probability of the "tail event" that the standard deviation will exceed this threshold.[117] The thresholds for each group can also be interpreted as the bounds within which most bias estimates are likely to fall for each marginalisation group. Table 2.1 show the prior specifications for each of the three marginalisation groups. This table shows that expected variability in CRVS bias terms are expected to increase with marginalisation.

Table 2.1: Penalised complexity prior specifications for the standard deviations of the CRVS bias terms in each of three municipality groupings. Column 3 specifies the range which 95% of municipality bias terms are likely to fall within if the standard deviation is equal to the threshold, U. Across all groupings, the true grouped standard deviations are estimated to exceed their associated thresholds with 5% probability.

Municipality grouping	Threshold	Likely range	Probability of exceeding
Less marginalized	0.048	(.91, 1.1)	0.05
Moderately marginalized	0.207	(.67, 1.5)	0.05
Most marginalized	0.821	(.2, 5)	0.05

I assigned priors to all model parameters and then fit the model using the Laplace approximation for mixed-effect parameter estimation.[118,119] The model was fit in R v.4.0.3 using the package Template Model Builder v.1.7.18.[118,120]

2.2.3 Simulation model

To determine the model's capacity to reconstruct true neonatal mortality and CRVS bias from two sources, I developed a simulation model under plausible conditions for neonatal mortality and CRVS bias in Mexico. First, I simulated neonatal mortality values by setting values for the intercept and five log-linear covariate effects, all of which were estimated based on data from the 2010 Mexican Population and Housing Census:

$$log(Q_{SIM}(s)) = -5.5 - 0.25 \times Avg\ Years\ of\ School + 0.2 \times Pct\ Low\ Wage$$
$$+ 0.5 \times Pct\ No\ Health\ Care - 0.3 \times Pct\ Electrified\ Home$$
$$- 0.1 \times Pct\ Own\ Refrigerator$$

Birth history and CRVS death observations were simulated as biased binomial draws from this surface, using the existing birth denominators set in the real data, with CRVS bias draws generated from the log-normal distributions specified above.

In small area spatial models, the correlated random effect Z_s is understood to account for latent variables that affect the outcome but are not directly observed.[121] In the simulation model, the final two covariates used to simulate the mortality surface (rates of household electrification and refrigerator ownership) are excluded from the fixed effect terms available to the model, meaning that variation from these terms must be captured by the correlated random surface. I checked the model's goodness of fit by comparing estimated parameters for covariate fixed effects, CRVS bias parameters by municipality, and predicted neonatal mortality by municipality to the underlying values generated from simulation.

2.2.4 Application to neonatal mortality data in Mexico

I fit two models for neonatal mortality across Mexican municipalities using births and deaths from the 2010 census as well as 2009-2010 CRVS records.[109,110] I included six covariates observed at the municipality level, all of which came from the 2010 census: years of schooling among adult women, proportion of employed adults earning less than 2 times the minimum wage, proportion of adults without access to health care, proportion of electrified households, proportion of households owning a refrigerator, and proportion of households with piped water. I selected these covariates as indicators of spatial variation in maternal education, household wealth, and access to health care, which previous studies have identified as three key determinants of child survival.[122–124] Both model fits included a term for BH bias; the first model fit CRVS bias terms according to the model formulation described above, while the second model forced all CRVS terms to zero, essentially treating CRVS as an unbiased data source. The first model is reported below as the primary source of model results, while the no-CRVS-bias model is used as a baseline to explore the effect of the CRVS bias terms.

2.3 Results

2.3.1 Predictive validity from simulation

In general, the simulation model accurately estimated neonatal mortality across the municipalities of Mexico, and accurately recovered known parameters. Table 2.2, below, lists simulated values for fixed effects and the BH bias term along with the values recovered by the model. All covariate fixed effects and the BH bias term were accurately predicted by the model within the bounds of uncertainty.

Table 2.2: Comparison between true underlying parameter terms and estimated values, with mean and 95% uncertainty intervals, from the simulation model.

Parameter	Simulated value	Model fitted value	Overlapping UI?
FE: Years of school	-0.25	-0.27 (-0.30 to -0.23)	Yes
FE: Low wage	0.20	0.20 (0.16 to 0.23)	Yes
FE: No health care	0.50	0.50 (0.47 to 0.52)	Yes
BH bias	1.05	1.05 (0.99 to 1.12)	Yes

Figure 2.4, below, shows simulated mortality rates (left) and CRVS bias terms (right) produced in the simulation on the x-axis, while the model estimates for these terms are displayed with uncertainty on the y-axis. As shown on the left side of this figure, the simulation model produced unbiased estimates of the neonatal mortality rate by municipality, with the true mortality value falling within the model's 95% uncertainty bounds for 2434 of 2441 municipalities. The recovered estimates for the CRVS bias terms are more mixed: while the majority of mean estimates for CRVS bias match the direction (over-reporting or under-reporting) of the true term, and the 95% uncertainty intervals for CRVS bias encompass the true simulated value in 146 of 152 municipalities in the widest exclusion group (96.0%), wide uncertainty intervals preclude confident estimation about the direction of CRVS bias. Stark differences in the widths of uncertainty intervals across the three groups suggest that CRVS completeness terms for each municipality are sensitive to prior specifications.

2.3.2 Neonatal mortality across Mexico

2.3.2.1 Relationship between social marginalisation and covariates predictive of mortality

Figure 2.5, below, shows the distribution of seven social and economic variables at the municipality level by marginalisation group. By comparing the mean and inter-quartile range (IQR) of values

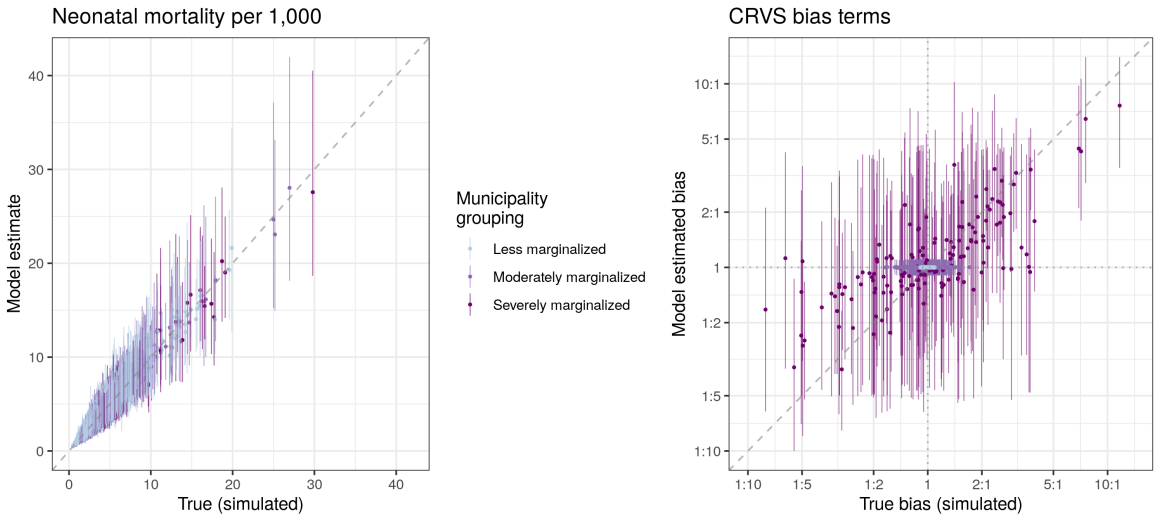


Figure 2.4: *Left* : Simulated neonatal mortality rates across Mexican municipalities compared to neonatal mortality rates recovered from the spatial statistical model. Points indicate the mean NMR estimated by the model, while vertical line spans represent the 95% uncertainty intervals for NMR estimated by the model for each municipality. *Right* : CRVS bias terms simulated for model input data, on the x-axis, compared to the mean and 95% uncertainty intervals for these bias parameters recovered by the spatial statistical model, on the y-axis.

for municipalities in the "Less marginalised" and "Severely marginalised" groupings, it becomes apparent that the differences in these groups correlate not just to factors that may affect birth and death registration, but may also have an impact on neonatal mortality rates. These differences include rates of household crowding, with an IQR of 26.5% to 39.8% among less-marginalised municipalities and 45.4% to 59.0% among severely-marginalised municipalities; piped water access, with respective IQRs of 81.6%-98.4% versus 57.3%-92.0%; and maternal literacy, with respective IQRs of 93.0%-98.1% versus 67.5%-83.7%.

2.3.2.2 Joint estimation model of neonatal mortality and CRVS bias

Figure 2.6 shows the mean estimated neonatal mortality rates across Mexican municipalities in 2009-2010. At the national level, the neonatal mortality rate was estimated to be 7.34 (7.27 to 7.42) deaths per 1,000 live births. This estimate falls slightly below the NMR estimate of 9.0 (8.4 to 9.6) produced by the UN IGME, although these differences may be due to additional data sources used by UN IGME to generate national mortality estimates.

At the municipal level, the neonatal mortality rate varied substantially, from 2.3 (1.4-3.6) in

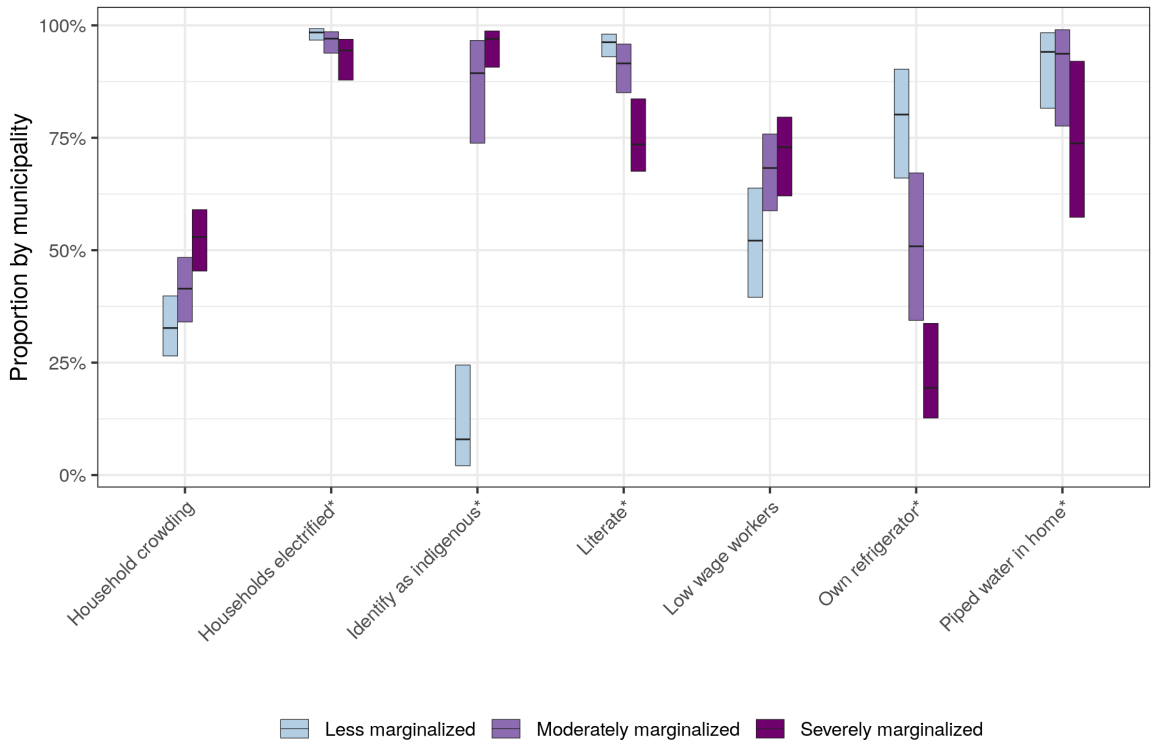


Figure 2.5: Distribution of socio-economic indicators across municipalities within each exclusion grouping. The centre line of each bar represents the median value for each indicator across less marginalised (N=1852), moderately marginalised (N=437), and severely marginalised (N=152) municipalities across Mexico, while the vertical range of each bar represents the inter-quartile range of indicator values for each grouping.

Oxchuc, Chiapas to 15.9 (13.3-18.9) in Cordoba, Veracruz. The state of Nayarit in central-west Mexico had the highest proportion of low-mortality municipalities, with twelve of its 20 municipalities measuring an NMR below 5. Conversely, the states of Mexico and Puebla in central-south Mexico, to the west and south of the Federal District, had the highest number of municipalities with mortality estimated above 10 per 1,000 (16 of 122 municipalities [13.1%] in Mexico state, and 28 of 217 [12.9%] in Puebla). Estimated counts of neonatal deaths are highly concentrated in the capitol regions of most states: of the approximately 35,000 neonatal deaths estimated in 2009-2010, 25% are concentrated in just 25 of the 2441 municipalities (the top 1%), of which 13 are state capitols.

Figure 2.7 shows estimated CRVS bias terms across the municipalities of Mexico. Estimated CRVS bias is highly heterogeneous within several states, with substantial over-reporting and under-reporting of neonatal mortality across the municipalities in the states of Guerrero, Oaxaca, Chiapas,

and Yucatan in the south of Mexico. Most municipalities with large estimated bias also rank highly on the index for social exclusion; within these; similar magnitudes of bias are observed across moderately-marginalised and severely-marginalised municipalities. Smaller bias corrections can also be observed in less-marginalised municipalities across these Mexican states.

By comparing these results to a baseline model formulation that does not include a term for municipality-specific CRVS bias, we can visualise how CRVS bias terms influence the results of the joint NMR estimation model. Figure 2.8, below, shows the difference in estimated mortality rate between the full joint model and a model where CRVS bias is forced to zero in all municipalities. At the state level, the addition of the CRVS bias term only has a minor affect, ranging from a decrease of .05 deaths per 1,000 live births in Puebla (with an NMR of 9.1 (7.0-11.9) compared to 9.2 (7.0-12.0) in the baseline) to an increase of .14 deaths per 1,000 live births in Chiapas (with an NMR of 5.3 (3.8-7.3) compared to 5.2 (3.8-7.1) in the baseline). These changes are small compared to the inter-state variation in neonatal mortality, which ranges from 5.3 (4.1-6.7) in Coahuila to 9.1 (7.0-11.9) in Puebla. However, this stability at the state level masks large differences among municipalities. Of the 2441 municipalities, 27 exhibit absolute differences of greater than .5 deaths per 1,000 between the two models, including 7 that exhibit absolute differences of greater than 1 death per 1,000. These changes can dramatically change a municipality's neonatal mortality ranking within a state: for example, out of the 570 municipalities in the state of Oaxaca, the municipality of Santa Maria Xadani's ranking in terms of neonatal mortality rose from 427 to 271 (a change of 156) after accounting for death under-reporting in the municipality. Conversely, among the 210 municipalities in the state of Veracruz, the municipality of Mixtla de Altamirano's ranking in terms of neonatal mortality fell from 59 to 157 (a change of 98) after correcting for estimated under-reporting of births in CRVS data.

2.4 Discussion

2.4.1 Performance of the simulation model

Simulation testing for this joint model formulations suggests a strong capacity to recover underlying estimates for neonatal mortality and relationships to predictive covariates under conditions that mimic the proposed data-generating process. Notably, the model generated unbiased estimates for

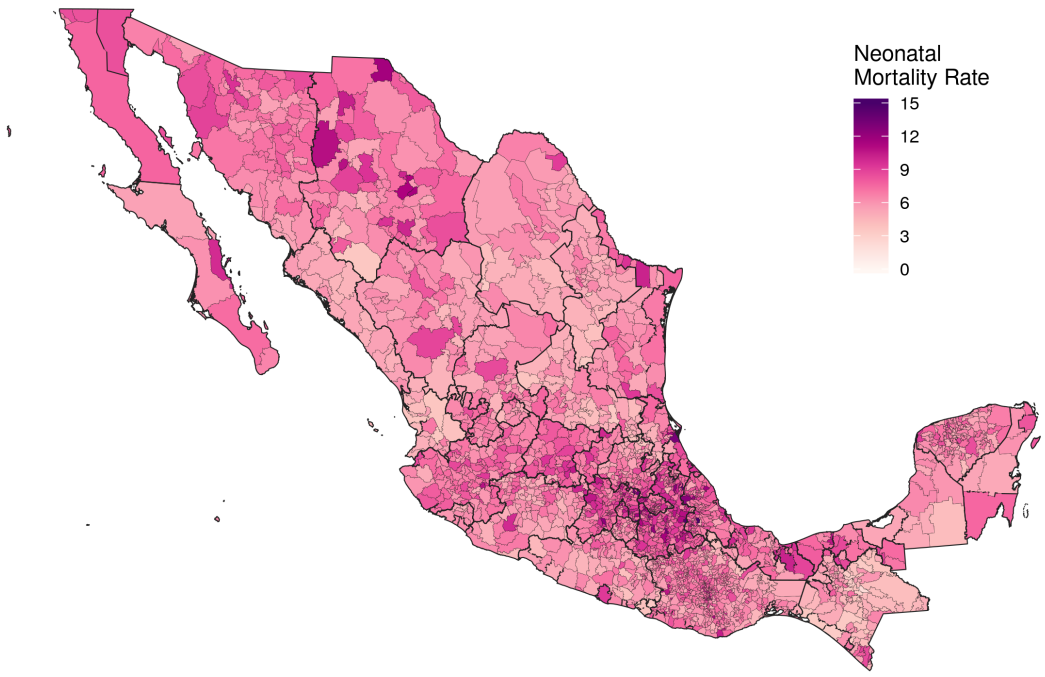


Figure 2.6: Neonatal mortality rate per 1,000 live births by Mexican municipality, 2009-2010, estimated by a joint model that incorporates both census and CRVS data.

mortality across all marginalisation groups, with relatively conservative uncertainty intervals that overlapped the true simulated mortality estimates in over 99% of municipalities. The joint model also recovered covariate fixed effects within the bounds of uncertainty even though two fixed effects used to generate the data were not included as covariates in the fitting model; this suggests that the spatial structured random effect is able to fit to excess variation not captured by covariate fixed effects. However, evident differences in the uncertainty intervals of CRVS bias estimates by marginalisation group, as shown on the right hand side of Figure 2.4, suggest that the model's estimates of CRVS bias are sensitive to misclassification between bias groupings. Additional completeness audits at the municipality level could validate CRVS bias estimates for particular municipalities and refine the process for assigning marginalisation groupings.

While the model estimates of CRVS bias terms included the simulated bias terms in 95% of the most-marginalised municipalities, the uncertainty intervals surrounding the fitted estimates were

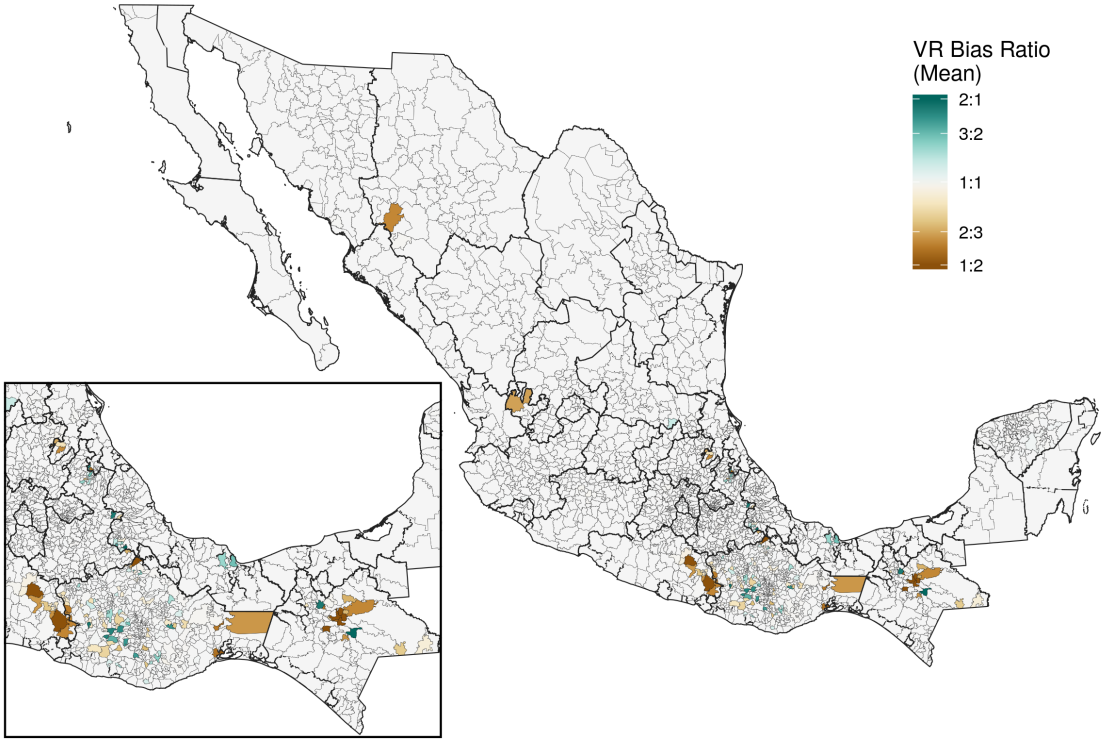


Figure 2.7: Mean estimates for CRVS bias terms predicted by the joint model for neonatal mortality. Municipalities shaded in green indicate under-registration of births relative to neonatal deaths, leading to inflated estimates of mortality based on raw CRVS data. Municipalities shaded in brown indicate under-registration of neonatal deaths relative to births, leading to under-estimation of mortality from raw CRVS data.

so wide as to preclude interpretation in most municipalities. Notably, the model precisely fits bias terms that correspond to over-reporting greater than 5:1; conversely, the estimated CRVS bias term becomes more uncertain in municipalities where CRVS deaths have been substantially under-reported. This result can be explained by the relative rarity of neonatal mortality within the study area: under these circumstances, separating observations that are biased downwards from unbiased binomial draws yielding zero or very few deaths can be problematic. This suggests that in the case of Mexico, estimates of over-reporting can be more informative for policy. In other countries and for less rare outcomes, this model may be able to better identify downwards bias in routine surveillance data.

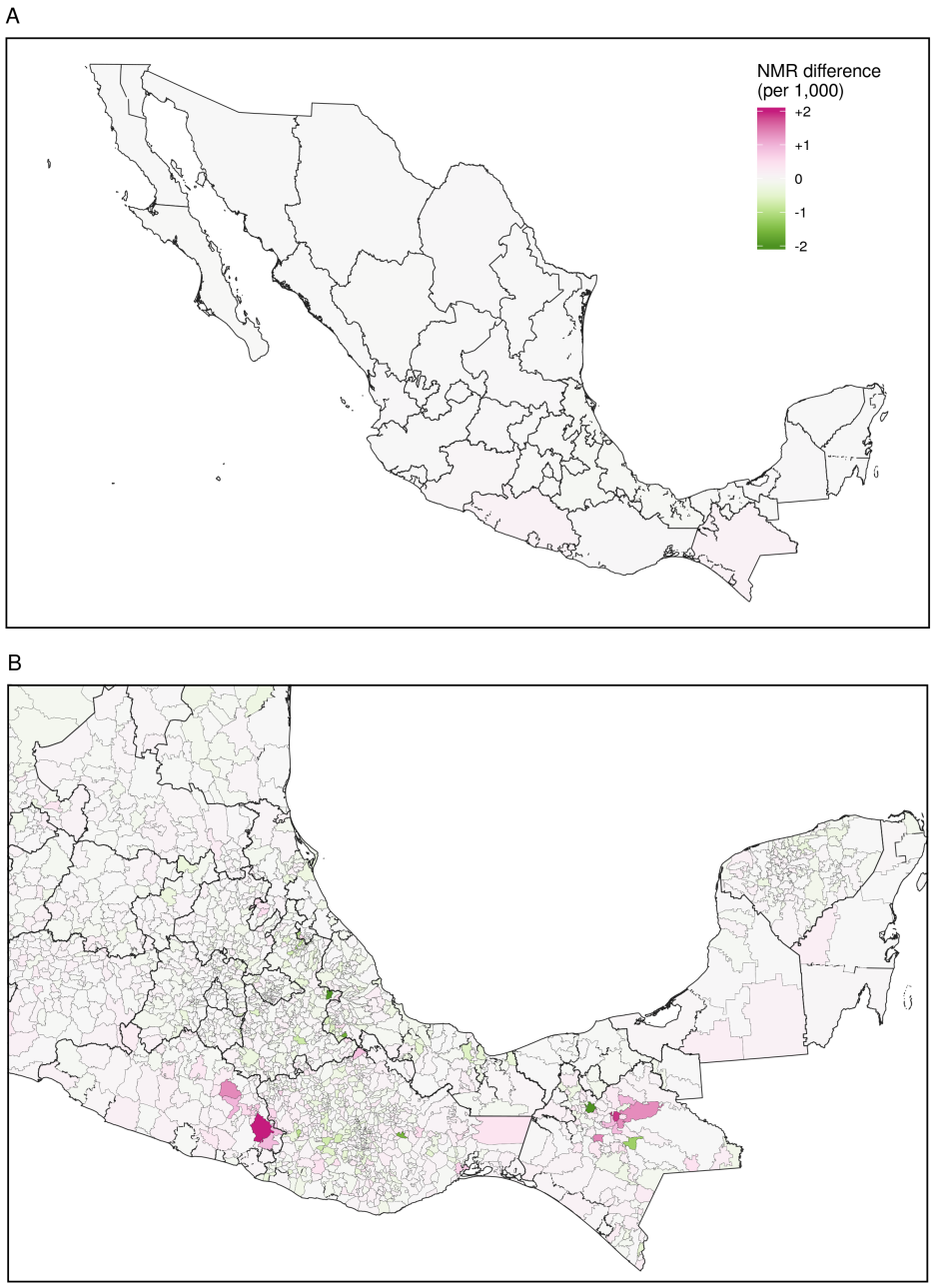


Figure 2.8: Comparison between a joint estimation model with CRVS bias terms and a baseline model where CRVS bias is set to zero. *Panel A* : The two models estimate minor differences in the neonatal mortality rate by Mexican state, with overlapping 95% uncertainty intervals for all states. *Panel B* : The two models generate substantially different and non-overlapping NMR estimates for municipalities in Guerrero, Oaxaca, and Chiapas states in the far south as well as Yucatan state in the far east.

2.4.2 Relationship between neonatal mortality and vital statistics performance

Thanks to a history of health system reform and past investigations of vital statistics completeness, the Mexican CRVS system is widely considered to be one of the highest-quality registration systems across Latin America.[84,90] Mexican CRVS data is used directly by international modelling consortia to estimate neonatal mortality at the state and national levels.[88,89] This analysis demonstrates that while Mexican CRVS data estimates levels of neonatal mortality consistent with birth histories at the national and state levels, the relationship between these data sources is more heterogeneous at the municipality level. Any spatial model of neonatal, infant, or child mortality across Mexico should account not only for small-number variation and spatial autocorrelation, but also diverse sources of bias arising by data source and municipality. These sources of bias present both a challenge for mortality estimation as well as an opportunity to further improve access to civil registration.

Serving all Mexican citizens regardless of wealth, location, or ethnic background is a major challenge for the Mexican health system;[90] this challenge extends to vital registration, which both enables service provision and is a human right in its own respect. Previous studies have identified multifarious barriers inhibiting access to formal health care among indigenous Mexican communities, including distance to health facilities, perceptions of low quality of care, limited human resources, language differences, and lack of trust between healthcare providers and recipients.[113,114] This structural violence shapes the perinatal experiences of indigenous women, who are more likely to give birth at home, separated from supportive health services and from the vital registration system.[16,112,116] In this chapter, I grouped municipalities into three categories associated with well-known dimensions of social and economic marginalisation that could present barriers to civil registration. As demonstrated in Figure 2.5, these municipality groupings also exhibit wide variation in factors that can affect neonatal health and survival, such as access to health services and household water supply. The positive correlation between the magnitude of CRVS bias and neonatal mortality rates by municipality suggests that both issues are rooted in social marginalisation: therefore, attempts to extend universal health care across the country must be linked to efforts to improve vital registration completeness, and vice versa.

2.4.3 Model limitations

Findings from this chapter should be interpreted with model and data limitations in mind. Because CRVS bias parameters are only estimated indirectly through the relationship between CRVS and BH estimates, these can suffer from wide uncertainty even when both BH and CRVS sample sizes are relatively high. This suggests that the CRVS completeness estimates should be treated with great caution, and this aspect of the results may need further refinement before it can be used to inform policy decisions. Other predictive factors could also be included to estimate the CRVS completeness surface: for example, estimates of participation in the *Seguro Popular* health insurance system as well as the *Prospera* conditional cash transfer program could be included as covariates predicting greater community integration with the formal health care system.

There are other methodological limitations to this model that should be considered in other country contexts. Because mortality estimates are primarily grounded by BH data, this model is not applicable to countries where death registration is complete but recent BH surveys have not been conducted. The current model also does not account for source-specific biases in particular BH surveys, although a larger regional model incorporating many surveys might be able to include a survey-specific random effect rather than the single BH bias term presented in this chapter.

2.4.4 Conclusions

The wide applicability of the joint BH and CRVS model makes it an appealing starting point for future research. Most major household surveys include BH questions as part of their standard questionnaire, making this method potentially usable in most low and middle income countries with a functioning CRVS system. This estimation approach also overcomes some limitations of BH-only geospatial modelling strategies, namely insufficient space-time coverage of data observations as well as relatively low sample sizes.[2,108] In countries with high-quality CRVS data, CRVS can increase the precision of neonatal mortality estimates when BH estimates are uncertain or biased.

Because this model generates estimates both of child mortality and of CRVS completeness, it can be used programmatically to target multiple aspects of health system performance. Finally,

the Bayesian modelling framework captures uncertainty both in estimates of neonatal mortality and CRVS bias, allowing for appropriately cautious interpretations of the results.

Moving beyond mortality, the multiple source estimation approach described in this chapter can be extended to map disease prevalence and incidence. In many low- and middle-income countries, notifiable infectious diseases such as tuberculosis, malaria, and HIV offer similar opportunities for spatial estimation based on a survey data source in conjunction with biased surveillance data sources.[26,125] A growing set of countries have also implemented electronic health information systems, such as the District Health Information System 2, that capture records from medical institutions and reports from field workers in a unified web platform;[126] records collected from these systems could serve as a valuable, if biased, source of health information across a growing number of countries. In the following chapter, I develop a multi-source mapping approach based on routine case notifications and a survey to describe space-time variation in tuberculosis prevalence across Uganda.

2.5 References

1. UN General Assembly. *Universal declaration of human rights* 1948.
2. Burstein, R. *et al.* Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature* **574,** 353–358. http://www.nature.com/articles/s41586-019-1545-0 (Oct. 2019).
9. AbouZahr, C. *et al.* Towards universal civil registration and vital statistics systems: The time is now. *The Lancet* **386,** 1407–1418. http://dx.doi.org/10.1016/S0140-6736(15)60170-2 (2015).
10. Rao, C. Elements of a strategic approach for strengthening national mortality statistics programmes. *BMJ Global Health* **4,** 1–10 (2019).
11. Diggle, P. J. & Giorgi, E. Model-based geostatistics for prevalence mapping in low-resource settings. *Journal of the American Statistical Association* **111,** 1096–1120. arXiv: 1505.06891. http://dx.doi.org/10.1080/01621459.2015.1123158 (2016).
13. Setel, P. W. *et al.* A scandal of invisibility: making everyone count by counting everyone. *Lancet* **370,** 1569–1577 (2007).
14. United Nations Statistics Division. *Principles and Recommendations for a Vital Statistics System Revision 3* **19.** http://unstats.un.org/unsd/Demographic/standmeth/principles/M19Rev3en.pdf (2014).
15. Fisker, A. B., Rodrigues, A. & Helleringer, S. Differences in barriers to birth and death registration in Guinea-Bissau: implications for monitoring national and global health objectives. *Tropical Medicine and International Health* **24,** 166–174 (2019).
16. Hernández, B. *et al.* Subregistro de defunciones de menores y certificación de nacimiento en una muestra representativa de los 101 municipios con más bajo índice de desarrollo humano en México. *Salud Pública de México* **54,** 393–400 (2012).
26. Rood, E. *et al.* A spatial analysis framework to monitor and accelerate progress towards SDG 3 to end TB in Bangladesh. *ISPRS International Journal of Geo-Information* **8,** 1–11 (2019).
43. Riebler, A. *et al.* An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research* **25,** 1145–1165. arXiv: 1601.01180 (2016).

44. Besag, J., York, J. & Mollié, A. A Bayesian image restoration with two applications in spatial statistics Ann Inst Statist Math 43: 1-59. *Find this article online* **43,** 1–20 (1991).
57. Banerjee, S., Carlin, B. P. & Gelfand, A. E. *Hierarchical modeling and analysis for spatial data* 2nd Editio (CRC Press, 2014).
64. Wakefield, J., Okonek, T. & Pedersen, J. Small Area Estimation for Disease Prevalence Mapping. *International Statistical Review,* insr.12400. https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12400 (July 2020).
81. Abouzahr, C. & Boerma, T. Health information systems: the foundations of public health. *Bulletin of the World Health Organization* **83,** 578–583. arXiv: arXiv:1011.1669v3 (2005).
82. Adair, T. & Lopez, A. D. Estimating the completeness of death registration: An empirical method. *PLoS ONE* **13,** 1–19 (2018).
83. Målqvist, M. *et al.* Unreported births and deaths, a severe obstacle for improved neonatal survival in low-income countries; a population based study. *BMC International Health and Human Rights* **8,** 1–7 (2008).
84. Mikkelsen, L. *et al.* A global assessment of civil registration and vital statistics systems: Monitoring data quality and progress. *The Lancet* **386,** 1395–1406. http://dx.doi.org/10.1016/S0140-6736(15)60171-4 (2015).
85. Murray, C. J., Rajaratnam, J. K., Marcus, J., Laakso, T. & Lopez, A. D. What can we conclude from death registration? Improved methods for evaluating completeness. *PLoS Medicine* **7** (2010).
86. Silva, R. *A Preliminary Evaluation of the Completeness and Quality of Death Registration Data in Selected Arab Countries* tech. rep. (Economic and Social Commission for Western Asia, United Nations, 2015), 1–20.
87. Bhat, P. N. Completeness of India's sample registration system: An assessment using the general growth balance method. *Population Studies* **56,** 119–134 (2002).
88. UN Inter-agency Group on Mortality Estimation (UN IGME). *Levels and trends in child mortality report 2020* tech. rep. (United Nations, New York, 2020). https://www.unicef.org/reports/levels-and-trends-child-mortality-report-2020.
89. Dicker, D. *et al.* Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **392,** 1684–1735. https://linkinghub.elsevier.com/retrieve/pii/S0140673618318919 (Nov. 2018).
90. Frenk, J. Bridging the divide: global lessons from evidence-based health policy in Mexico. *Lancet* **368,** 954–961 (2006).
91. Schmid, B. & da Silva, N. N. Estimation of live birth underreporting with a capturerecapture method, Sergipe, northeastern Brazil. *Revista de Saude Publica* **45,** 1088–1098 (2011).
92. De Frias, P. G., Szwarcwald, C. L., de Souza, P. R. B., da Silva de Almeida, W. & Lira, P. I. C. Correcting vital information: Estimating infant mortality, Brazil, 2000-2009. *Revista de Saude Publica* **47,** 1048–1058 (2013).
93. Szwarcwald, C. L., de Frias, P. G., deSouza Júnior, P. R. B., da Silva de Almeida, W. & de Morais Neto, O. L. Correction of vital statistics based on a proactive search of deaths and live births: Evidence from a study of the North and Northeast regions of Brazil. *Population Health Metrics* **12,** 1–10 (2014).
94. Lima, E. E., Queiroz, B. L. & Zeman, K. Completeness of birth registration in Brazil: an overview of methods and data sources. *Genus* **74** (2018).
95. Ribotta, B. S., Acosta, L. M. S. & Bertone, C. L. Evaluaciones subnacionales de la cobertura de las estadísticas vitales. Estudios recientes en América Latina [Evaluations of the Vital Statistics Coverage at a Subnational Level. Recent Studies in Latin America]. *Revista Gerencia y Políticas de Salud* **18** (2019).
96. Smith, P. J. Bayesian Methods for Multiple Capture-Recapture Surveys. *Biometrics* **44,** 1177–1189. http://www.jstor.com/stable/2531745 (1988).

97. Chandra Sekar, C. & Deming, E. On a Method of Estimating Birth and Death Rates and the Extent of Registration. *Journal of the American Statistical Association* **44,** 101–115 (1949).
98. Tilling, K., Sterne, J. A. & Wolfe, C. D. Estimation of the incidence of stroke using a capture-recapture model including covariates. *International Journal of Epidemiology* **30,** 1351–1359 (2001).
99. Hook, E. B. & Regal, R. R. Capture-recapture methods in epidemiology: Methods and limitations. *Epidemiologic Reviews* **17,** 243–264 (1995).
100. Yip, S. F. *et al.* Capture-recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology* **142,** 1047–1058 (1995).
101. Becker, S., Waheeb, Y., El-deeb, B., Khallaf, N. & Black, R. Estimating the Completeness of Under-5 Death Registration in Egypt. *Demograph* **33,** 329–339. https://www.jstor.org/stable/2061765 (1996).
102. Tilling, K. Capture-recapture methods - Useful or misleading? *International Journal of Epidemiology* **30,** 12–14 (2001).
103. Cormack, R. M. Problems with using capture-recapture in epidemiology: An example of a measles epidemic. *Journal of Clinical Epidemiology* **52,** 909–914 (1999).
104. De Almeida, W. D. S. *et al.* Captação de óbitos não informados ao ministério da saúde: Pesquisa de busca ativa de óbitos em municípios brasileiros [Capturing deaths not informed to the Ministry of Health: proactive search of deaths in Brazilian municipalities]. *Revista Brasileira de Epidemiologia* **20,** 200–211 (2017).
105. National Adminstrative Department of Statistics (DANE). *La Mortalidad Materna y Perinatal en Colombia en los albores del siglo XXI. Estimación del subregistro de nacimientos y defunciones y estimaciones ajustadas de nacimientos, mortalidad materna y perinatal por departamentos* tech. rep. (2006), 46. http://docplayer.es/40107048-Estudio-la-mortalidad-materna-y-perinatal-en-colombia-en-los-albores-del-siglo-xxi.html.
106. Schmertmann, C. P. & Gonzaga, M. R. Bayesian Estimation of Age-Specific Mortality and Life Expectancy for Small Areas With Defective Vital Records. *Demography* **55,** 1363–1388 (2018).
107. Golding, N. *et al.* Mapping under-5 and neonatal mortality in Africa, 2000-15: a baseline analysis for the Sustainable Development Goals. *The Lancet* **390,** 2171–2182. http://dx.doi.org/10.1016/S0140-6736(17)31758-0 (2017).
108. Wakefield, J. *et al.* Estimating under-five mortality in space and time in a developing world context. *Statistical Methods in Medical Research* **28,** 2614–2634 (2019).
109. National Institute of Statistics and Geography (INEGI) (Mexico). *Mexico Vital Registration 2009-2010* Aguascalientes, Mexico. https://en.www.inegi.org.mx/programas/mortalidad/.
110. National Institute of Statistics and Geography (INEGI) (Mexico). *Mexico Population and Housing Census 2010* Aguascalientes, Mexico, 2010. https://en.www.inegi.org.mx/programas/ccpv/2010/.
111. National Institute of Statistics and Geography (INEGI) (Mexico). *Geostatistical framework 2010 version 5.0 (Population and Housing Census 2010)* Aguascalientes, Mexico, 2010. http://en.www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825292812.
112. Enciso, G. F., Del Pilar Ochoa Torres, M. & Hernández, J. A. M. The subsystem of information on births. Case study of an indigenous region of Chiapas, Mexico. *Estudios Demográficos y Urbanos* **32,** 451–486 (2017).
113. Paulino, N. A., Vázquez, M. S. & Bolúmar, F. Indigenous language and inequitable maternal health care, Guatemala, Mexico, Peru and the Plurinational State of Bolivia. *Bulletin of the World Health Organization* **97,** 59–67. http://www.who.int/entity/bulletin/volumes/97/1/18-216184.pdf (Jan. 2019).
114. Gamlin, J. & Osrin, D. Preventable infant deaths, lone births and lack of registration in Mexican indigenous communities: health care services and the afterlife of colonialism. *Ethnicity and Health* **25,** 925–939. https://doi.org/10.1080/13557858.2018.1481496 (2020).

115. Dickman, P. W., Sloggett, A., Hills, M. & Hakulinen, T. Regression models for relative survival. *Statistics in Medicine* **23,** 51–64 (2004).
116. Luis, N. B. *El subregistro de nacimiento en los grupos vulnerables de México: análisis de las políticas públicas implementadas para abatirlo entre los años 2010 y 2014* PhD thesis (INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY, 2014), 1–121.
117. Simpson, D., Rue, H., Riebler, A., Martins, T. G. & Sørbye, S. H. Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science* **32,** 1–28. arXiv: 1403.4630 (2017).
118. Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H. & Bell, B. M. TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* **70.** arXiv: 1509.00660 (2016).
119. Thorson, J. T. & Kristensen, K. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research* **175,** 66–74 (2016).
120. R Core Team. *R: A Language and Environment for Statistical Computing* Vienna, Austria, 2018. https://www.r-project.org/.
121. Divino, F., Egidi, V. & Salvatore, M. A. Geographical mortality patterns in Italy: A Bayesian analysis. *Demographic Research* **20,** 435–466 (2009).
122. Gakidou, E., Cowling, K., Lozano, R. & Murray, C. J. Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: A systematic analysis. *The Lancet* **376,** 959–974. http://dx.doi.org/10.1016/S0140-6736(10)61257-3 (2010).
123. Chalasani, S. Understanding wealth-based inequalities in child health in India: A decomposition approach. *Social Science and Medicine* **75,** 2160–2169. http://dx.doi.org/10.1016/j.socscimed.2012.08.012 (2012).
124. Rutherford, M. E., Mulholland, K. & Hill, P. C. How access to health care relates to under-five mortality in sub-Saharan Africa: Systematic review. *Tropical Medicine and International Health* **15,** 508–519 (2010).
125. Dwyer-Lindgren, L. *et al.* Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature* **570,** 189–193. http://www.nature.com/articles/s41586-019-1200-9 (June 2019).
126. Dehnavieh, R. *et al.* The District Health Information System (DHIS2): A literature review and meta-synthesis of its strengths and operational challenges based on the experiences of 11 countries. *Health information management : journal of the Health Information Management Association of Australia* **48,** 62–75 (2019).

Chapter 3

Mapping the relationship between tuberculosis burden and case notifications in Uganda

3.1 Introduction

In 2015, the World Health Assembly committed its member states to the ambitious goal of ending the tuberculosis (TB) epidemic by 2035. In response, the World Health Organization (WHO) developed the global End TB strategy, which succeeded the Directly Observed Therapy (DOTS, 1994-2005) and Stop TB (2006-2014) plans as the coordinating principles for addressing the tuberculosis epidemic worldwide.[127] The End TB strategy outlines three ambitious targets to achieve by 2030: compared to 2015, the plan aims to reduce the number of TB deaths by 90%, reduce the number of incident TB cases by 80%, and ensure that 100% of families are protected from catastrophic costs due to TB.[128] These goals are consistent with Sustainable Development Goal 3.3, which commits United Nations member countries to end the TB epidemic by 2030.[129] Timely and accurate TB case detection serves as a critical component for two of the three pillars central to the End TB strategy: early detection enables for adequate patient-centred care, while mandatory and universal registration of TB disease allows countries to properly allocate resources and track progress towards goals.[130]

To achieve the End TB Strategy, the World Health Organization has developed tools for systematically building health system capacity[7,131,132] and has tracked progress and shared best practices through the release of annual Global Tuberculosis Reports.[129] Both these reports and external anal-

yses have concluded that while many countries have increased their capacity to identify and treat tuberculosis, particularly drug-resistant TB and TB among people living with HIV, the majority of high-burden countries are not on track to meet the End TB 2030 targets.[129,133] Developing national health system capacity to detect and treat people with active TB is a necessary condition for the eradication of the TB epidemic in high-burden countries.[129] Among its many benefits, successful tuberculosis surveillance allows policy-makers in high-burden countries to develop equitable interventions for communities with the highest TB burden. According to the WHO:

> Disaggregation [can] inform analysis of within-country inequalities and associated assessments of equity, with findings used to identify areas or subpopulations where progress is lagging behind and greater attention is needed. Such disaggregation is also an important consideration for the TB community, given the influence of sex, age, socioeconomic status and differential access to health care on the risks and consequences of TB infection and disease.[129]

The recent history of TB control in Uganda, one of the 30 high-burden countries recognised by the End TB strategy, exemplifies both the possibilities for health system progress and the remaining challenges to TB elimination in low-resource settings. In Uganda, TB control is overseen by the National Tuberculosis and Leprosy Programme (NTLP) within the Ugandan Ministry of Health. The NTLP 2015/16-2019/20 strategic plan mirrors the core objectives of the global End TB strategy, calling for a 5% reduction in TB incidence as well as a 30% increase in case notifications over a four-year time period.[134] While the 2019-2020 NTLP report indicated impressive gains in TB treatment since 2015, the case notification target was barely missed (26% increase as opposed to 30%), while the estimated number of total TB cases increased slightly from 2017 to 2019.[135] Two of the principle challenges to universal TB detection and treatment identified in the 2020 report were (1) low case detection rates in the far eastern and south-western regions of the country; and (2) the threat of service disruptions related to health crises such as COVID-19, which could reduce health system capacity and adherence to TB treatment regimens.[135]

Tuberculosis control programs in high-burden countries need tools to track progress towards TB case detection and treatment goals across the country using routinely-collected data. In this

chapter, I present a model for tracking both tuberculosis prevalence and case notification completeness at the district level across Uganda. This approach synthesises multiple data sources available to the Uganda NTLP to provide more precise and reliable estimates of district-level TB burden than existing methods.

3.1.1 TB mapping in high-burden settings

In recent years, TB control programs in lower-resource settings have shown increasing interest in small-area and spatial analyses of tuberculosis burden.[136] The prevalence and incidence of active TB disease are expected to vary locally within a country based on differences in underlying risk factors for the disease; by identifying these differences, programs can more efficiently allocate resources to target locations and sub-populations with the highest TB burden.[26]

In countries with high TB burden, primarily in low-income settings, two data sources have historically been used to generate estimates of TB burden. TB prevalence surveys, considered to be the "gold standard" of TB data collection, estimate TB prevalence at the national level based on sampled cluster locations across a country. These surveys use a well-defined screening process to identify potential TB cases within the eligible population, typically 15 years of age and above, and then conduct laboratory tests on the sputum of potential cases to confirm the presence of TB.[137] Depending on the underlying TB prevalence in a country and the overall sample size, each TB prevalence survey may identify as few as several dozen TB cases. Given the high costs associated with conducting TB prevalence surveys, most high-burden countries have conducted no more than two of these surveys over the past two decades.[129,137]

A more abundant data source for TB burden is available to countries in the form of aggregate case notifications reported to national TB control programs. Most countries mandate hospitals to report diagnosed TB cases, along with select demographic and diagnostic information, at regular intervals; in countries using electronic health information systems such as the DHIS2, reporting and aggregation of notified TB cases may happen continuously.[126,132] Aggregate case notifications are a potentially appealing source of subnational data due to their broad geographic coverage and wide availability. However, in many cases, counts of active TB diagnoses do not fully capture true TB

incidence due to lack of access to health care, insufficient capacity to diagnose, or misdiagnosis with another lung condition.[138] Even clinically-diagnosed TB cases, particularly those diagnosed in private settings, may be missed in national statistics due to gaps in the reporting cascade.[25,136,137]

Although previous studies have explored spatial variation in TB incidence and prevalence within high-burden countries, no widely-accepted standards exist for assessing subnational variation in TB. In a recent systematic review, Shaweno and colleagues identified 168 studies that conducted spatial analyses of TB, including 57 in high-burden settings.[139] Of the 168 identified studies, 161 (96%) used TB case notifications as the data underlying the spatial analysis, although no studies in the systematic review accounted for spatial variation in case notification under-reporting. This oversight may be deeply problematic in high-burden settings: in another spatial assessment conducted in Bangladesh, Nepal, and Pakistan, Van Gurp and colleagues found that the estimated completeness of TB case notifications were correlated with programmatic factors such as the testing rate, facility density, and test positivity rate.[140] The authors concluded that subnational TB case notification rates were more likely to reflect subnational variation in healthcare access and programmatic factors than underlying TB burden. Factors such as physical location, gender, knowledge of TB, and stigma can all influence treatment seeking behaviour,[141–143] further complicating the interpretation of case notification rates across space and over time.

The limitations identified in case notifications present a major challenge to assessing spatial variability in TB burden across low-resource settings. Given the small sample sizes and low spatial coverage of many TB prevalence surveys, data from these surveys alone may be insufficient to estimate subnational variation in TB burden without the addition of other predictive spatial data.

In the current chapter, I investigate whether case notification data can be combined with TB prevalence data to estimate spatial variation in TB burden under the assumption that notification completeness also varies across space and time. I develop this exploration in the context of Uganda, a high-burden country with a single prevalence survey conducted in 2014/15 as well as an annual series of subnational notifications.

Uganda's diverse geography and population both imply that the risk factors for tuberculosis are unevenly distributed across the country: for example, the prevalence of HIV, a leading risk

factor for TB, was estimated to vary up to eight-fold across the districts of Uganda as of 2017.[125] Previous studies and the experiences of the NTLP confirm that TB burden can vary widely across Uganda's districts and regions.[134,138,144] As of 2019, Uganda was divided among 128 administrative districts grouped into four regions. Figure 3.1 shows Uganda's regions and major population centres to provide context for the data sources and results presented below.



Figure 3.1: Regions and population centers of Uganda.

3.2 Methods

In this chapter, I develop a small area modelling approach that simultaneously estimates TB prevalence and case notification completeness by district across Uganda. This approach synthesises data from a single national TB prevalence survey and an annual time series of case notifications, two data sources that are widely used for TB surveillance in high-burden settings. Figure 3.2, below, summarizes the modelling process described in this chapter. I first develop cross walks between the differing

age groups, case definitions, and indicators reported across annual notifications and the national TB prevalence survey. I then present a spatial model that incorporates TB prevalence survey data *directly* as a sample of underlying TB prevalence, and incorporates TB case notifications *indirectly* as a sample of underlying TB incidence modified by spatially-varying completeness. District-level estimates of prevalence and case reporting completeness are hypothesised to vary in accordance with predictive space-time covariates as well as the neighbourhood structure of districts across Uganda.

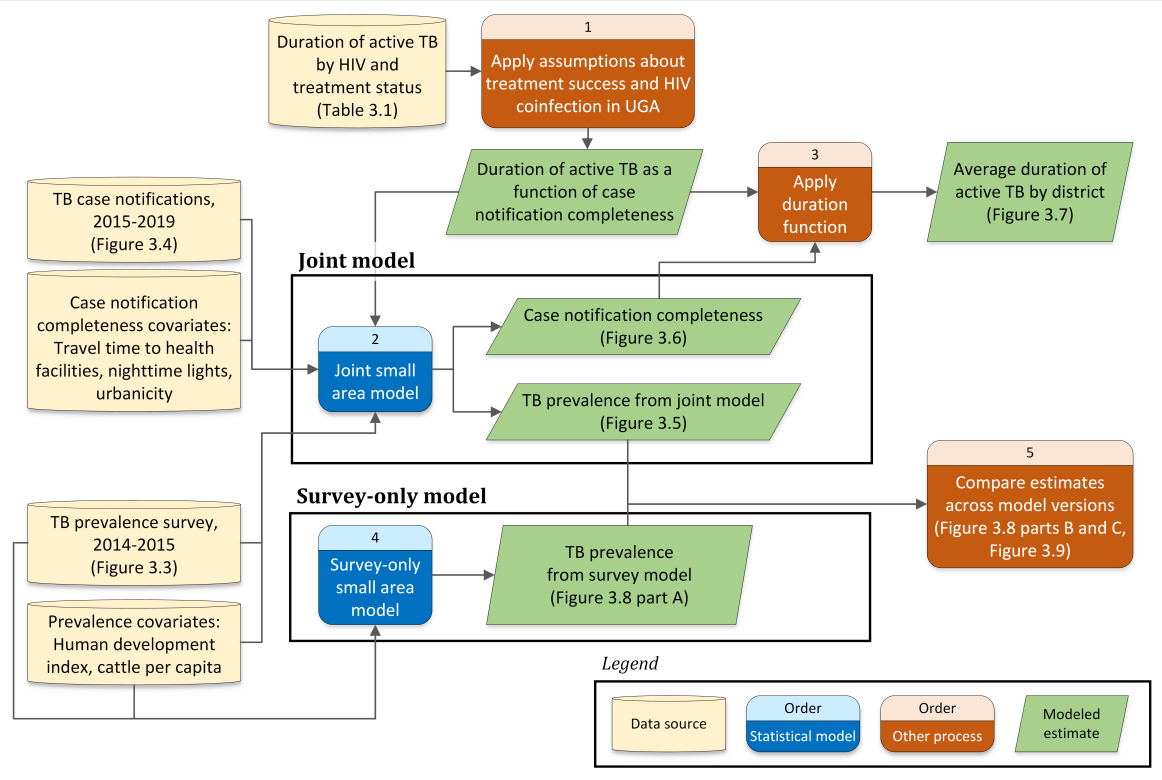


Figure 3.2: Summary of data sources, analyses, and results described in Chapter 3.

3.2.1 Data preparation

The joint spatial modelling approach combines data from a national TB prevalence survey conducted in 2014-2015, a time series of TB case notifications from 2015-2019, and predictive spatial covariates.

The number of administrative districts across Uganda increased from 122 to 128 between 2015 and 2019, with six districts being subdivided throughout this period. To simplify analysis and avoid unnecessary assumptions about the distribution of TB burden prior to district splits, I assessed

burden across 122 analytical districts that correspond to districts of Uganda as of 2015. Cases and population counts recorded in split districts were grouped and summed to match the totals that would have been recorded under the 2015 administrative groupings.

I downloaded and extracted cluster-level results from the 2014-15 Uganda National TB Prevalence Survey, including place names associated with each cluster.[145] Using spatial information from Google Maps and OpenStreetMap, I matched each cluster with the GPS location of the sampled cluster. I then extracted the number of lab-confirmed TB cases as well as the total tested population within each cluster, and summed these values to get district-level totals for each cluster. The prevalence survey tested sputum samples using both a smear test and a culture test; I included individuals with a positive culture test to match the case definition for a confirmed pulmonary TB case in a clinical setting.[146] Of the 122 analytical districts in Uganda, 56 were associated with at least one cluster from the 2014-15 prevalence survey. Figure 3.3 shows these raw prevalence estimates by district.

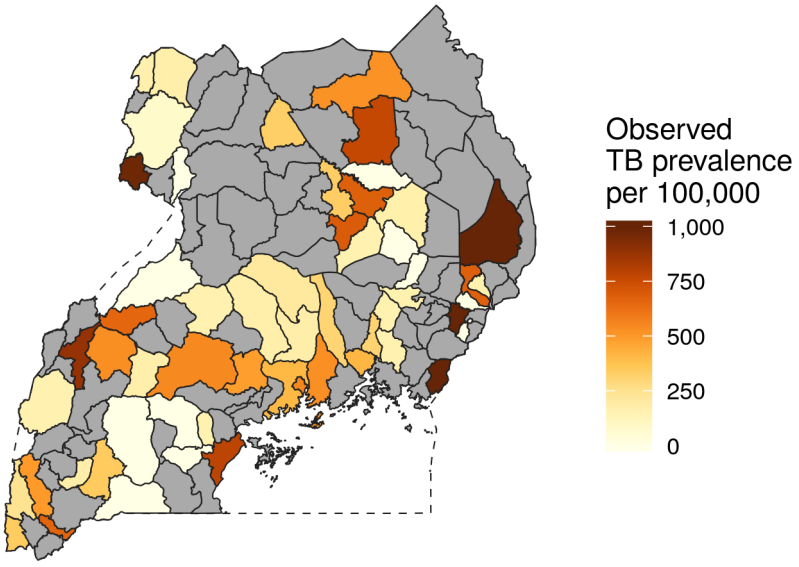


Figure 3.3: Unadjusted prevalence data from the 2014-2015 Uganda National TB Prevalence Survey, aggregated by district. Districts not containing any sampled clusters are symbolised in grey.

Total case notifications were reported by district across Uganda from 2015 through 2019 in annual reports published by the Uganda National TB and Leprosy Programme. I extracted the

total number of diagnosed individuals associated with each district and year from these reports. Although district-level case notifications were reported for all age groups combined, including both pulmonary and extra-pulmonary cases, another set of tables reported total cases by age group and pulmonary/extra-pulmonary cases across 12 program regions in the country. I associated each district with its associated TB program region, and then corrected the total number of cases downwards based on the estimated fraction of total TB cases which were pulmonary cases among adults aged 15 and over. Across the 12 TB program regions, the proportion of diagnosed extra-pulmonary cases ranged from 2% in the Hoima program region, near Lake Albert in the west, to 19% in the Moroto program region, which borders Kenya in the far east. Meanwhile, the proportion of cases diagnosed in children under age 15 varied from 6% in the Lira program region in Northern Uganda to 19% in the Moroto program region.[147] While the age structure of districts could affect the proportion of children under age 15 diagnosed within each district, a preliminary investigation revealed that the proportion of under-15s varied by less than 5% across any two districts within the same program region. Gridded estimates of the over-15 population from the WorldPop project were aggregated to the district level and used as population denominators.[55]

Figure 3.4 shows the case notification rate for pulmonary TB among adults aged 15 and above for the two years where district-level case notifications were reported. At the national level, the case notification rate for TB increased from 119 per 100,000 people in 2015 to 150 per 100,000 in 2019.[135]

I also include spatial covariates in this model to inform variation in both the TB burden and case reporting completeness surfaces. The TB prevalence surface varies according to known risk factors for tuberculosis in the context of Uganda: these included subnational estimates of the Human Development Index as well as cattle per capita, used as a proxy to identify pastoral communities that may have different TB transmission dynamics.[148] Following Van Gurp *et al.*,[140] I included covariates for the reporting completeness surface that are associated with health system access: these included average travel time to health facilities, night-time light intensity, and urban land cover.[66,149] All gridded covariates were aggregated to the district level using population weights derived from the WorldPop project.[55]

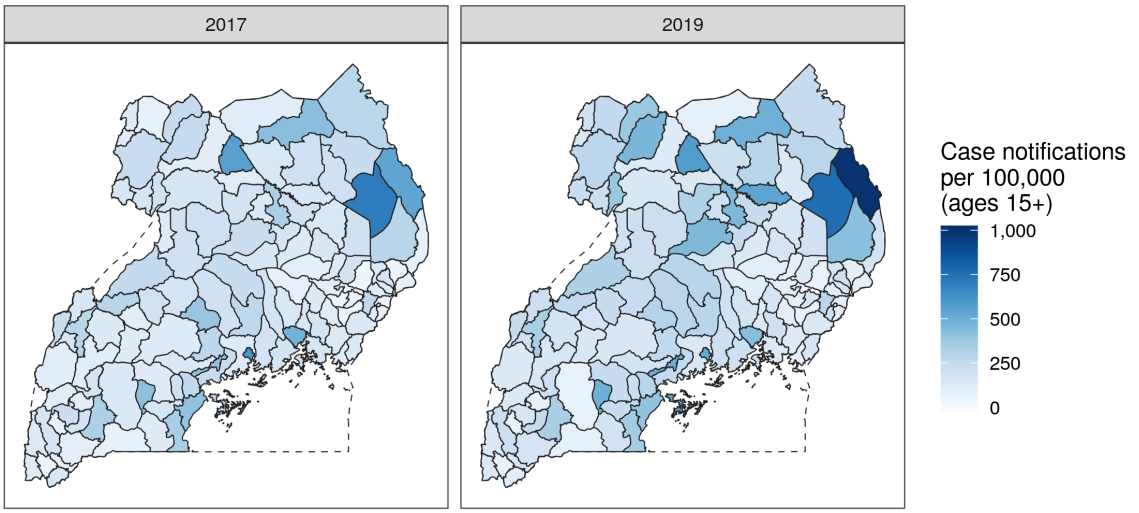


Figure 3.4: Reported case notification rates for pulmonary TB among adults aged 15 and above in 2017 and 2019, the two years for which notification data was available at the district level.

Because the reporting completeness surface is fit based on relationships between case notifications and estimated prevalence, the three covariates used to fit reporting completeness were truncated to the range of values observed in the 56 districts with both recorded case notifications and TB prevalence survey data. This trimming process, conducted before model fitting and analysis, removed unobserved covariate values that might have yielded extreme model predictions for completeness in districts without prevalence survey data.

3.2.2 Estimated duration and incidence-prevalence ratio of TB

In populations where disease burden is relatively stable, the prevalence and incidence of a disease are related by the expected (mean) duration of the disease: $Prevalence = Incidence * E[Duration]$.[150] Estimates of TB duration by district are therefore necessary to cross-walk between data sources designed to measure TB prevalence and incidence.

The mean duration of tuberculosis in a population can vary based on factors such as HIV prevalence, rates of case detection and treatment, and delays between the onset of active TB and the beginning of treatment. If an active TB case is immediately treated, the preferred treatment strategy for drug-susceptible TB in HIV negative adults includes 26 weeks, or approximately 6 months, of chemotherapy.[151] At the upper range of duration, natural history studies suggest that the mean

duration of untreated pulmonary TB among HIV negative adults is approximately three years.[152] The average duration of clinically active TB among adults with HIV is likely to vary based on antiretroviral therapy (ART), possible interactions between drugs used to treat TB and HIV, and higher mortality among individuals co-infected with HIV-TB.[152,153] Treatment duration is longer for people with drug-resistant TB patients, although the recommended treatment regimen has recently been shortened to 9-12 months for people without a prior history of TB treatment, compared to historical courses of nearly two years.[154,155] A reasonable estimate for the expected duration of active TB in a population should combine the experiences of untreated and treated individuals, accounting for HIV co-infection and possible treatment delay, and should fall between the extremes of 6 months and three years.

Two research groups at the Institute for Health Metrics and Evaluation (IHME) and the WHO produce annual estimates of global TB burden,[129,156] and these groups apply differing approaches to estimate the expected duration of TB by country. The IHME TB group estimates duration as a function of health care access and quality in each country, quantified by a Health Access and Quality (HAQ) Index.[157] Average duration is calculated as a linear interpolation between six months (for the highest possible HAQ score) and three years (for the lowest possible HAQ score). Using this approach, IHME estimates the average duration of active tuberculosis in Uganda to be 2.15 (95% uncertainty interval [UI]: 2.11 to 2.18) years as of 2019.[156] Conversely, the WHO TB group estimates plausible duration ranges of TB disease by HIV and treatment status, then estimates duration at the national level as a population-weighted combination of these groupings.[158] These groupings, plausible duration distributions for individuals within each grouping, and expected duration across each grouping are shown in Table 3.1, below. The expectation of duration across each grouping is calculated as $E[Uniform(a,b)] = (a+b)/2$.

Table 3.1: Plausible duration of clinically active tuberculosis by HIV coinfection and treatment status, as published alongside the WHO Global Tuberculosis Report 2020.

Group	Case category	Distribution of duration (years)	E[duration]
a	Treated, HIV negative	Uniform(0.2-2)	1.100
b	Not treated, HIV negative	Uniform(1-4)	2.500
c	Treated, HIV positive	Uniform(0.01-1)	0.505
d	Not treated, HIV positive	Uniform(0.01-2)	1.005

In this chapter, I apply the WHO approach to estimate TB duration by Ugandan district based on expected rates of TB-HIV co-infection and TB treatment. This approach was preferred because estimates for the Health Access and Quality score have not been produced at the subnational level and would be difficult to infer given available data. The formula I apply to estimate expected duration in each district is a weighted average across the four groups identified in Table 3.1, adjusting for treatment delay:

$$E[D_{Total}] = (E[D_a] + L_a)Prop_a + E[D_b]Prop_b + (E[D_c] + L_c)Prop_c + E[D_d]Prop_d$$

In this equation, $E[D_{<a,b,c,d>}]$ represents the expected (average) duration within each WHO grouping specified in Table 3.1, while $Prop_{<a,b,c,d>}$ represents the proportion of individuals with active TB that fall within each grouping. The variables L_a and L_c denote the average delay between the onset of active TB and the beginning of treatment for HIV negative (a) and HIV positive (c) adults who receive treatment. Based on a 2014 study of two districts in Uganda which estimated an average treatment delay of four weeks in which the median delay to treatment for active TB was found to be four weeks (0.077 years),[159] I set $L_a = L_c = 0.077$.

While Dwyer-Lindgren *et al.* estimated that HIV prevalence among adults varied from 1.7% to 11.5% across the districts of Uganda as of 2017,[125] the relationship between HIV prevalence, TB prevalence, and HIV-TB co-infection at the subnational level is not well-understood. The Global Burden of Disease Study estimates that 41.3% of all prevalent TB cases in Uganda were HIV-TB co-infections as of 2017.[156] This aligns with the proportion of adults diagnosed with TB who are co-infected with HIV in Uganda as reported by case notifications: as of 2017-2018, the Uganda NTLP reported that the TB/HIV co-infection rate was 40% based on case notifications data.[147] However, the 2014-2015 National TB Prevalence Survey in Uganda estimated just 27% adults with active TB were co-infected with HIV. A WHO modelling study observed this same discrepancy, with case notifications reporting higher TB/HIV co-infection rates than prevalence surveys, across seven countries with high TB and HIV prevalence.[160] Because the causes underlying this discrepancy are not well-understood, it is unclear which data type more accurately estimates the TB/HIV co-infection rate in Uganda.

Given uncertainty around district-level variation in HIV-TB co-infection and treatment rates across Uganda, I made simplifying assumptions to estimate $Prop_{<a,b,c,d>}$ as a function of TB case notification reporting. I assumed the nationwide TB/HIV co-infection rate to be 41.3% based on Global Burden of Disease estimates, which also align with findings from case notifications.[156] I applied this ratio uniformly across the country to estimate the relative prevalence of HIV positive and HIV negative TB cases by district. Proportions of treated versus untreated cases were estimated by multiplying the TB case detection rate by the treatment success rate, which was calculated to be 71% nationwide in the Uganda NTLP annual report.[147] Finally, treatment rates were assumed to be constant for TB cases among individuals with and without HIV, so that $Prop_a = Prop_{HIV} * Prop_{Treat}$, and so on. These assumptions allow for the estimation of the proportion of all prevalent TB cases that fall into each WHO category based on the case detection rate, denoted as π:

$$Prop_a = Prop_{Treat} * (1 - Prop_{HIV}) = .71\pi * .587 = .42\pi$$

$$Prop_b = (1 - Prop_{Treat}) * (1 - Prop_{HIV}) = (1 - .71\pi) * .587 = .587 - .42\pi$$

$$Prop_c = Prop_{Treat} * Prop_{HIV} = .71\pi * .413 = .29\pi$$

$$Prop_d = (1 - Prop_{Treat}) * Prop_{HIV} = (1 - .71\pi) * .413 = .413 - .29\pi$$

$$Prop_a + Prop_b + Prop_c + Prop_d = 1$$

Using these simplifying assumptions, the average duration of TB cases in a district can be estimated as a function of the case reporting completeness π:

$$\begin{aligned} D(\pi) &= E[D_{Total}] \\ &= (E[D_a] + L_a)Prop_a + E[D_b]Prop_b + (E[D_c] + L_c)Prop_c + E[D_d]Prop_d \\ &= (1.1 + .077)(.42\pi) + 2.5(.587 - .42\pi) + (.505 + .077)(.29\pi) + 1.005(413 - .29\pi) \\ &= 1.88 - .67\pi \end{aligned}$$

I deploy this formula below to estimate the average duration of TB cases by district, and therefore the relationship between TB incidence and prevalence, as a function of TB case notification

completeness. I conducted a sensitivity analysis on the effect of varying treatment delays, HIV-TB proportions, and differential treatment between HIV positive and HIV negative TB cases on expected duration. Additionally, I compared estimates from the joint model presented below to a similar model where duration was fixed at 2.15 years, as estimated by IHME. The results of these sensitivity analyses are presented in Appendix B.

3.2.3 Small area model incorporating prevalence data and incomplete notifications

I developed a geostatistical model to jointly estimate TB prevalence and TB case notification reporting completeness by district between 2015 and 2019. TB prevalence, p, is represented by a time-invariant spatial surface that varies according to underlying covariates as well as a spatially-structured latent process. Notification completeness, π, is represented as a space-time surface that varies according to underlying covariates as well as a spatially-structured latent process and fixed effect with a random slope by year.

Specifically, the TB prevalence surface is estimated to vary as follows:

$$log(p_s) = \alpha^p + \overrightarrow{\beta^p} X_s^p + Z_s^p$$

Here, TB prevalence p is a log-linear surface that is indexed by district (s). Prevalence varies according to an intercept, α, fixed effects $\overrightarrow{\beta^p}$ on spatial covariates X_s^p that vary by district, and a structured random effect Z_s^p that fits excess variation not captured by fixed effects. The latent spatial surface is parametrised using the BYM2 spatial model formulation described by Riebler and colleagues.[43] The superscript (p) differentiates these terms from a similar spatial field used to estimate reporting completeness (π).

Simultaneously, the reporting completeness surface is estimated to vary as follows:

$$logit(\pi_{s,t}) = \alpha^\pi + \overrightarrow{\beta^\pi} X_{s,t}^\pi + (\gamma^\pi + \delta_s^\pi)t' + Z_s^\pi$$

In this formulation, reporting completeness π is a logit-linear surface that is indexed by district (s) and year (t). Completeness varies according to fixed effects on covariates that vary by district and

year, as well as a structured random effect Z_s^{π} that is parametrised in the same manner as Z_s^{p}, according to the adjacency structure of Ugandan districts. Reporting completeness also varies according to a fixed effect on time with a random slope ($\gamma^{\pi} + \delta_s^{\pi}$). This term is multiplied by the year (t') of a particular observation, where the years 2015-2019 have been rescaled to have mean zero and range one.

Data from the TB prevalence survey, with numerators Y_i^{prev} and denominators N_i^{prev}, are evaluated directly against the underlying prevalence surface:

$$Y_i^{prev} = Poisson(N_i^{prev} * p_{s(i)})$$

In this formula, the subscript i is a generic index indicating a unique observation: for example, Y_1^{prev} indicates the number of people who tested positive for TB in survey cluster 1, N_1^{prev} indicates the total number of people sampled in this cluster, and $s(1)$ is the spatial location (district) associated with the first cluster. If only data from the TB prevalence survey was incorporated into the model, this modelling approach correspond to a standard small-area estimation framework. However, in addition, case notifications with reported cases Y_i^{notif} and corresponding population denominators N_i^{notif} are evaluated against the fitted surfaces as follows:

$$Y_i^{notif} = Poisson(p_{s(i)} \ / \ D(\pi_{s(i),t(i)}) \ * \ N_i^{notif} \ * \ \pi_{s(i),t(i)})$$

The observed number of case notifications are estimated to be centred around the true incident cases in the population multiplied by case notification incompleteness. The true incidence rate is estimated as the fitted underlying prevalence in a given district, $p_{s(i)}$, multiplied by the inverse of the duration, $1/D(\pi_{s(i),t(i)})$, which is calculated based on the duration formula presented in the previous section. The true underlying incidence of pulmonary TB is multiplied by the over-15 population in a given district-year, N_i^{notif}, to obtain the true estimated incident cases in each district and year. Finally, the true estimated incident cases are multiplied by the fitted case reporting completeness in a given district and year, $\pi_{s(i),t(i)}$, to obtain an estimate of the number of observed cases in a particular district and year.

I assigned priors to all model parameters and then fit the model using the Laplace approximation for mixed-effect parameter estimation.[118,119] I fit two versions of the model: one incorporating both TB prevalence survey results as well as case notifications, and a prevalence-only model incorporating just data from the TB prevalence survey to estimate the effects of adding notification data in the joint model. The model was fit in R v.4.0.3 using the package Template Model Builder v.1.7.18.[118,120]

3.3 Results

The joint spatial model identified that TB prevalence varies over 10-fold across districts in Uganda, while the completeness of TB case notification completeness varied from less than 40% to nearly 80% as of 2019, the final year of analysis. A joint model that incorporated both TB case notifications and prevalence data predicted TB prevalence with substantially greater precision than a model using only data from the 2014-15 national TB prevalence survey.

3.3.1 Spatial variation in tuberculosis prevalence across Uganda

Across the 122 analysis districts of Uganda, estimated pulmonary TB prevalence in adults ranged from 161 per 100,000 people (95% uncertainty interval: 132 to 193) in the district of Butaleja in the south-east of Uganda, to a maximum of 2,300 per 100,000 (uncertainty 1645 to 3186) in Moroto district which borders Kenya in the north-east. As shown in Panel A of Figure 3.5, estimated prevalence was generally lowest in Uganda's Eastern region and highest in the far north-eastern districts of the country. The densely-populated districts of Kampala and Arua experienced moderately high tuberculosis burden relative to the rest of the country, with estimated TB prevalence of 631 (519 to 765) and 550 (477 to 636) per 100,000, respectively.

Panel B of Figure 3.5 shows the relationship between unadjusted TB prevalence data from the 2014-2015 National TB Prevalence Survey (on the x axis) and prevalence estimates generated from the joint model (on the y axis). Given the relatively small numbers of observed cases in any individual prevalence survey cluster, the joint model generally smooths extreme observations from the prevalence survey, increasing estimates where only zero or one case was observed, and decreasing one extremely high district prevalence estimate of nearly 2,500 per 100,000 (1 in 40). These modifications

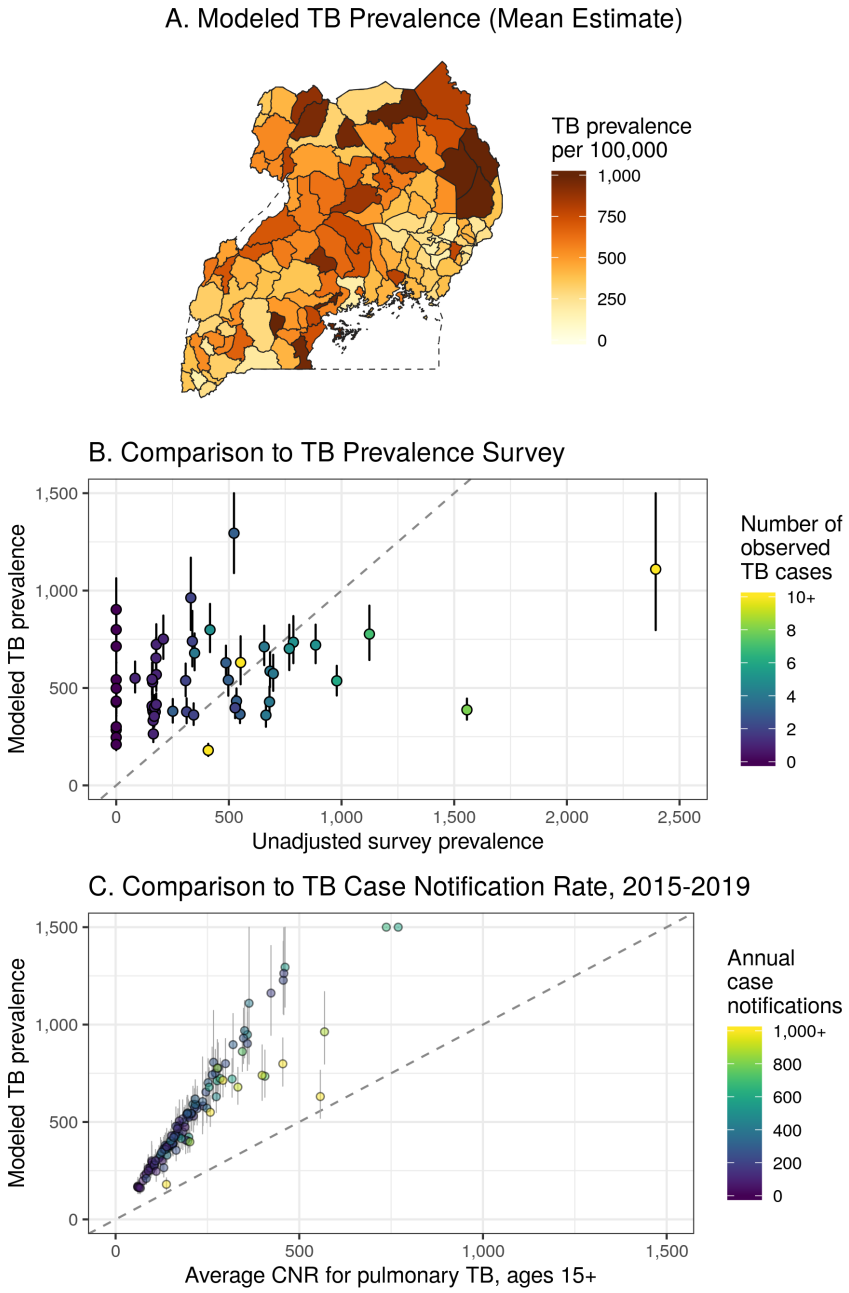


Figure 3.5: Results from the joint model compared to underlying data sources. *Panel A* : TB prevalence estimated from the joint data model. *Panel B* : scatter plot comparing modeled estimates, with uncertainty, against raw TB prevalence survey data in 56 districts. Panel C: scatter plot comparing modeled estimates by district, with uncertainty, against TB case notification rates among adults averaged across the years 2015-2019.

are due to the combined effect of the smoothing spatial random effect in the prevalence surface as well as the effects of the notifications data.

Panel C of Figure 3.5 shows the relationship between the average TB case notification rate by district (on the x axis) and prevalence estimates (on the y axis). Noting the estimated TB incidence-to-prevalence ratio of approximately 1.5 in Uganda, the model estimates higher prevalence than indicated by case notifications in every district, although the ratio between the case notification rate and estimated TB prevalence varies by district. This difference reflects the combined effects of case notification incompleteness and estimated average TB duration that fell above 1 in most districts, both of which increased estimated prevalence relative to recorded case notifications.

3.3.2 Completeness of case notifications increases over time

Figure 3.6, below, displays model estimates of case notification completeness by district for the years 2017 and 2019, the two years for which case notifications were available by district rather than by TB program region. This figure demonstrates the general increasing trend in case notification completeness over time, matching the trend of increasing completeness estimated by the NTLP at the national level.[135] At the national level, case notification completeness increased from 49% (45% to 53%) in 2017 to 55% (51% to 59%) in 2019. In 2017, estimated case notification completeness across districts varied from 27% (23% to 31%) in the rural Otuke district, in the Northern region, to 86% (74% to 94%) in Kampala. In 2019, case notification completeness ranged from an estimated 36% (32% to 41%) in the south-western Kiruhira district to 80% (68% to 88%) reporting completeness in Kampala. Population was associated with greater reporting completeness, and the populous districts of Kampala, Wakiso, and Gulu all had an estimated reporting completeness of 65% or greater in 2019.

The change in estimated reporting completeness from 2017 to 2019 in Kampala is not significant within the bounds of uncertainty. On the other hand, many districts with the lowest reporting completeness experienced significant increases between 2017 and 2019. As of 2017, 21 of 122 districts were highly likely to have reporting completeness of less than 45%, meaning that the upper bounds of the 95% uncertainty interval fell below the 45% mark; by 2019, only Kiruhira district's reporting completeness fell below the 45% threshold with high certainty.

The large majority of districts, 100 out of 122, experienced an increase in reporting completeness over the duration of the study, with an average increase of six percentage points. Although the joint

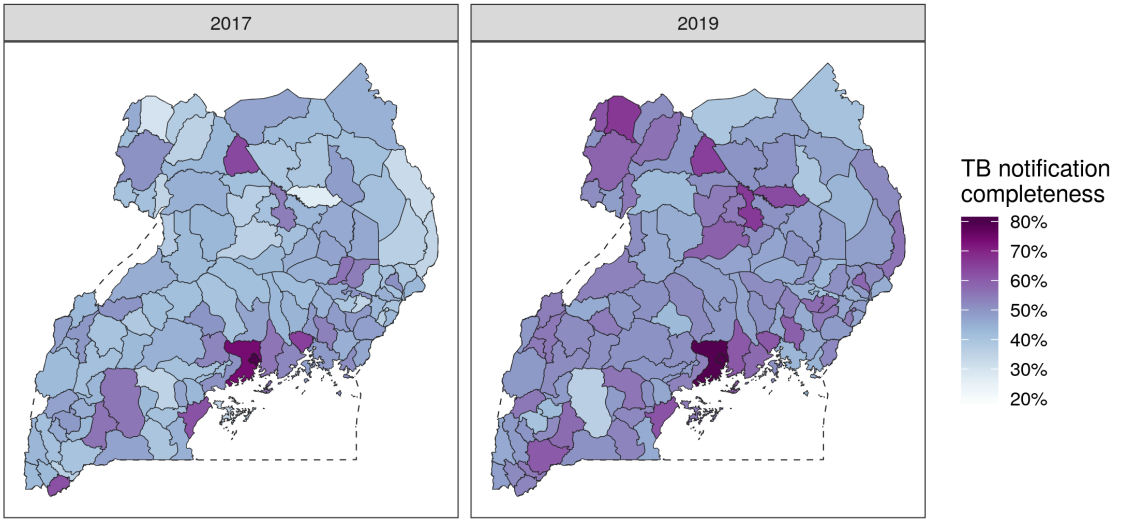


Figure 3.6: Mean estimated completeness for district-level TB case notifications in 2017 and 2019, as predicted by the joint data model. Completeness is estimated for pulmonary TB cases in people aged 15 and above.

model estimated that the remaining 22 districts experienced a decline in case notifications, that decline was not significant within the bounds of uncertainty for any district.

Based on the fitted model parameters, TB case notification rates were negatively associated with travel time to health facilities, and positively associated with night-time lights, urbanicity, and time, although collinearity between these covariates hinders straightforward interpretation of fixed effect coefficients. Additionally, the spatially-structured effect in the completeness model identified district groupings with higher or lower completeness than would be expected from the covariate fixed effects alone: this included a group of high-completeness districts surrounding Kampala in the southern Central region, and a group of lower-completeness districts in the far east of the Northern region.

Using the simplified formula for TB duration presented above, $D(\pi) = 1.88 - .67\pi$, the average duration of TB cases was estimated as a function of case notification completeness. Figure 3.7 shows the estimated duration of pulmonary TB among adults by district in 2017 and 2019. At the national level, the average duration of pulmonary TB among adults was estimated to have fallen slightly from 1.55 (1.52-1.58) years to 1.51 (1.48-1.54) years between 2017 and 2019 as a result of greater case reporting and subsequent treatment. As of 2019, the average duration of pulmonary TB was estimated to range from 1.35 (1.29-1.42) years in Kampala to 1.64 (1.61-1.67) in Kiruhira, a difference

of over 15 weeks. Given that the relationship between TB incidence and prevalence is defined based on the average duration of disease, $Prevalence = Incidence * E[Duration]$, space-time variation in TB duration mediated the relationship between prevalence estimates from the 2014-2015 National TB Prevalence Survey and incidence estimates from completeness-adjusted case notifications.

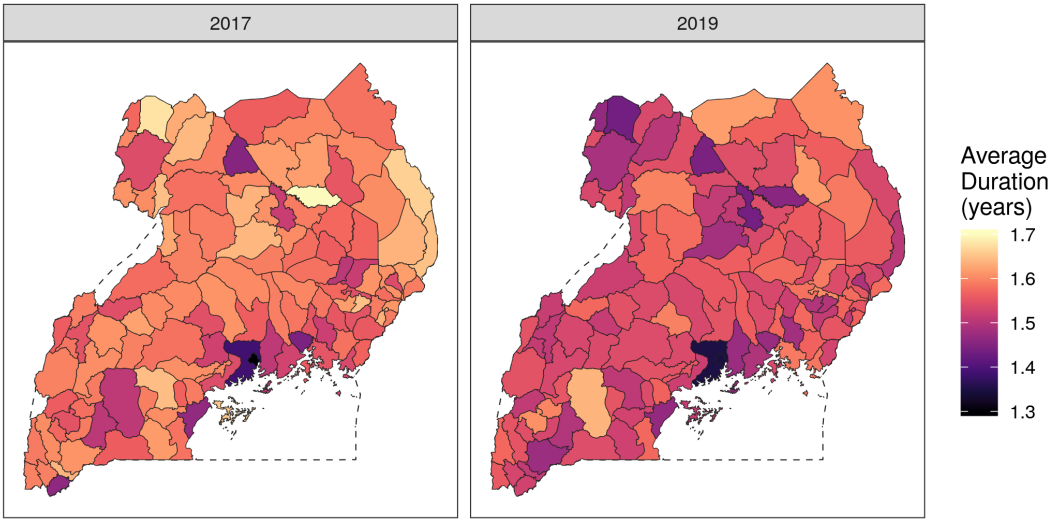


Figure 3.7: Average duration of active pulmonary TB among adults by Ugandan district, estimated as a function of case notification completeness.

3.3.3 Effect of including notifications on estimates of TB burden

A second, prevalence-only model was run using only raw data from the 2014-15 national TB prevalence survey along with the same suite of spatial covariates. This model iteration represents an attempt at district-level TB prevalence estimation without incorporating data from case notifications. Panel A of Figure 3.8 below shows the mean estimated TB prevalence based on this model. Compared to the raw data shown in Figure 3.3, the prevalence-only model captures and smooths extreme values within districts observed in the prevalence survey, but strongly smooths towards the national mean in unobserved districts. Estimates of prevalence generated from this model are also much more uncertain than the joint model. In the joint model, the width of the 95% UI for prevalence varies from 61 to 1,541 cases per 100,000, with an average uncertainty interval width of 197. In the prevalence-only model, the corresponding widths of the 95% UI for estimated prevalence ranges from 423 to 5,079, with an average spread of 1,007.

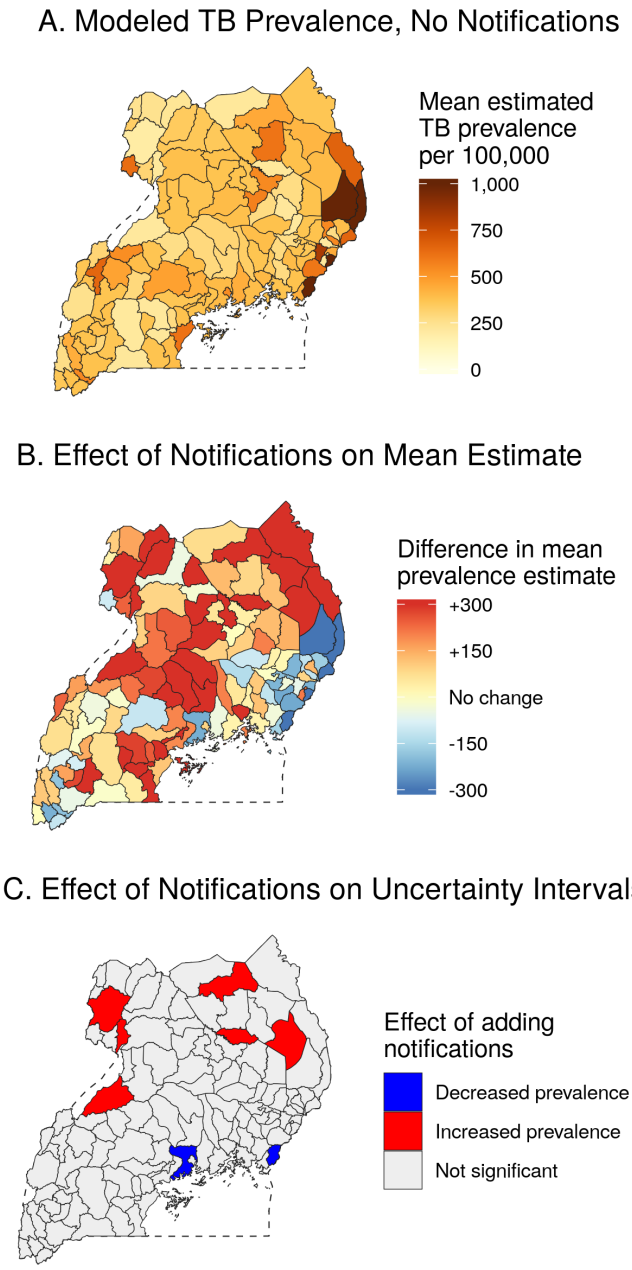


Figure 3.8: Effect of adding notifications to the joint estimation model. Panel A: Mean estimated TB prevalence based on a spatial model incorporating only data from the Uganda National TB Prevalence Survey. Panel B: Difference in mean estimated TB prevalence between the joint data model and a model that only incorporated prevalence data. Panel C: Effect of adding case notifications, taking uncertainty into account.

Panels B and C in Figure 3.8 visualise the effect of indirectly incorporating case notifications into the modelled prevalence surface. As shown in Panel B, incorporating case notifications changes

the mean estimates for TB prevalence across Ugandan districts, increasing mean estimated prevalence in the north-east and the districts surrounding Hoima while decreasing estimated prevalence in the coastal south and south-east. However, given the wide uncertainty associated with the prevalence-only estimates, just eight districts of 122 experience significant changes between the two model formulations. Of these eight districts, seven were sampled in the National TB Prevalence Survey.

Figure 3.9 demonstrates how the addition of case notification data increases model precision, enabling the identification of low- and high-burden districts with greater confidence. The figure shows model predictions for districts where TB prevalence falls below 253 cases per 100,000, the national estimate for Uganda reported by the 2014-2015 National TB Prevalence Survey, as well as districts exceeding a high prevalence threshold of 759 per 100,000, or triple that national estimate. In the context of this figure, low-confidence predictions indicate that the model's mean estimated prevalence either fell below the lower threshold of 253 or exceeded the higher threshold of 759, while high-confidence predictions indicate that both bounds of the 95% uncertainty interval for prevalence in a given district fall entirely beyond one of these thresholds. Greater uncertainty in the model without notifications, shown on the right of this figure, precludes the observation of low-burden districts with high uncertainty. In addition, the survey-only model only predicts one district to have high burden with high confidence: this district, Nakapiripirit, was the location of the survey cluster with the highest recorded prevalence from the National TB Prevalence Survey. The joint data model, shown on the right, makes more precise estimates due to the inclusion of corrected case notifications that supplement the prevalence survey data. With high confidence, the joint model predicts that 8 of the 122 districts fall below the lower prevalence threshold and 13 of 122 districts exceed the upper prevalence threshold.

3.4 Discussion

In this chapter, I demonstrate a novel framework for estimating TB prevalence by district across Uganda by synthesising data from a TB prevalence survey and annual case notifications. A joint estimation model compensates for the weaknesses of each data source: it increases the precision of estimates available from either source alone while also correcting for systematic under-reporting in case

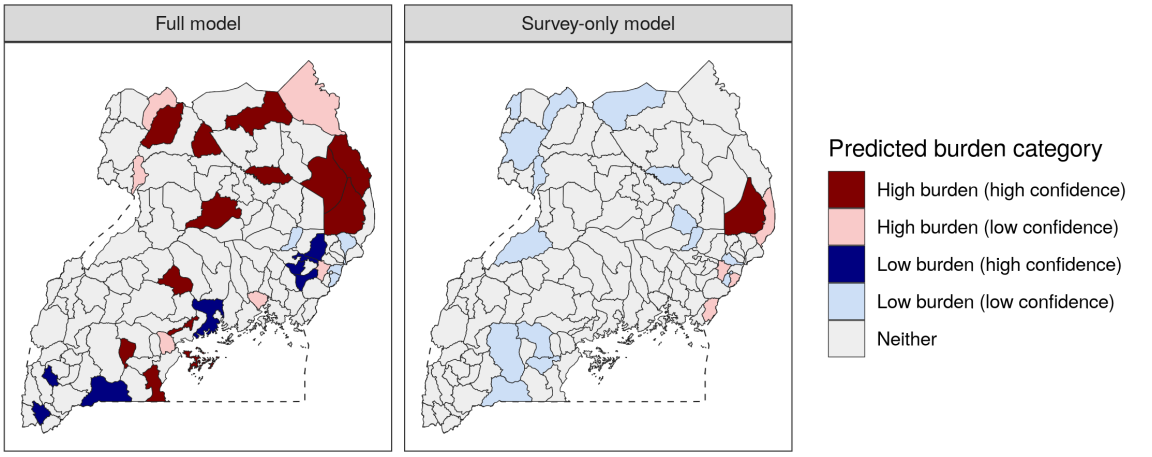


Figure 3.9: Comparison of estimates from the joint data model (left) and a model using only prevalence survey data (right) against a low prevalence threshold of 253/100,000 and a high threshold of 506/100,000. A high-confidence estimate indicates that the 95% uncertainty bounds uniformly pass a threshold, while a low-confidence estimate indicates that the model's mean estimate in a district passes a threshold but the 95% uncertainty bounds include both sides of that threshold.

notifications data. This approach corrects for weaknesses in previous spatial modelling frameworks for tuberculosis, which have primarily relied solely on just one of these two data types.[139]

TB prevalence surveys are generally not powered to predict subnational variation in TB prevalence.[137] Because of the larger sample sizes and time series associated with TB case notifications, this joint model generates more precise estimates of TB prevalence than a comparable model based on data from a single TB prevalence survey alone. This precision facilitates the identification of exemplars and hot spots for TB, allowing for more targeted and locally-appropriate policymaking. The south-eastern district of Butaleja, with an estimated TB prevalence of 161 (132 to 193), is the only district that falls below the prevalence threshold of 200 per 100,000 with high confidence; on the other hand, the four districts of Moroto and Napak have 95% UIs that fall entirely above the high prevalence threshold of 1,500 per 100,000. The stark differences between these predictions, and their high level of confidence, could be used as the basis of future studies to determine the elements of successful TB control in high-performing districts; they could also serve as the evidence base to develop new interventions in the sub-populations with the greatest burden, as recommended by the WHO End TB strategy.[129,161]

Previous research has demonstrated that unadjusted TB case notifications are an inappropriate proxy for TB incidence in high-burden settings due to confounding factors related to TB program activity and access to health care.[26,139] The space-time surface of notification completeness generated by the joint model offers a new tool for adjusting future case notifications and for targeting improvements in TB case-finding. Case notification completeness is a key indicator for tuberculosis control in Uganda, as described in the NTLP 2021-2025 strategic plan.[148] Measures of notification completeness have previously been estimated only through expensive audits in a small number of districts; this joint modelling approach, while limited by assumptions, offers a supplementary method to tracking notification completeness that can be estimated nationwide each year. For example, the model identified 10 districts where case reporting was less than half complete with 95% certainty as of 2019. These estimates could be used as the basis for developing programs that strengthen the TB case reporting cascade in the south-western districts surrounding the Mbarara regional referral hospital, where reporting remains low; as well as for increasing access to health services in north-eastern districts that are far from population centres and their regional referral hospitals. In the near future, this model could be applied to 2020 notifications to explore spatially-varying barriers to accessing TB care that may have arisen in the wake of the COVID-19 pandemic.[162]

3.4.1 Extension to other contexts

While this chapter focused on the context of Uganda, similar methods could be applied in other high-burden countries with a TB prevalence survey. According to the Global Tuberculosis Report 2020, of the 30 high TB burden countries classified by the WHO, 16 countries had completed just one TB prevalence survey since 2011.[129] Many high burden countries rely on aggregated case notifications to track annual change in TB burden: by combining notifications with data from a single prevalence survey, these programs could learn a substantial amount about subnational variation in TB control nationwide. In Ethiopia, Pakistan, and Bangladesh, previous studies have already explored factors associated with subnational variation in case reporting completeness, and this information could be used to develop highly predictive spatial covariates for a latent completeness surface.[26,140,163]

Although this model makes use of the two most common data sources for TB surveillance available in high-burden countries, additional data types could easily be incorporated into the model framework to improve the precision of estimates. Countries may conduct case inventory studies to estimate under-reporting in particular districts.[136] Capture-recapture analyses of tuberculosis cases may also reveal spatial variation in TB notification completeness, subject to the limitations of this method that have been previously discussed in Chapter 2.[99,164] Completeness estimates generated from these analyses can be directly incorporated to improve the precision of case reporting completeness, which is otherwise fit only indirectly in the model through the relationship between notifications and prevalence survey data.

The paradigm of intermittent nationwide prevalence surveys combined with incomplete case notifications data arises in the context of other infectious diseases. In sub-Saharan Africa, Malaria Indicator Surveys (MIS) and AIDS Indicator Surveys (AIS) are often considered to be the gold standard for surveillance, but small sample sizes and gaps in temporal coverage limit their use for routine surveillance. Previous spatial modelling studies have combined these surveys with routine surveillance data, using an estimated cross-walk between data types;[125,165] these models could be made more flexible by adapting the synthesis approach described in this chapter.

3.4.2 Limitations

While this model generates relatively precise estimates of TB prevalence and case notification completeness across districts in Uganda, it relies on an assumption that may not hold over long time periods. Because WHO and IHME modelling studies agree that nationwide TB incidence in Uganda fell by less than 5% between 2015 and 2019,[156,166] the model presented in this chapter assumes that subnational variation in TB prevalence also remains constant over this time period. Conversely, case notification reporting completeness was modelled as a space-time surface to capture documented increases in case notifications over the same time period.[135,167]. While such an assumption simplifies the comparisons between prevalence and incidence data sources, it cannot hold if the End TB goals are to be met. In settings where TB burden is decreasing, the underlying TB prevalence surface

should vary in both space and time, with district-specific censuses directly grounding prevalence estimates in the years following a prevalence survey.

To relate data sources measuring tuberculosis incidence and prevalence, I roughly approximated variation in TB duration by district as a function of case detection rates. This approximation required strong assumptions about rates of HIV-TB co-infection, drug resistance, and TB treatment and their effects on duration by district in Uganda. While I conducted a sensitivity analysis showing that the effect of relaxing any individual assumption would have a relatively minor effect on estimated duration, further research is needed to better understand subnational variation in HIV-TB and TB treatment across Uganda. The IHME and WHO use differing approaches to calculate disease duration;[156,158] while this chapter extends the WHO approach due to the greater availability of relevant inputs at the district level, additional research is needed to compare and validate these approaches in a variety of settings.

3.4.3 Conclusions

In this chapter, I demonstrate that subnational TB case notifications can provide meaningful information about the spatial distribution of TB burden within a country. This study extends the neonatal mortality model developed in Chapter 2 to account for health surveillance data that is expected to be incomplete, with incompleteness varying in both space and time. Whether the measured outcome is all-cause mortality or infectious disease prevalence, the principle of indirect data synthesis offers programmatically useful insights about variation in both underlying risk and data quality.

3.5 References

7. World Health Organization. *Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies* tech. rep. 1 (2010), 1–92. http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.35.021103.105711.
25. Uplekar, M. *et al.* Mandatory tuberculosis case notification in high tuberculosis-incidence countries: Policy and practice. *European Respiratory Journal* **48,** 1571–1581. http://dx.doi.org/10.1183/13993003.00956-2016 (2016).
26. Rood, E. *et al.* A spatial analysis framework to monitor and accelerate progress towards SDG 3 to end TB in Bangladesh. *ISPRS International Journal of Geo-Information* **8,** 1–11 (2019).
43. Riebler, A. *et al.* An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research* **25,** 1145–1165. arXiv: 1601.01180 (2016).
55. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* **4,** 2–5 (2017).

66. Thomson, D. R. *et al.* Extending Data for Urban Health Decision-Making: a Menu of New and Potential Neighborhood-Level Health Determinants Datasets in LMICs. *Journal of Urban Health* **96,** 514–536 (2019).
99. Hook, E. B. & Regal, R. R. Capture-recapture methods in epidemiology: Methods and limitations. *Epidemiologic Reviews* **17,** 243–264 (1995).
118. Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H. & Bell, B. M. TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* **70.** arXiv: 1509.00660 (2016).
119. Thorson, J. T. & Kristensen, K. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research* **175,** 66–74 (2016).
120. R Core Team. *R: A Language and Environment for Statistical Computing* Vienna, Austria, 2018. https://www.r-project.org/.
125. Dwyer-Lindgren, L. *et al.* Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature* **570,** 189–193. http://www.nature.com/articles/s41586-019-1200-9 (June 2019).
126. Dehnavieh, R. *et al.* The District Health Information System (DHIS2): A literature review and meta-synthesis of its strengths and operational challenges based on the experiences of 11 countries. *Health information management : journal of the Health Information Management Association of Australia* **48,** 62–75 (2019).
127. Onozaki, I. & Raviglione, M. Stopping tuberculosis in the 21st century: goals and strategies. *Respirology* **15,** 32–43 (2010).
128. World Health Organization. *The End TB Strategy* tech. rep. (World Health Organization, Geneva, 2015). https://www.who.int/teams/global-tuberculosis-programme/the-end-tb-strategy.
129. World Health Organization. *Global Tuberculosis Report 2020* tech. rep. (World Health Organization, Geneva, 2020), 1–208. https://www.who.int/publications/i/item/9789240013131.
130. World Health Organization. *Implementing the End TB Strategy* tech. rep. (World Health Organization, Geneva, 2015). https://www.who.int/teams/global-tuberculosis-programme/the-end-tb-strategy.
131. World Health Organization. *Framework and standards for country health information systems (2E)* tech. rep. (World Health Organization, Geneva, 2012), 1–63. http://www.who.int/healthmetrics/documents/hmn_framework200803.pdf.
132. World Health Organization. *Definitions and reporting framework for tuberculosis: 2013 revision* tech. rep. (World Health Organization, Geneva, 2020), 1–40. https://apps.who.int/iris/handle/10665/79199.
133. Kyu, H. H. *et al.* Global, regional, and national burden of tuberculosis, 1990-2016: results from the Global Burden of Diseases, Injuries, and Risk Factors 2016 Study. *The Lancet Infectious Diseases* **18,** 1329–1349. https://linkinghub.elsevier.com/retrieve/pii/S147330991830625X (Dec. 2018).
134. Uganda National Tuberculosis and Leprosy Programme. *National Tuberculosis and Leprosy Control Programme: revised national strategic plan 2015/16-2019/20* tech. rep. (Uganda Ministry of Health, Kampala, 2017), 1–52. file:///C:/Users/nyomb/Downloads/Revised-NTLP-Stragetic-Plan-2015-2020_25th-Jul2017-final.pdf.
135. Uganda National Tuberculosis and Leprosy Programme. *Uganda National TB and Leprosy Program: July 2019-June 2020 report* tech. rep. October 2020 (Uganda Ministry of Health, Kampala, 2020), 1–72.
136. Glaziou, P. & Floyd, K. *Latest developments in WHO estimates of TB disease burden* tech. rep. (WHO Global Task Force on TB Impact Measurement, Glion, 2018), 1–13.
137. Glaziou, P., Van Der Werf, M. J., Onozaki, I. & Dye, C. Tuberculosis prevalence surveys: Rationale and cost. *International Journal of Tuberculosis and Lung Disease* **12,** 1003–1008 (2008).

138. Karamagi, E. *et al.* Improving TB case notification in northern Uganda: evidence of a quality improvement-guided active case finding intervention. *BMC Health Services Research* **18,** 954. https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-018-3786-2 (Dec. 2018).
139. Shaweno, D. *et al.* Methods used in the spatial analysis of tuberculosis epidemiology: A systematic review. *BMC Medicine* **16,** 1–18 (2018).
140. Van Gurp, M. *et al.* Finding gaps in TB notifications: spatial analysis of geographical patterns of TB notifications, associations with TB program efforts and social determinants of TB risk in Bangladesh, Nepal and Pakistan. *BMC Infectious Diseases* **20,** 1–14 (2020).
141. Sudhakar, M. & Admassu, B. Determinants of help seeking and treatment seeking behavior of Tuberculosis patients - Gender perspective: A systematic review. *JBI Database of Systematic Reviews and Implementation Reports* **11,** 199–271 (2013).
142. Abebe, G. *et al.* Knowledge, health seeking behavior and perceived stigma towards tuberculosis among tuberculosis suspects in a rural community in Southwest Ethiopia. *PLoS ONE* **5,** 1–7 (2010).
143. Ahsan, G. *et al.* Gender difference in treatment seeking behaviors of tuberculosis cases in rural communities of Bangladesh. *Southeast Asian Journal of Tropical Medicine and Public Health* **35,** 126–135 (2004).
144. Kirirabwa, N. S. *et al.* A four-year trend in pulmonary bacteriologically confirmed tuberculosis case detection in Kampala-Uganda. *BMC Pulmonary Medicine* **19,** 1–8 (2019).
145. Uganda Ministry of Health. *The Uganda National Tuberculosis Prevalence Survey, 2014-2015: survey report* tech. rep. (Uganda Ministry of Health, Kampala, 2015), 1–162. https://www.health.go.ug/cause/the-uganda-national-tuberculosis-prevalence-survey-2014-2015-survey-report/.
146. Uganda National Tuberculosis and Leprosy Programme. *Manual for management and control of tuberculosis and leprosy (3E)* tech. rep. (Uganda Ministry of Health, Kampala, 2017), 1–162.
147. Uganda National Tuberculosis and Leprosy Programme. *National Tuberculosis and Leprosy Division: July 2017-June 2018 report* tech. rep. (Uganda Ministry of Health, Kampala, 2018), 1–49.
148. Uganda National Tuberculosis and Leprosy Programme. *National strategic plan for tuberculosis and leprosy control: 2020/21-2024/25* tech. rep. November 2017 (Uganda Ministry of Health, Kampala, 2020), 1–106.
149. Weiss, D. J. *et al.* Global maps of travel time to healthcare facilities. *Nature Medicine* **26,** 1835–1838 (2020).
150. Freeman, J. & Hutchison, G. B. Prevalence, incidence, and duration. *American Journal of Epidemiology* **112,** 707–723. https://academic.oup.com/aje/article/78756/PREVALENCE, (Nov. 1980).
151. World Health Organization. *Guidelines for treatment of drug-susceptible tuberculosis and patient care* 1–56. https://apps.who.int/iris/bitstream/handle/10665/255052/9789241550000-eng.pdf (World Health Organization, Geneva, 2017).
152. Tiemersma, E. W., van der Werf, M. J., Borgdorff, M. W., Williams, B. G. & Nagelkerke, N. J. Natural history of tuberculosis: Duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: A systematic review. *PLoS ONE* **6** (2011).
153. Payne, B. & Bellamy, R. Managing tuberculosis in people with HIV. *BMJ Clinical Evidence* (2007).
154. World Health Organization. *WHO consolidated guidelines on tuberculosis. Module 4: drug-resistant tuberculosis treatment* tech. rep. (World Health Organization, Geneva, 2020), 1–98. https://www.who.int/publications/i/item/9789240007048.
155. Falzon, D. *et al.* WHO guidelines for the programmatic management of drug-resistant tuberculosis: 2011 update. *European Respiratory Journal* **38,** 516–528 (2011).

156. Ledesma, J. R. *et al.* Global, regional, and national sex differences in the global burden of tuberculosis by HIV status, 1990-2019: results from the Global Burden of Disease Study 2019. *The Lancet Infectious Diseases* **3099,** 1–20 (2021).
157. Fullman, N. *et al.* Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. *The Lancet* **391,** 2236–2271. https://linkinghub.elsevier.com/retrieve/pii/S0140673618309942 (June 2018).
158. Glaziou, P., Dodd, P. J., Dean, A. S. & Floyd, K. *Methods used by WHO to estimate the global burden of TB disease: global tuberculosis report 2020* tech. rep. October (World Health Organization, Geneva, 2020). https://cdn.who.int/media/docs/default-source/hq-tuberculosis/global-tuberculosis-report-2020/tb2020_technical_appendix_20201014.pdf.
159. Buregyeya, E., Criel, B., Nuwaha, F. & Colebunders, R. Delays in diagnosis and treatment of pulmonary tuberculosis in Wakiso and Mukono districts, Uganda. *BMC Public Health* **14,** 1–10 (2014).
160. Glaziou, P., Dodd, P. J., Dean, A. S. & Floyd, K. *Methods used by WHO to estimate the global burden of TB disease: global tuberculosis report 2019* tech. rep. October (World Health Organization, Geneva, 2019). https://www.who.int/docs/default-source/documents/tuberculosis/technical-appendix-global-tb-report-2019methods-used-by-who-to-estimate-the-global-burden-of-tb-disease.pdf?sfvrsn=e9c1d7b0_2.
161. Glaziou, P., Dodd, P., Zignol, M., Sismanidis, C. & Floyd, K. *Methods used by WHO to estimate the global burden of TB disease: global tuberculosis report 2018* tech. rep. September (World Health Organization, Geneva, 2018), 1–36. https://www.who.int/tb/publications/global_report/gtbr2018_online_technical_appendix_global_disease_burden_estimation.pdf.
162. Togun, T., Kampmann, B., Stoker, N. G. & Lipman, M. Anticipating the impact of the COVID-19 pandemic on TB patients and TB control programmes. *Annals of Clinical Microbiology and Antimicrobials* **19,** 1–6. https://doi.org/10.1186/s12941-020-00363-1 (2020).
163. Shaweno, D., Trauer, J. M., Denholm, J. T. & McBryde, E. S. A novel Bayesian geospatial method for estimating tuberculosis incidence reveals many missed TB cases in Ethiopia. *BMC Infectious Diseases* **17,** 1–8 (2017).
164. Van Hest, R., Grant, A. & Abubakar, I. Quality assessment of capture-recapture studies in resource-limited countries. *Tropical Medicine and International Health* **16,** 1019–1041 (2011).
165. Lucas, T. C. D. *et al.* Model ensembles with different response variables for base and meta models: malaria disaggregation regression combining prevalence and incidence data. *BioRxiv,* 2–7. arXiv: 548719 (2019).
166. World Health Organization. *Global Tuberculosis Report 2019* 1–283 (World Health Organization, Geneva, 2019).
167. Uganda National Tuberculosis and Leprosy Programme. *National Tuberculosis and Leprosy Programme: annual report for 2017* tech. rep. (Uganda Ministry of Health, Kampala, 2017), 1–72.

Chapter 4

A space-time-age model for subnational child mortality estimation in India

4.1 Introduction

In 2017, the Indian Ministry of Health and Family Welfare released the National Health Policy (NHP), the first strategic plan in 15 years to clearly lay out the Indian government's priorities and targets related to health.[168] This document's importance extends even beyond its goal-setting function for the next ten years of health policy in India: it also represents an attempt by India's national governing party to fulfil campaign promises centred around universal health coverage by the year 2025.[169] To carry out these core principles, Indian health policy-makers first require local and focal insights about health burden nationwide. A situation analysis released alongside the NHP emphasises how measuring inequality in health outcomes is a first step towards achieving universal good health:

> We also need to keep in mind that high degree of inequity in health outcomes and access to health care services exists in India. This is evidenced by indicators disaggregated for vulnerable groups and between and within States. Identifying the deprived areas/vulnerable population groups (including special groups) through disaggregated data is a first step to address the existing inequities in health outcomes between and within States in India.[170]

The NHP places particular emphasis on the health and survival of one vulnerable sub-population: children under five years of age. The NHP's concrete goals for the coming decade include commitments

to reduce the neonatal mortality rate (NMR) to fewer than 16 deaths per 1,000 live births by 2025; to reduce the infant mortality rate (IMR) to fewer than 28 deaths per 1,000 live births by 2018; and to reduce the under-5 mortality rate (U5MR) to fewer than 23 deaths per 1,000 live births by 2025. To meet these targets universally and equitably, policy-makers need information about disparities in the health of children under 1 month, 1 year, and 5 years of age, respectively, across the country. However, none of India's three primary mortality surveillance systems have traditionally offered spatially-resolved information about child survival across the country.

In this chapter, I describe a model to estimate neonatal, infant, and child mortality at the district level based on household survey data, the only data source for child mortality for which location data is published below the state level. Based on results from this model, I demonstrate how district-level estimates of mortality can reveal policy-relevant information hidden by state-level results. I then compare these estimates to the data available from the other two mortality surveillance systems to demonstrate how subnational reporting across all three systems is needed to support equitable child health outcomes across India.

4.1.1 Mortality across Indian states: progress, transition, and inequality

The NHP's emphases on child welfare and equity reflect India's past experience in child health provision. While India halved the under-5 mortality rate from 2000 to 2017, from 80 deaths per 1,000 live births to less than 40,[89] state-level estimates of child welfare indicate that some parts of the country are being left behind amid general progress towards improved child health. The NHP reports that as of 2013, the states of Madhya Pradesh and Assam both experienced infant mortality rates of 54/1,000, or more than 1 in 20, more than five times higher than the IMR observed across the country in Goa (9/1,000) or Manipur (10/1,000).[168]

Previous research has contextualised these striking disparities as part of a discontinuity in the standard epidemiological and demographic transitions. In some regions of India, such as Kerala state in the south, life expectancy has increased dramatically, leading to an increase in non-communicable disease burden and corresponding strain on the health system related to elder care;[171] meanwhile, particularly in rural settings and among marginalised groups, infectious diseases and maternal and

child disorders are more deleterious to health.[172] These competing needs complicate national health policymaking, leading some researchers to contend that the Indian epidemiological context must be understood as many "nations within a nation."[172]

Facing a rapidly-changing health context marked by fundamental differences in health needs, Indian policy-makers need health information systems that accurately reflect local variation in health status in order to deploy appropriate interventions within communities facing the greatest disease burden. The World Health Organization recognises health information systems as a necessary building block for any successful health system, following the principle that public health relies on evidence to function.[7,81] This chapter evaluates the capacity of India's health information systems to track progress towards the National Health Policy's child mortality targets.

4.1.2 Tracking local variation in mortality across India

According to the 1969 Registration of Births and Deaths Act passed by the Parliament of India, birth and death registration are owed to every Indian citizen.[173] This Act, passed alongside the Census and Statistics Act for India,[174] recognises that civil registration and vital statistics (CRVS) systems are a key both to efficient public administration as well as to maintaining the human rights to documented citizenship and social security.[1] Given the logistical challenges associated with registering all births and deaths across India, the Indian government historically developed and maintained a number of overlapping health information systems to meet these needs. In this section, I discuss three information systems that are crucial to understanding child mortality across India: the Sample Registration System, the Civil Registration System, and a system of regular household surveys focused on maternal and child health. Table 4.1 summarizes key features of these data sources as well as differences between them.

Following the passage of the 1969 Registration of Births and Deaths Act, the Indian government implemented the Sample Registration System (SRS) in 1970, then expanded it to a nationwide system in 1976-77.[87] As of 2017, the SRS covered a population of approximately 7.9 million people (approximately 0.6% of the population of India) across 3,892 urban and 4,961 rural sampling units.[175]

Table 4.1: Key features of mortality data sources in India.

Source	Population sampled	Temporal resolution	Spatial resolution
Civil registration system (CRS)	Nationwide (incomplete)	Annual	State urban/rural
Sample registration system (SRS)	7.9 million people	Annual	State urban/rural, natural divisions (infant mortality only)
Household surveys: NFHS, DLHS, AHS	Between 100,000 and 4.3 million households	Every 1-5 years	Survey cluster GPS or district

In rural areas, a sampling units correspond to individual villages or subsets of villages with a population totalling less than 2,000 people each; in urban areas, sampling units correspond to Census Enumeration Blocks.[175] The system is based on a two-tiered sampling strategy: when a new sampling unit is added, a baseline census of the area is taken. Vital events are then registered continuously, with a survey of the sampling area conducted every six months for an independent count and demographic update.[176] Indicators of fertility, population, and age-specific mortality are then generated at the national and state levels, disaggregated by urban and rural status. Infant mortality is also reported at the sub-state natural division level for select large states.[175]

Conversely, the Indian Civil Registration System (CRS) aims for universal registration of all births and deaths across India. Complete coverage of the CRS is the pathway by which all Indian citizens can access the legal and civil protections afforded by birth and death registration, as guaranteed by the 1969 Registration of Births and Deaths Act.[173,177] Registration in the CRS relies on individual reporting of vital events, which falls to household heads in cases where births and deaths occur in the home, or to facility heads for vital events that occur in institutional settings. This self-report principle can be challenging given the large number of births and deaths that occur at home or in private facilities.[178] To partially address these challenges, the Indian government has targeted improvements in the coverage of CRS birth and death registration based in part on digital registration; however, the estimated completeness of the CRS still remains low compared to the SRS.[35]

Three major survey series conducted by the Indian government also capture aspects of maternal and child mortality nationwide. These series are the District-Level Household Surveys (DLHS);

the National Family Health Surveys (NFHS), conducted in partnership with the international Demographic and Health Surveys program; and the Annual Health Surveys (AHS). Of these, the AHS is the largest, with over 4.3 million households captured in its 2013 sample.[34] While these surveys use differing sampling strategies, survey different household members, and collect information on diverse topics, all three capture retrospective information on child mortality by requesting birth histories from women of reproductive age.[34] While both the DLHS and NFHS surveys are conducted in five-year intervals, a two-year gap between them offers temporal survey coverage that is unparalleled in other countries. However, survey-specific differences in questions related to adult mortality and non-communicable disease burden limit their utility for understanding the course of the epidemiological transition across India.[171]

Among these data sources, the SRS is considered to be the gold standard for estimating fertility, births, and deaths across the country due to its relatively large sample size and representative sampling design.[176] Past studies have used the SRS as the baseline against which the completeness of CRS mortality reporting is estimated.[35] Against this standard, the coverage of the CRS has been increasing: using SRS as the baseline, a study estimated that CRS coverage increased from 55% to 77% nationally, with completeness approaching 100% in nine states.[35] However, the Registrar General of India, which maintains the SRS, has only infrequently published estimates of the source's completeness. An independent investigation of SRS completeness using the Brass Generalised Growth Balance method found that at the state level, the completeness of death registration in the SRS between 1981-1990 varied between 81% and 100% for males and 74% to 95% for females across all age groups.[87] A more recent study of SRS completeness using the Preston and Coale method estimated that SRS completeness varied between 77% and 99% between 1990 and 2007 without a clear trend showing improved completeness over time. Notably, this study also found that SRS in the south-eastern Indian state of Andhra Pradesh captured only 58% of deaths in 2007, the final year of estimation.[176]

As tools for identifying health disparities nationwide, the SRS and CRS are limited in the spatial and sub-population data they report. Neither of these two sources reports mortality at the sub-state level, except for the infant mortality rate, which the SRS has begun to report by district grouping in recent years.[176] The survey systems report township-level geographic information

associated with each sample cluster; however, these surveys capture only retrospective information about child mortality that is subject to possible response biases and do not include information about causes of death that are provided by CRS records.[34]

4.2 Methods

In this chapter, I develop a space-time modelling technique to estimate spatial variation in child mortality using data from the three major household survey series. This model takes advantage of the precise spatial information offered by the household surveys while attempting to correct for possible recall and omission biases associated with retrospective survey data. I then compare these modelled estimates of child mortality to the SRS and CRS data sources to explore possible cross-source differences in mortality estimates. This chapter offers the first analysis, to my knowledge, to systematically examine the relationship between the structure of birth history data and its limitations for space-time modelling. This chapter also offers a critical appraisal of the subnational mortality data published by the SRS, which no previous study has examined.

Figure 4.1, below, summarizes the analyses process described in this chapter. While I initially prepared this model as part of a collaborative project estimating causes of child mortality across India,[179] I have since personally produced all survey-based estimates of child mortality reported in this chapter. The comparison of estimates between SRS and survey-based data presented below are novel to this thesis.

4.2.1 Geo-locating complete birth history data

The NFHS, DLHS, and AHS all collect complete birth history from women of reproductive age. Complete birth history data provide the month and year of birth and death for each child of an interviewed woman. Complete birth histories from individual-level survey data were extracted between 2000 and 2017. These birth histories were then reshaped to reflect the number of children entering distinct age groups between birth and age five, as well as the number who died within each age group, by survey cluster and retrospective year.[2,180] Survey clusters were matched to precise spatial identifiers such as

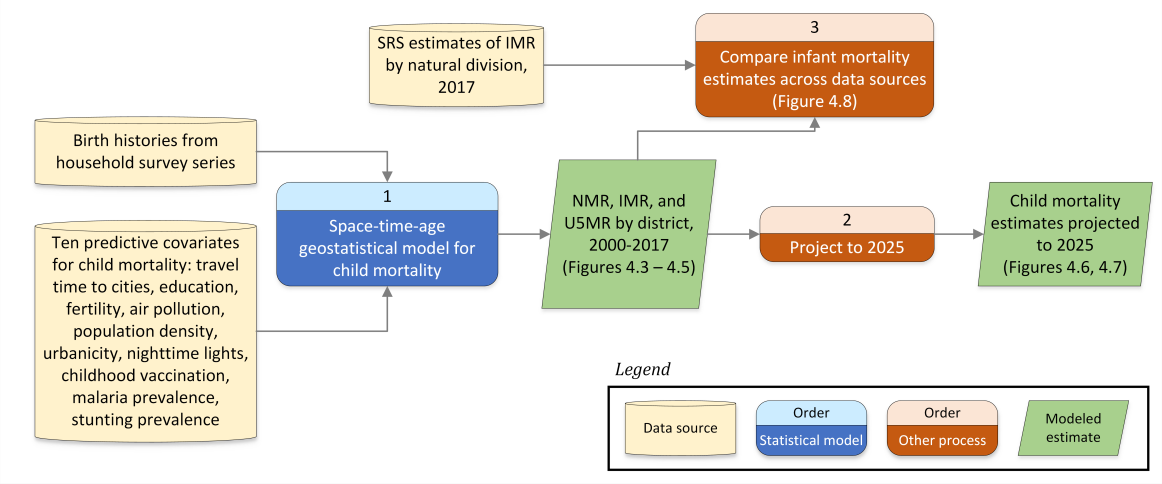


Figure 4.1: Summary of data sources, analyses, and results described in Chapter 4.

GPS points, precise township names, or districts. Across the three survey series, 3.3 million georeferenced birth histories were incorporated into the spatial model. Figure 4.2 shows the distribution of retrospective birth histories by year of birth across the study time period.

4.2.2 Space-time-age mortality estimation model

To synthesise information across various sources, and to make consistent estimates across space and time, I fitted a discrete hazards geostatistical model to the data. Age groups were represented in seven mutually exclusive bins (0, 1–5, 6–11, 12–23, 24–35, 36–47 and 48–59 months), each with a baseline mortality probability that was assumed to be constant nationwide. The model explicitly accounted for variation across age bin, year and space through inclusion of both fixed and random effects. Indicator variables for each age bin were included to form a discrete baseline mortality hazard function, representing the risk of mortality in discrete bins from birth to 59 months of age with covariates set at their means. Baseline hazard functions were allowed to vary in space and time in response to changing covariate values, as well as in response to linear effect on year. These estimated fixed effects were then applied to the gridded surface of covariate values to make predictions across the entire study geography. A latent process effect was also included to account for remaining correlation across age, time and physical space after accounting for fixed effects and source-specific

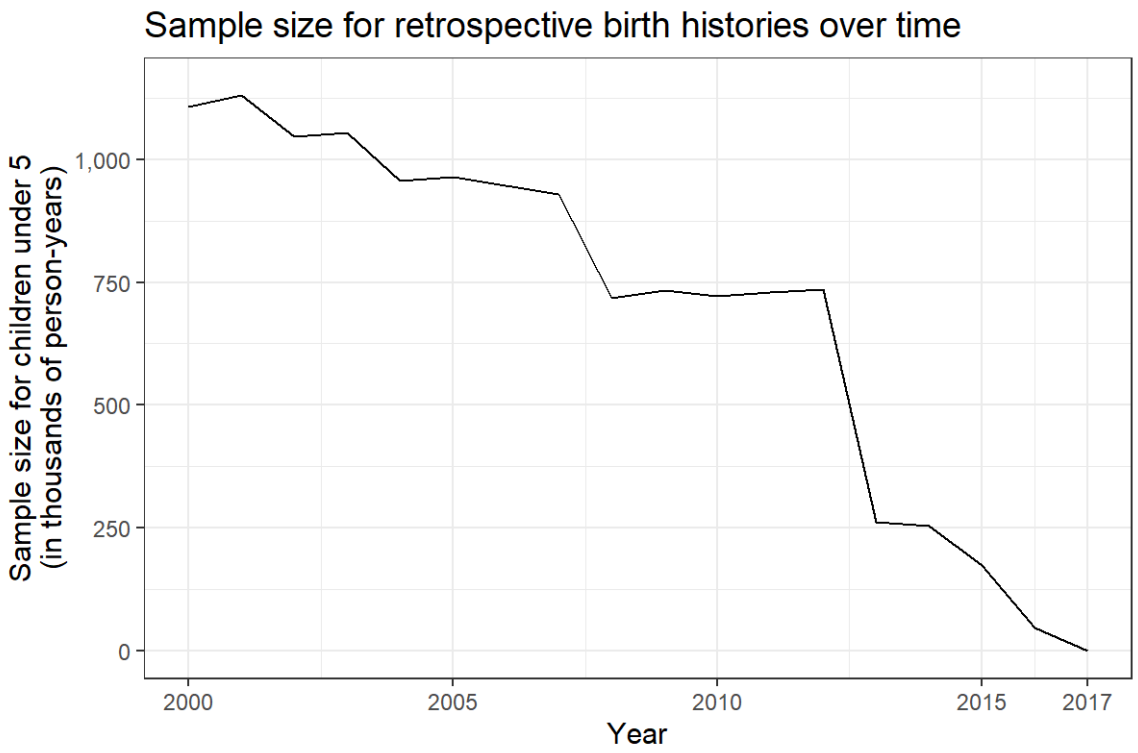


Figure 4.2: Sample size for under-5 mortality modeling based on survey data between 2000 and 2017 across India. The sample size for each year is approximated here as the number of individuals under age 5 entering each year according to retrospective complete birth history data. At the time of analysis, no complete birth history data was available after 2016, the year of the most recent included household survey.

biases. As such, estimates at a specific age, time or place benefited from drawing predictive strength from data points nearby in all of these dimensions.

All data were prepared such that we counted or estimated the number of children entering into (n) and dying within (Y) each period–age bin from each GPS-point location (s) in each survey (k). The number of deaths for children in age band (a) in year (t) at location (s) was assumed to follow a binomial distribution:

$$Y_{a,s,t} \sim Binomial(n_{a,s,t}, P_{a,s,t})$$

where $P_{a,s,t}$ is the probability of death in age bin (a), conditional on survival to that age bin for a particular space–time location. Using a generalised linear regression modelling framework, a logit link function is used to relate P to a linear combination of effects:

$$logit(P_{a,s,t}) = \beta_0 + \beta_1(a) + \beta_2 X_{s,t} + \beta_3 t + \nu_k + Z_{a,s,t}$$

The first term, β_0, is a global intercept, representing the mean for the first age band when all covariates are equal to zero, whereas $\beta_1(a)$ are fixed effects for each age band, representing the mean overall hazard deviation for each age band from the intercept, when all other covariates are equal to zero. β_2 are the linear fixed effects of geospatial covariates ($X_{s,t}$), while β_3 is a linear temporal effect by year. The term $\nu_k \sim Normal(0, \sigma_k^2)$ is a survey-level random effect used to account for systematic variation or biases across data sources: this term was fit to the deviation between overlapping survey estimates, and was excluded from the final model predictions as a bias term.

Finally, the term $Z_{a,s,t} \sim GP(0, K)$ is a correlated four-dimensional separable Gaussian process, accounting for structured residual correlation across the indices of space, time, and age that are not accounted for by any of the model's other effects. This model term can identify and leverage space-time autocorrelation in the residuals. Model separability assumes that the covariance function between any two observations across the dimensions of age, space, and time is equal to the outer product of the component covariance functions across each dimension. In other words, $K = \Sigma_a \otimes \Sigma_t \otimes \Sigma_s$ where "$\otimes$" is the Kronecker product or generalized outer product. Space-time separable models allow for computationally efficient estimation and inference,[181] and have been applied under a wide array of disease mapping contexts.[60,108,182]

The age covariance function, Σ_a, and the time covariance function, Σ_t, are represented as autoregressive processes with order 1: this means that on a single dimension, the value taken by a time or age grouping is dependent only on the values in the immediately adjacent groupings.[183] The spatial covariance function, Σ_s, is modelled using the popular approach described by Lindgren and Rue (2011) and implemented in the INLA software, where a continuous two-dimensional Gaussian process in space is approximated across a triangular mesh.[45,51] In the INLA approach, spatial covariance is represented by a Matérn covariance function.[45]

This model draws from previous studies investigating variation in child mortality based on survey data, particularly Burstein *et al.*'s study in low- and middle-income countries.[2,108]

4.2.3 Mortality forecasting

After estimating NMR, IMR, and U5MR between 2000 and 2017, I calculated a weighted annualised rate of change for all grid cells and draws, giving greater weight to more recent annual rates of change. I then applied this annual rate to each indicator cell-draw in the final year of mortality estimates, in 2017, to project estimates for all three indicators through 2025 while preserving uncertainty. Previous spatial analyses have established this method for projecting spatial estimates into the future.[62] I then applied a population-weighted aggregation to calculate estimated mortality at the district and state levels based on gridded mortality projections.

4.2.4 Comparisons to SRS data at the most detailed spatial level available

In recent years, the Indian Sample Registration System has begun to report estimates for infant mortality rates across 68 sub-state "natural divisions" in the larger states of India.[175] I used population-weighted aggregation to estimate infant mortality by natural division, with uncertainty, in 2017. I then compared these estimates based on survey data with point estimates provided by the SRS at the same spatial resolution in 2017. Rather than using either source as a "gold standard", I review significant differences between survey and SRS data as a starting point for future research and improvement.

4.3 Results

A small-area analysis of mortality reveals focal areas and within-state disparities in child mortality that would be obscured in a state-level analysis. While the neonatal, infant, and under-5 mortality rates dropped in almost all districts of India between 2000 and 2017, stark inequalities in child survival remain across the country. Simple projections to 2025 reveal that while absolute differences in mortality across district will generally converge in future years if current trends persist, many districts are not on track to meet the 2025 child survival goals set by the National Health Policy.[168]

A comparison of infant mortality estimates between the Sample Registration System and synthesised survey data by Indian natural division reveals that both sources identify the same focal regions for infant mortality nationwide. The Sample Registration System estimates significantly lower infant

mortality rates than survey-based estimates in 11 of 68 natural divisions, primarily in states where past analyses have identified deficiencies in SRS completeness. However, the larger sample size of the SRS facilitates precise estimates of infant mortality in each natural division, complementing a weakness of survey-based estimates at the end of the study time period.

4.3.1 Progress and local disparities in child survival

A synthesis of survey data indicates that at the national level, the under-5 mortality rate across India fell to 42.4 (37.2-49.0) per 1,000 live births by 2017, from 83.1 (77.6-88.6) in 2000. Across all states in the country, child mortality varies almost 6-fold, from 10.4 (7.5-14.2) in Kerala to 59.7 (50.4-71.8) in Uttar Pradesh. This inequality, while stark, pales in comparison to differences at the district level. Figure 4.3 shows the distribution of under-5 mortality across India, which ranged from a low of 16.4 (13.2-20.2, in Thrissur district in Kerala) to a high of 163.3 (147.4-179.6, in Panna district of Madhya Pradesh) in 2000 and from 8.4 (5.6-12.1, again in Thrissur district) to 87.9 (72.2-106.4, in Budaun district in Uttar Pradesh) in 2017. In 2017, under-5 mortality in 86 of the 723 districts nationwide exceeded the highest state-level U5MR; of these, more than half were found outside of Uttar Pradesh state, including 14 districts in Assam and 14 in Madhya Pradesh.

In many ways, estimates of neonatal and infant mortality rates from 2000-2017, also shown in Figure 4.3, reflect a similar pattern of unequal progress. Between 2000 and 2017, the neonatal mortality rate declined nationally from 37.8 (35.2-40.3) to 23.3 (20.4-26.9), while the infant mortality rate declined from 62.7 (58.5-66.9) to 36.0 (31.6-41.5). However, at the district level, neonatal mortality ranged 8-fold, from 5.8 (3.9-8.2) in Ernakulum, Kerala state to 46.2 (37.5-56.6) in Budaun, Uttar Pradesh. District-level rates of infant mortality ranged 10-fold, from 7.3 (4.9-10.6) in Thrissur district, Kerala state to 75.9 (62.1-92.3) in Budaun, Uttar Pradesh. As with under-5 mortality, districts with the highest rates of infant mortality were generally concentrated in the states of Uttar Pradesh, Madhya Pradesh, Chhattisgarh, and Rajasthan in 2017.

The annual rate of decline in under-5 mortality is an important summary indicator showing the net effect of child survival interventions across all stages of early childhood. While the under-5 mortality rate declined by 49 percent from 2000 to 2017, equivalent to a 3.8% annualised rate of decline,

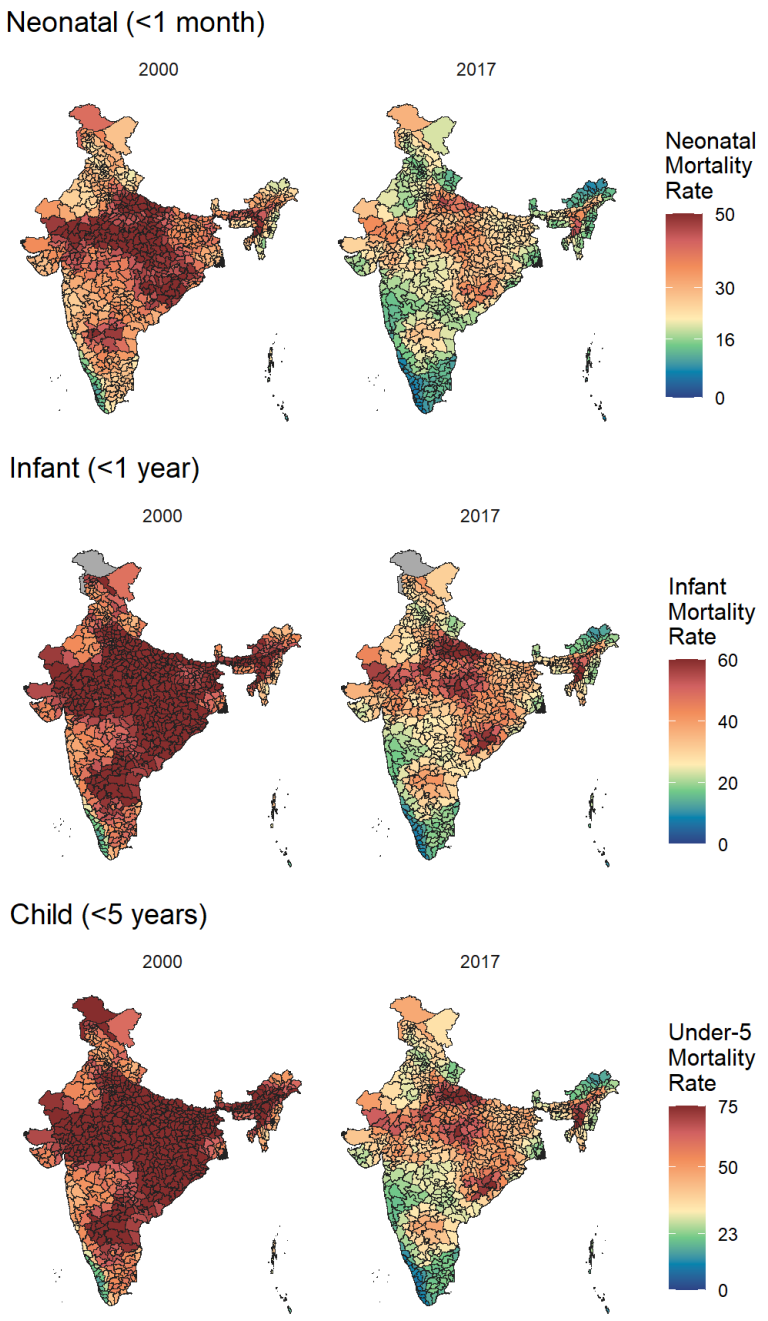


Figure 4.3: NMR, IMR, and U5MR per 1,000 live births in 2000 and 2017, estimated from survey data. The Indian National Health Plan 2017 calls for reducing neonatal mortality to less than 16 deaths per 1,000 live births, and under-5 mortality to less than 23 deaths per 1,000 live births, nationwide by 2025. The color scale for mortality varies across age groups, with green centered around the NHP mortality targets for each age group.

the speed of improvements in child survival varied considerably across the country. Figure 4.4 shows the annualised rate of decline in child mortality by district from 2000 to 2017. Across 103 districts concentrated along the south-eastern coast, Telangana state, and Arunachal Pradesh, annual decline in child mortality exceeded 5% per year, equivalent to a nearly 60% reduction in U5M across the study period; meanwhile, 23 districts concentrated primarily in Rajasthan and Himachal Pradesh in the north-west and Mizoram and eastern Assam in the north-east experienced less than 2% of annual decline in the child mortality rate. When observing the pattern of mortality decline across Indian districts, the space-time pattern in the north-east of the country is of particular interest: although Arunachal Pradesh and western Assam have experienced rapid declines in mortality, child mortality remains stubbornly high in a region covering eastern Assam, Mizoram, and Tripura. A number of factors could be influencing this unusual time trend across North-eastern India: the region's geographical separation from the rest of the country, heterogeneous ethnic composition, proximity to Myanmar and Bangladesh, and high prevalence of drug-resistant malaria could all play a role in shaping under-5 mortality.[184,185] Further investigation is needed to understand why gains in child survival observed in Arunachal Pradesh and western Assam have not been matched in other North-east Indian districts.

The differential improvement in child survival across the eastern and western halves of Assam is one of many insights revealed by a district-level analysis. Figure 4.3 also demonstrates spatial trends in neonatal, infant, and child mortality that would be obscured by a state-level analysis. The states of Karnataka, Maharashtra, and Gujarat experience significantly lower NMR, IMR, and U5MR in the coastal east than in the inland west. In the state of Rajasthan, U5MR is approaching the NHP target of 23 in 6 districts near Jaipur city, while 10 districts in the south-west of the state still suffer from a U5MR exceeding 60. The ratio between the under-5 mortality rates of any two districts in the same state reaches 3.6 (in Karnataka state, from 11.3 (7.8-15.7) in Dakshin Kannad district to 41.2 (30.4-54.2) in Koppal district). A similar ratio of 3.7 is observed across neonatal mortality rates in the districts of Karnataka, suggesting that many of the inequalities observed in child survival arise from disparities in health burden within the first month of life.

This space-time model also offers insight into changing within-state inequalities over time. Figure 4.5 shows the spread of NMR, IMR, and U5MR across districts within each Indian state and union

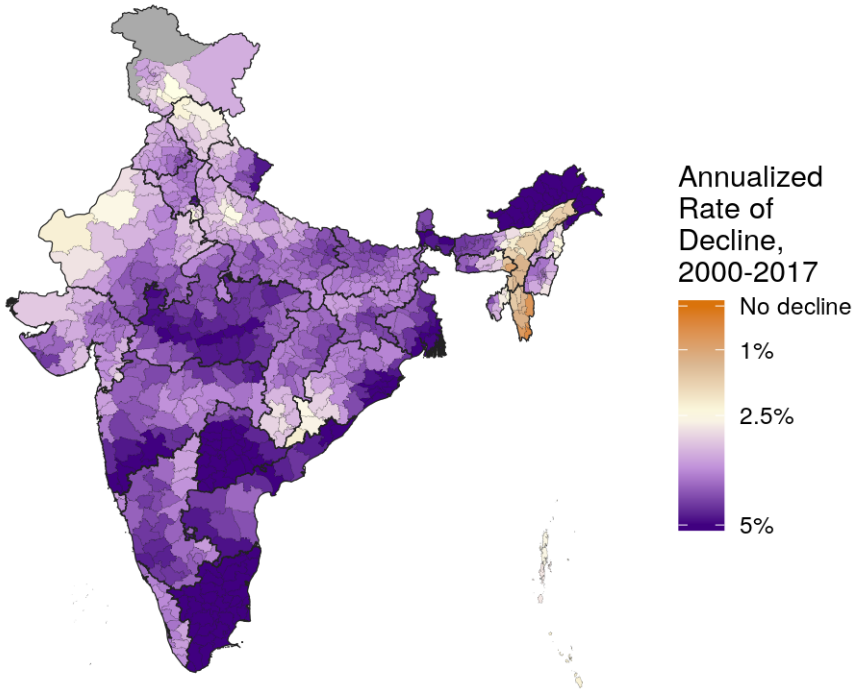


Figure 4.4: Annualised rate of decline (ARoD) in U5MR between 2000 and 2017. An ARoD of 5% is equivalent to a cumulative decline of 58% over 17 years, while an ARoD of 1% is equivalent to a cumulative decline of 16% over 17 years. ARoD rankings by district were stable across age-specific indicators.

territory in 2000 (grey) and 2017 (blue). Each dot represents a district: the lower bound of each vertical spread represents the district with the lowest U5MR in each state and year, while the upper bound of each vertical spread represents the district with the highest U5MR in the same district and year. The difference in the U5MR between the highest and lowest districts can be interpreted as one measure of absolute inequality in U5M within each state: by this metric, absolute inequality in U5M declined in almost all Indian states between 2000 and 2017. Figure 4.5 reveals the diverse relationship between declining state-level U5MR and increasing equity over the study time period. For example, in Uttar Pradesh, state-level U5M declined by 47%, from 112.8 (105-120.5) to 59.7 (50.4-71.8), but absolute inequality only declined slightly, from a difference of 63 to 46. In Bihar, state-level U5MR and inequality declined considerably between 2000 and 2017: state-level U5MR dropped 51%, from 88.8 (82-95.6) to 43.8 (35.9-54), while the U5MR spread across all districts declined from 31 to 13. Several north-eastern states were notable exceptions to the trend of declining inequality: Meghalaya

and Assam both displayed greater between-district inequalities in 2017 than in 2000.

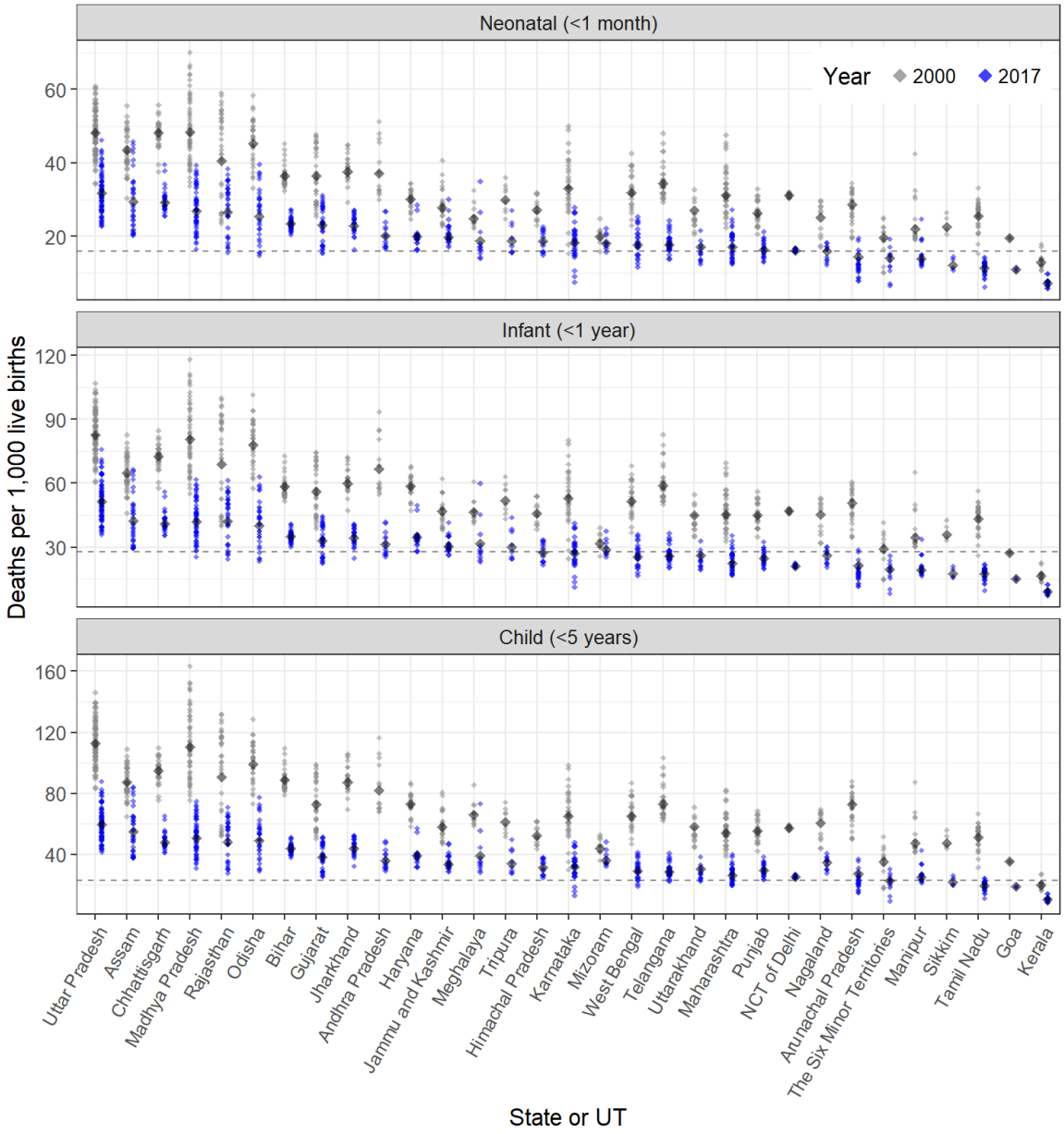


Figure 4.5: Absolute inequalities in U5MR, IMR, and NMR across districts within each Indian state and union territory in 2000 (grey) and 2017 (blue). Each dot represents a district: the lower bound of each vertical spread represents the district with the lowest U5MR in each state and year, while the upper bound of each vertical spread represents the district with the highest U5MR in the same district and year. The large diamond in each vertical spread shows the overall mortality rates across the state as a whole. A blue bar that is shorter than its grey counterpart, as evident across all age groupings in Madhya Pradesh and West Bengal, indicates that absolute between-district inequality has narrowed between 2000 and 2017. The horizontal dashed lines identify NHP targets for each indicator.

Although the NMR, IMR, and U5MR all represent *probabilities* of death rather than true *rates*, the number of deaths in each age group can be estimated by employing a simple demographic transformation and multiplying by underlying child populations in each district derived from WorldPop.[2,55] Estimates of the number of neonatal, infant, and under-5 deaths by year are therefore functions of mortality, district size, and population density in each district. Nevertheless, estimates of death counts are useful guides for child health interventions that can only cover a limited geographic area. In 2017, an estimated 1.03 million (.90-1.20 million) children under the age of 5 died across India. The three districts with the highest number of deaths were Allahabad (10,800 (8,300-13,800) deaths), Sitapur (8,600 (6,900-10,800) deaths), and Bareilly (8,400 (6,800-10,400) deaths), all in Uttar Pradesh: this largely reflects the large child populations of these three districts as well as their exceptionally high child mortality rates. Counter-factual analyses of child mortality rates can also be used to estimate possible lives saved under different interventions. For example, if all Indian districts were to meet the NHP 2025 goal of 23 or fewer child deaths per 1,000 live births, without any change in districts that have already met this goal, the lives of 480,000 children could be saved annually.

4.3.2 Comparison to 2025 and 2030 mortality targets

The left side of Figure 4.6 shows the mean projected under-5 mortality rate for each district in 2025. If under-5 mortality continues to decline at the same rate as it did in most districts from 2000-2017, with emphasis on the trend of decline in recent years, by 2025 only 44 of 723 districts are estimated to have a U5MR greater than 50 deaths per 1,000 live births. The coastal regions of the country as well as Andhra Pradesh are projected to have the lowest under-5 mortality rates in the country, with 22 districts having a mean projected U5MR of less than 10. While in-state inequalities are projected to remain in the central and western states of Rajasthan, Madhya Pradesh, and Uttar Pradesh, absolute differences between the highest-mortality and lowest-mortality districts in these states are all projected to decline between 2017 and 2025. Due to relatively slow rates of mortality decline between 2000 and 2017, eastern Assam and northern Uttar Pradesh are both projected to become the regions with the highest child mortality by 2025.

The right side of Figure 4.6 highlights areas that are likely to meet or not meet the Indian National Health Policy's 2025 target of fewer than 23 deaths per 1,000 live births, accounting for uncertainty in the projections. Districts with a 95% uncertainty intervals (UI) falling exclusively above 23 were marked as unlikely to meet the NHP 2025 goal, while districts with projected UIs falling exclusively below 23 were marked as likely to meet the goal. All other districts were marked in white, indicating that these districts had projected 95% UIs overlapping with the target and could not be assigned to either meeting or missing the target with high certainty. Of India's 723 districts, only 95 were projected to meet the target and 232 were projected to miss the target with high confidence. Only the states of Kerala and Goa and the Union Territory of Puducherry were projected to meet the target in all districts with high confidence. On the other hand, more than half of all districts in Rajasthan, Madhya Pradesh, Uttar Pradesh, Chhattisgarh, Jharkhand, and Odisha were projected to miss the NHP 2025 target with high confidence. Notably, no states included both districts projected to meet and districts projected to miss the target with high confidence.

Figure 4.7, below, plots districts ranked by their levels of under-5 mortality in 2000, 2017, and 2025 projections across the states of Andhra Pradesh, Rajasthan, and Telangana. This figure demonstrates that the trend of converging mortality rates across districts, noted above in reference to Figure 4.5, are projected to continue to 2025 within most states. Uncertainty in district estimates is represented by coloured vertical bars. Due to a larger data sample available in the year 2000 compared to 2017 as well as the uncertainty inherent in the projection method, mortality uncertainty intervals tend to widen from 2000 to 2017 and from 2017 to the 2025 mortality projections.

4.3.3 Comparing local estimates from SRS and survey-based estimates

Estimates of infant mortality rates are the only indicator that the Sample Registration System releases at a sub-state geographic resolution. A spatial comparison between SRS and survey-based estimates of infant mortality can offer insight into the relative strengths of each data source and highlight the implications of using either source to track progress in child survival across India.

Infant mortality estimates from the SRS are published annually at the natural division level across 19 large Indian states. Although smaller than states, natural divisions typically contain sev-

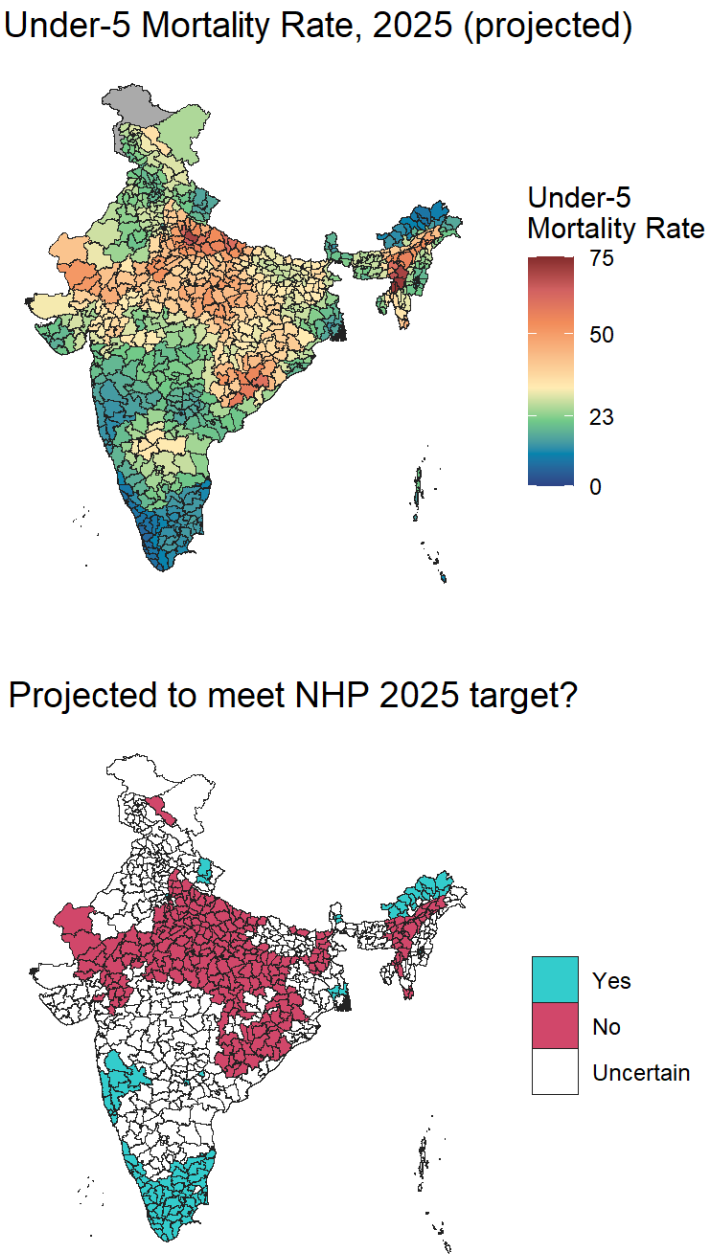


Figure 4.6: *Top* : Mean estimated U5MR by district in 2025, projected forward by extending the annualised rate of decline from 2000-2017. *Bottom* : Comparison between projected 2025 U5MR estimates and the Indian National Health Policy 2017 target of 23 under-5 deaths per 1,000 live births, accounting for uncertainty. Districts shaded in white have a 95% uncertainty interval (UI) that overlaps with the NHP 2025 target; districts shaded in red have a non-overlapping UI that falls above the target; districts shaded in blue have a non-overlapping UI that falls below the target.

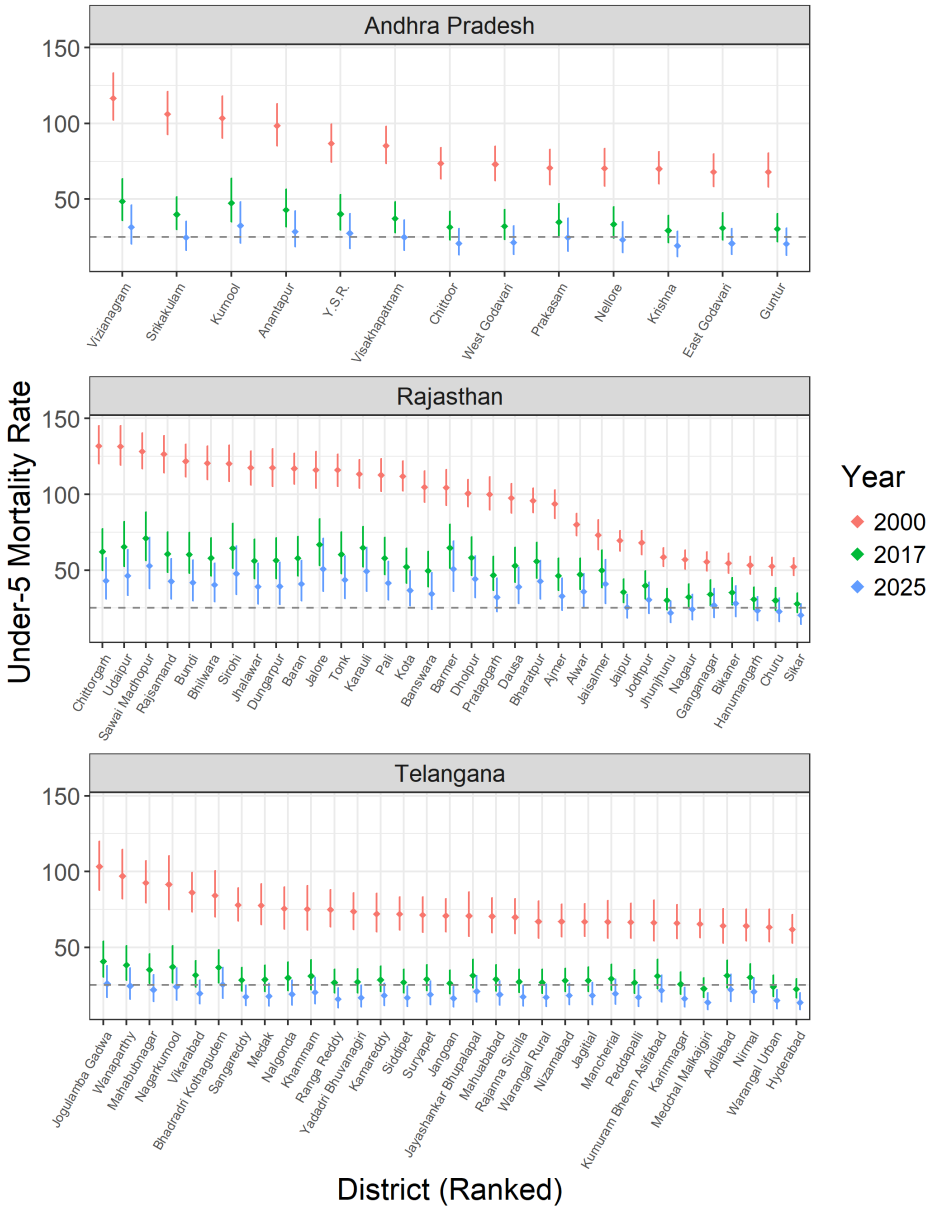


Figure 4.7: Estimated U5MR by district in 2000, 2017, and projected to 2025 across three exemplar Indian states. Each vertical line displays the bounds of the 95% uncertainty interval in a given district and year, and the point within the vertical line displays the corresponding mean U5MR estimate for that district and year. Districts are sorted in decreasing order of their estimates U5MR in 2000. The dotted horizontal line displays the Indian NHP target of fewer than 23 under-5 deaths per 1,000 live births by 2025.

eral districts: according to WorldPop estimates, the 723 districts of India had an average infant population of 32,000 in 2017, while the SRS reported IMR estimates across 68 natural divisions with an average infant population of 324,000.[55]

In 2017, the SRS estimated the infant mortality rate across India to be 33 per thousand live births, compared to a national IMR of of 36.0 (31.6-41.5) estimated by the survey-based model.[175] Figure 4.8 Panel A shows variation in infant mortality across India according to SRS data. SRS data showed similar national trends as the survey-based model, with the highest-mortality areas in the Northern division of Madhya Pradesh state as well as the Central Brahamputra Plains and Cachar Plains divisions in Assam state to the north-east. SRS data shows 7.5-fold variation in infant mortality by division: the lowest-mortality division is the Southern division of Kerala. Although natural divisions obscure local variation occurring at finer spatial resolutions, within-state variation is still apparent in Karnataka state, where recorded IMR varies 2-fold across its constituent natural divisions.

Although the difference in the national IMR calculated between SRS data and the survey-based model falls within the bounds of model uncertainty, a subnational analysis demonstrates why the SRS model estimate falls slightly lower than the mean survey-based estimate. Figure 4.8 Panel B shows the ratio between the IMR mean estimates derived from SRS versus survey data: from this map, it is apparent that the SRS estimates lower infant mortality rates in almost all natural divisions, with the notable exceptions of Madhya Pradesh and two natural divisions in Assam. In the Northern Upper Ganga Plain division of Uttar Pradesh state and Ladakh, the IMR estimated by the SRS is more than 40% lower than the estimate derived from survey data. Figure 4.8 Panel C highlights the natural divisions where SRS infant mortality estimates fall outside the 95% uncertainty intervals of survey-based estimates. Of the 68 natural divisions reported by the SRS, only three divisions are significantly higher in SRS than in the survey-based estimated, including the Plains Western division in Assam; but 11 are significantly lower, including the four northernmost natural divisions in Jammu and Kashmir/Ladakh as well as the Western Plains division of West Bengal.

4.4 Discussion

To develop programs that foster child health across a populous and diverse country, Indian policy-makers need health information systems that can identify local variation in child mortality. Estimates of neonatal, infant, and under-5 mortality across the 723 districts of India reveal key trends and focal areas that should inform the implementation of the National Health Policy's child mortality priorities.

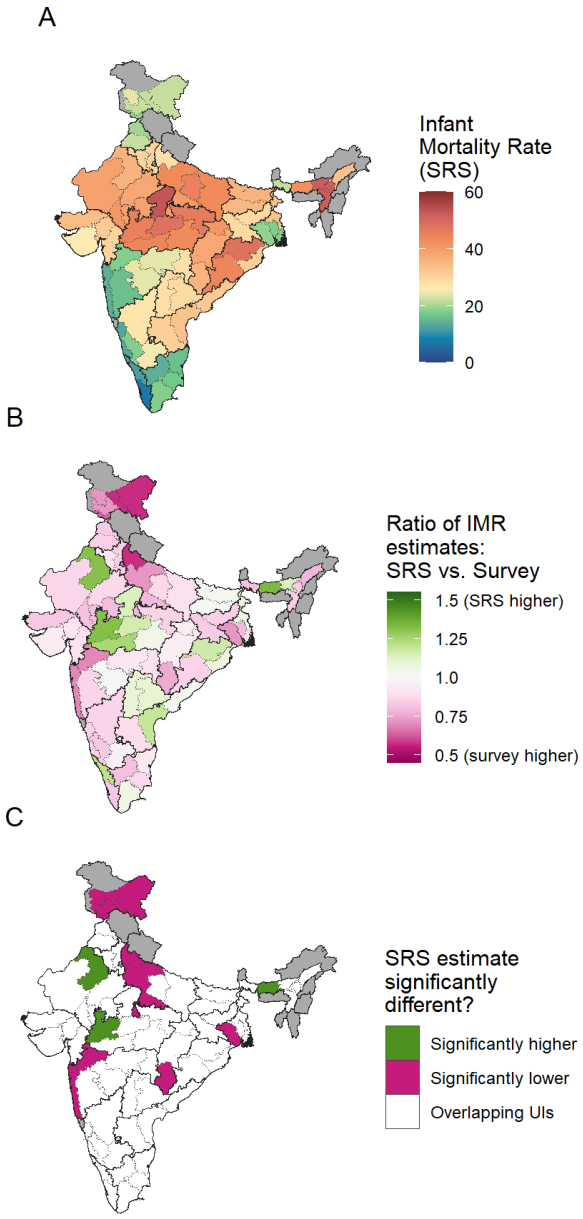


Figure 4.8: *A* : IMR per 1,000 live births in 2017 across 68 natural division in large Indian states, as estimated by the Indian SRS 2017 report. *B* : By aggregating survey-based estimates of IMR to the natural division level, estimates of IMR from SRS and survey sources can be compared. The ratio between SRS-based IMR estimates and survey-based mean IMR estimates are shown at the natural division level. *C* : Comparison between point estimates of IMR provided by the SRS with the 95% uncertainty interval for IMR estimated by survey data. Compared to survey-based estimates in 2017, the SRS estimated significantly lower IMR for 11 of 68 natural divisions, and significantly higher IMR for 3 of 68 natural divisions.

Given that the probability of a child dying before their fifth birthday ranges from 9 to 88 per thousand (1 to 9 percent) depending on their district of residence, child health program administrators should

consider expanding programs in areas where child health burden is highest. This observation also holds for the 36 states and union territories of India, where the under-5 mortality rate can vary up to 6-fold. Only three states and union territories, Goa, Kerala, and Puducherry, are on track to meet the NHP U5MR target of 23 deaths per 1,000 live births in all of their districts with high certainty, meaning there is an urgent need to increase both maternal and child health services as well as local health data collection across most of the country. This analysis also reveals regional hot spots of high under-5 mortality—notably in eastern Assam, Tripura, and Mizoram—that are obscured by state-level estimation projects. These hot spots offer an opportunity for further investigation of the underlying conditions that are impeding progress in child survival.

These insights are enabled by georeferenced birth histories provided by surveyed mothers. These birth histories have several advantage over data from the Sample Registration System and Civil Registration System, particularly the advantage of a known, and complete, denominator for child mortality estimation. However, if reported at the district level, SRS and CRS data could also supplement some of the limitations evident in survey-based estimation. Unlike SRS and CRS data, retrospective survey data suffers from diminished sample sizes in more recent years of reporting, leading to increased uncertainty in the years of analysis with the greatest policy relevance. Additionally, surveys only reliably capture mortality in young age groups, while CRS and SRS data could report age-specific mortality across the entire life course.

4.4.1 Scale and inequality in mortality

This analysis demonstrates that spatial scale matters for identifying disparities in child survival across India. According to the spatial model presented here, infant mortality varies 5.6-fold at the state level (the level reported by the CRS); 7.3-fold at the natural division (the level reported by the SRS); and 10.3-fold at the district level. As results are presented at more aggregate levels, focal areas with persistently high mortality as well as exemplar areas displaying substantial improvements over time are convolved with other districts, obscuring patterns that could be used to uncover mortality drivers or develop interventions. Given that approximately 160,000 children under 5 live in the average Indian district,[55] this analysis presumably obscures even more local disparities in child mortality.

This analysis also allows policy-makers to target and prioritise interventions to specific districts with the highest mortality rates or greatest number of child deaths. As shown above, both mortality rates and counts are distributed highly unequally across the country and within states. Given previous evidence about the context dependence of child health burden across India, interventions must be tailored to survival in particular age groups and urban-rural context, balancing between the priorities of efficacy and equity in service delivery.[170,172]

4.4.2 Limitations of survey-based spatial mortality mapping

As the only data source that includes the spatial identifiers needed to conduct a district-level analysis, retrospective birth histories from India's three household survey series were essential to this analysis. In addition to their spatial metadata, these surveys facilitate analysis thanks to a set of questionnaires designed to minimise bias, reporting of sampling weights associated with each household, and a documented methodology that is comparable across survey series and even across countries.[32,34] However, survey-based spatial models of mortality are also subject to limitations inherent in the data and modelling approach that limit some aspects of their policy utility. These limitations could be alleviated if SRS or CRS data was made available to compare with survey-based estimates.

The first programmatic challenge of this survey-based model relates to data availability over time. Although policy makers tend to be interested in the most recent year of analysis, birth history, as a retrospective data source, increases in quantity farther back in time. Figure 4.2, which shows the combined sample sizes of all birth history survey sources used in this analysis, demonstrates this issue across India between 2000 and 2017. Thanks to retrospective data collection, birth histories covering the year 2000 are available from all eight DLHS, NFHS, and AHS surveys collected since the year 2000. Conversely, in 2016, only a partial year of data is available from the 2016 NHFS—and, in the final year of analysis, no retrospective data is available. While the spatial model leverages known relationships over the dimensions of space, time, and age and with covariates to produce estimates of mortality in the final year, this estimate tends to be more uncertain due to the scarcity of underlying data.

In many Indian states, decreasing sample sizes and converging between-district mortality rates complicate district rankings in 2017 as well as projections forward to 2025. Figure 4.7 demonstrates

this limitation by showing under-5 mortality estimates by district in 2000, 2017, and 2025 for three states. Districts are ordered according to their estimated under-5 mortality rate in 2000. The states of Andhra Pradesh, Rajasthan, and Telangana, showed stark disparities in U5MR between their lowest- and highest-burden districts in 2000. These disparities, while reflecting an unconscionable situation on the ground, facilitate making clear and highly-certain distinctions between areas with low and high mortality within a state: for example, the difference in U5MR between the districts of Chittorgarh and Sikar in Rajasthan was 80 per thousand in the year 2000. The progress that all three states have made in reducing between-district mortality inequality, while laudable, increases the analytical difficulty of identifying the highest-mortality districts in each state in 2017. Ranking the relative mortality of districts in 2025 becomes even more difficult given both the convergence in mortality rates and uncertainty in the underlying annualised rates of change from 2000 to 2017, a symptom of a shrinking sample size. Uncertainty in the most appropriate projection method also inhibits the interpretability of results: although this analysis used a simple projection strategy that applied annualized rates of change to the final year of data, a more sophisticated approach might apply historic rates of change in low-mortality districts to better predict the likely mortality trajectory of medium- and high-burden districts. There is a danger for the simpler projection applied in this chapter to under-estimate future child mortality by missing asymptotic patterns in mortality reductions over time. While future district rankings are a highly sought-after metric for targeting program funding within states, these rankings should be communicated extremely carefully due to their inherent uncertainty.

Another concern relates to potential bias in unsampled locations, which can be sensitive to model specification. The spatial model described in this chapter makes predictions across 5 x 5 km grid cells for each estimated year and age group. In location-age-years not overlapping sampled data points, the model estimates the underlying mortality level using covariate information and a latent surface that is influenced by nearby observations in the dimensions of space, time, and age.[2,11] This introduces concerns that relationships between child mortality and underlying covariates may not hold outside of the sampled data points. In the current model, covariates are capped at their minimum and maximum observed values: this avoids unfounded extrapolation of covariates at unobserved values, but may also miss relationships at the extremes of child mortality that were not found in any sampled location.

While different functional forms can be applied to estimate the relationship between mortality and underlying covariates, additional model flexibility can come at the expense of interpretability in a policy context.[186,187] SRS and CRS data sources can alleviate the issue of extrapolation: because every district is surveyed, they open up the possibility of using small-area model frameworks where data has been sampled across every space-time-age combination in the study area.[64]

4.4.3 Opportunities offered by spatial reporting of CRS and SRS data

When SRS and a survey-based model produce conflicting estimates of infant and child mortality across India, it begs the question: which source is correct? Other studies comparing multiple sources of mortality data across India have measured data quality against a selected "gold standard."[35] When comparing SRS and survey-based estimates, I assert that survey data should be used as a gold standard, subject to the limitations of model specification and uncertainty. While the SRS has a larger sample size than all household surveys combined, and includes a census component that is intended to ensure completeness, past evidence as recently as 2007 suggests that in practice, past rounds of SRS data can be highly incomplete.[87,176] Mortality data in the SRS is also sensitive to the population denominator in each sampled location, which may miss early neonatal deaths as well as migration in and out of the study area. While survey-based estimates are limited by relatively small sample sizes in recent years, this disadvantage should be reflected in the uncertainty surrounding mortality estimates in 2017, given that the model is correctly specified. This chapter identified 14 natural divisions where SRS estimated infant mortality rates falling outside of the 95% uncertainty intervals estimated by the survey-based model: these differences offer an opportunity for further investigation, with an emphasis on barriers that may be impeding complete reporting for the SRS.

Despite these limitations, the SRS is still the most promising data source for future district-level analysis of mortality across all age groups in India. Birth history data is generally used only to predict mortality rates among children, due to sparse data and potential recall bias for predicting mortality among older age groups.[108] Conversely, SRS data completeness is likely to be higher for reported adult deaths than among children. Lower completeness in under-5 mortality reporting may be due in part to different cultural conceptions of mortality in the neonatal context or social insurance

programs offered to the relatives of deceased adults.[35] If SRS data was reported at the district level or by sampling unit, its much larger sample size and annual time trends would enable policy-makers to assess spatial variation in mortality with greater precision, and to more confidently assess progress towards future mortality targets. It is a near-certainty that adult mortality, like child mortality, is distributed unequally across the country in patterns not fully reflected by state-level tabulations. Even an imperfect accounting of these differences would greatly increase our understanding of the diverse epidemiological contexts across India.

When discussing estimates of all-cause mortality rates, the discussion so far has centred around survey data and the Sample Registration System. When estimating all-cause mortality rates, the Indian Civil Registration System is widely understood to be incomplete compared to these other two data sources.[35,188] To understand why the CRS is a vital component needed to achieve the National Health Policy, we must turn to the four principles of a successful vital registration system: completeness, accuracy, universality, and confidentiality.[14] While source-specific completeness has been one focus of this chapter, detailed CRS tabulations could also inform mortality estimates by acting as a floor for the total number of deaths that occurred in a given district, age group, and year. However, we should understand the CRS not just as a source of health statistics, but as the foundation for legal rights offered to all Indian citizens that can only be obtained through birth and death registration. When it comes to the principle of universality, and the legal right of all Indians to access the benefits of civil registration, only the CRS is designed to ultimately cover all residents of India. The CRS also adds cause-of-death assignment as another dimension of mortality reporting. While cause-of-death classification accuracy remains challenging for deaths that occurred in the home,[189] even flawed data on causes of mortality provides a wealth of information about local variation in the epidemiological transition taking place across the country.

Finally, while local reporting of SRS and CRS data would enable new policy-relevant insights about the state of health across India, strict practices must be maintained to ensure that all records remain confidential. Reporting standards for spatially-resolved SRS and CRS data could follow restrictions developed in other countries, such as the United States and United Kingdom, that require aggregation of any mortality data where counts fall below a fixed threshold.[190] The potential

benefits of publicly sharing additional health information must also be balanced by the need for proper data governance so that new insights benefit the communities being measured.[74] Opportunities for additional data sharing must be reviewed through the proper channels in the Indian Ministry of Health, which has historically governed both the collection and responsible dissemination of health data nationwide.[191]

4.4.4 Conclusions

India's adoption of the 2017 National Health Policy indicates the need for local estimates of health burden, particularly among children in the first five years of life. Despite this need, no data system currently provides district-level estimates of neonatal, infant, and child mortality. In this chapter, I develop a model for district-level mortality estimation based on household surveys, the only data source reported at the district level and below. This model provides new insights that can inform the drive to meet neonatal and under-5 mortality targets by 2025; however, spatially-resolved data from the Sample Registration System is better-suited to describe local variation in mortality among older age groups. Ultimately, only the birth and death certificates administered by the Civil Registration System can extend the full legal rights afforded by vital records to all Indian citizens.

This chapter has also investigated the relationship between household survey data and the Sample Registration System across 68 natural divisions of India. Comparisons between the two data sources were limited by the relatively large size of each natural division, which concealed local spatial variation; instead, this analysis relied on past literature estimating the relative quality of SRS data. Were a time series of SRS mortality estimates made available at the district level, a joint estimation model such as the one presented in Chapter 2 could be applied to synthesise the SRS and survey data sources.

4.5 References

1. UN General Assembly. *Universal declaration of human rights* 1948.
2. Burstein, R. *et al.* Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature* **574,** 353–358. http://www.nature.com/articles/s41586-019-1545-0 (Oct. 2019).
7. World Health Organization. *Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies* tech. rep. 1 (2010), 1–92. http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.35.021103.105711.

11. Diggle, P. J. & Giorgi, E. Model-based geostatistics for prevalence mapping in low-resource settings. *Journal of the American Statistical Association* **111,** 1096–1120. arXiv: 1505.06891. http://dx.doi.org/10.1080/01621459.2015.1123158 (2016).
14. United Nations Statistics Division. *Principles and Recommendations for a Vital Statistics System Revision 3* **19.** http://unstats.un.org/unsd/Demographic/standmeth/principles/M19Rev3en.pdf (2014).
32. Corsi, D. J., Neuman, M., Finlay, J. E. & Subramanian, S. V. Demographic and health surveys: A profile. *International Journal of Epidemiology* **41,** 1602–1613 (2012).
34. Dandona, R., Pandey, A. & Dandona, L. A review of national health surveys in India. *Bulletin of the World Health Organization* **94,** 286–296A (2016).
35. Kumar, G. A., Dandona, L. & Dandona, R. Completeness of death registration in the Civil Registration System, India (2005 to 2015). *Indian Journal of Medical Research* **149,** 740. http://www.ncbi.nlm.nih.gov/pubmed/23144490%20http://www.ijmr.org.in/text.asp?2019/149/6/740/265955 (2019).
45. Lindgren, F. & Rue, H. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society. Series B* **73,** 423–498 (2011).
51. Krainski, E. *et al. Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA* **September** (2018).
55. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* **4,** 2–5 (2017).
60. Weiss, D. J. *et al.* Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000-17: a spatial and temporal modelling study. *The Lancet* **394,** 322–331. http://dx.doi.org/10.1016/S0140-6736(19)31097-9 (2019).
62. Osgood-Zimmerman, A. *et al.* Mapping child growth failure in Africa between 2000 and 2015. *Nature* **555,** 41–47. http://www.nature.com/articles/nature25760 (Mar. 2018).
64. Wakefield, J., Okonek, T. & Pedersen, J. Small Area Estimation for Disease Prevalence Mapping. *International Statistical Review,* insr.12400. https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12400 (July 2020).
74. Cinnamon, J. Data inequalities and why they matter for development. *Information Technology for Development* **26,** 214–233 (2020).
81. Abouzahr, C. & Boerma, T. Health information systems: the foundations of public health. *Bulletin of the World Health Organization* **83,** 578–583. arXiv: arXiv:1011.1669v3 (2005).
87. Bhat, P. N. Completeness of India's sample registration system: An assessment using the general growth balance method. *Population Studies* **56,** 119–134 (2002).
89. Dicker, D. *et al.* Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **392,** 1684–1735. https://linkinghub.elsevier.com/retrieve/pii/S0140673618318919 (Nov. 2018).
108. Wakefield, J. *et al.* Estimating under-five mortality in space and time in a developing world context. *Statistical Methods in Medical Research* **28,** 2614–2634 (2019).
168. Government of India Ministry of Health and Family Welfare. *National Health Policy 2017* tech. rep. (Government of India, New Delhi, 2017), 28. https://www.nhp.gov.in/nhpfiles/national_health_policy_2017.pdf.
169. Sundararaman, T. National Health Policy 2017: a cautious welcome. *Indian Journal of Medical Ethics* **2,** 69–71 (2017).
170. Government of India Ministry of Health and Family Welfare. *Situation Analyses: Backdrop to the National Health Policy 2017* tech. rep. (Government of India, New Delhi, 2017), 14.
171. Yadav, S. & Arokiasamy, P. Understanding epidemiological transition in India. *Global Health Action* **7** (2014).
172. Dandona, L. *et al.* Nations within a nation: variations in epidemiological transition across the states of India, 1990-2016 in the Global Burden of Disease Study. *The Lancet* **390,** 2437–2460 (2017).

173. Parliament of the Republic of India. *The Registration of Births and Deaths Act, 1969* New Delhi, 1969.
174. Subramanian, S. A Census and Statistics Act for India. *Calcutta Statistical Association Bulletin* **18,** 85–96. http://journals.sagepub.com/doi/10.1177/0008068319690204 (June 1969).
175. Office of the Registrar General & Census Commissioner. *Sample Registration System Statistical Report 2017* tech. rep. (Ministry of Home Affairs, Government of India, New Delhi, 2017), 337. http://www.censusindia.gov.in/vital_statistics/SRS_Report/9Chap%202%20-%202011.pdf.
176. Mahapatra, P. *An Overview of the Sample Registration System in India* in *Prince Mahidol Award Conference & Global Health Information Forum* (Institute of Health Systems, Bangkok, 2010), 1–13. http://www.ihs.org.in/IHS/PMahapatra-SRSInIndiaAnOverview.pdf.
177. Abouzahr, C. *et al.* The way forward. *Lancet Health Metrics Network, World Health Organization* **370,** 1791–99 (2007).
178. Mohanty, I. & Gebremedhin, T. A. Maternal autonomy and birth registration in India: Who gets counted? *PLoS ONE* **13,** 1–19 (2018).
179. Dandona, R. *et al.* Subnational mapping of under-5 and neonatal mortality trends in India: the Global Burden of Disease Study 2000-17. *The Lancet* **395,** 1640–1658 (2020).
180. Ahmad, O. B., Lopez, A. D. & Inoue, M. The decline in child mortality: a reappraisal. *Bulletin of the World Health Organization* **78,** 1175–91. http://www.ncbi.nlm.nih.gov/pubmed/11100613%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2560617 (2000).
181. Huang, H. C., Martinez, F., Mateu, J. & Montes, F. Model comparison and selection for stationary space-time models. *Computational Statistics and Data Analysis* **51,** 4577–4596 (2007).
182. Local Burden of Disease HIV Collaborators. Mapping subnational HIV mortality in six Latin American countries with incomplete vital registration systems. *BMC Medicine* **19,** 1–25. https://link.springer.com/article/10.1186/s12916-020-01876-4%20https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-020-01876-4 (Dec. 2021).
183. Marquez, N. *ar.matrix: Simulate Auto Regressive Data from Precision Matricies* 2018. https://cran.r-project.org/package=ar.matrix.
184. Ghosh, R. Child mortality in India: A complex situation. *World Journal of Pediatrics* **8,** 11–18 (2012).
185. Zomuanpuii, R. *et al.* Epidemiology of malaria and chloroquine resistance in Mizoram, northeastern India, a malaria-endemic region bordering Myanmar. *Malaria Journal* **19,** 1–11. https://doi.org/10.1186/s12936-020-03170-3 (2020).
186. Pletcher, S. D. Model fitting and hypothesis testing for age-specific mortality data. *Journal of Evolutionary Biology* **12,** 430–439 (1999).
187. Lucas, T. C. A translucent box: interpretable machine learning in ecology. *Ecological Monographs* **90,** 1–55 (2020).
188. Gupta, M., Rao, C., Lakshmi, P., Prinja, S. & Kumar, R. Estimating mortality using data from civil registration: a cross-sectional study in India. *Bulletin of the World Health Organization* **94,** 10–21 (2016).
189. Kotabagi, R., Chaturvedi, R. & Banerjee, A. Medical certification of cause of death. *Medical Journal Armed Forces India* **60,** 261–272 (2004).
190. Thacker, S. B. & Berkelman, R. L. Public Health Surveillance in the United States. *Epidemiologic Reviews* **10,** 164–190 (1988).
191. Raban, M. Z., Dandona, R. & Dandona, L. Essential health information available for India in the public domain on the internet. *BMC Public Health* **9,** 1–19 (2009).

Chapter 5

Variation in COVID-19 excess mortality by age, sex, and province within Italy

5.1 Introduction

Italy received international attention as one of the first countries outside of China to experience a major COVID-19 outbreak. On March 9, 2020, Italy announced a nationwide lock-down to stem community transmission, and deaths peaked two weeks later during the week of March 25[192]. After that peak, deaths declined for the following two months, and lock-down restrictions were gradually eased starting in late May[193]. In total, between February and August 2020, approximately 35,500 COVID-19 deaths were registered across Italy, equivalent to approximately 60 deaths per 100,000 people[194].

Aside from the timing and magnitude of the first wave of COVID-19 transmission within Italy, two notable features set its COVID-19 epidemic apart from subsequent epidemics in other European countries. First, registered COVID-19 cases and deaths were unevenly distributed across the regions of Italy[195], and that cause-specific mortality from COVID-19 varied by age and sex[196]. Registered COVID-19 deaths were highest in the northern regions of the country, particularly in the region of Lombardy, where the registered COVID-19 death total amounted to over 160 deaths per 100,000 people (Figure 5.1). Another salient feature of the Italian epidemic was the early recognition among health authorities that registered COVID-19 deaths under-counted the full mortality burden of the epidemic.[197,198] In May 2020, the Italian National Institute of Statistics (Istat) re-

ported that while 13,700 COVID-19 deaths had been registered across Italy between 20 February and 31 March, deaths from any cause had increased by 25,300 compared to an expected baseline during the same time period, suggesting that the full mortality burden of the COVID-19 epidemic was nearly double what had been previously reported[199].

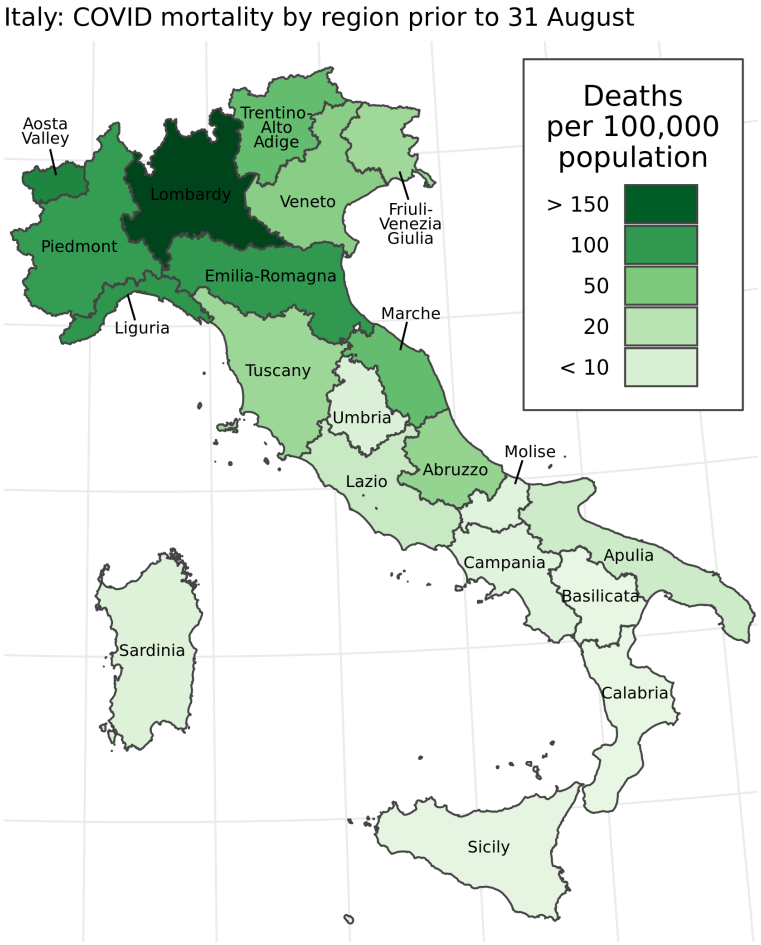


Figure 5.1: Registered COVID-19 deaths per 100,000 people by Italian region prior to August 31, 2020. Nationally, 35,500 deaths were registered during this time, or approximately 59 deaths per 100,000 population.

Istat and other research groups drew conclusions about the mortality burden of the COVID-19 pandemic based on a series of excess mortality analyses. Excess mortality analyses attempt to measure the net effect of a discontinuity, such as a COVID-19 outbreak, on all-cause mortality. This is a two-step process: first, the investigator constructs an estimated baseline number of deaths expected during the period in question. While various methods have been used in the past to construct this

baseline[70,200,201], previous studies in Italy have averaged the number of deaths recorded by week in the years from 2015 through 2019 to generate a baseline estimate for the same weeks in 2020[193,202]. Next, the investigator compares that expected baseline with the count of recorded deaths during the same time period. Excess deaths, a count, are measured as the difference between the observed death count and the expected baseline. Standardised mortality ratios (SMRs) are measured as the ratio between the observed death count and the expected baseline[115,203]. Excess mortality analyses of the COVID-19 pandemic are becoming widely used in the media: among others, the Economist, the Financial Times and the New York Times have estimated excess mortality across dozens of countries in Europe, the Americas, and Asia[204–206]. Previous investigations have also examined how excess mortality analysis can capture deaths caused by COVID-19 which were attributed to other causes, as well the indirect mortality burden of the COVID-19 pandemic[70,207].

This study explores another central issue for excess mortality estimation: how can public health officials detect increases in mortality at the local level or across multiple age groups, where the expected number of baseline deaths in each sub-population of interest is relatively low? This question addresses a tension central to many forms of public health surveillance, where the imperative to identify clustering and high-risk subgroups in health surveillance data must be balanced against reduced study power and possible biases associated with small samples[190]. Any suitable approach for estimating small-group excess mortality must quantify uncertainty due to stochastic variation as well as limited data informing the baseline. Even at the national level, weekly estimates of COVID-19 excess mortality presented without uncertainty intervals can leave viewers confused about what constitutes a meaningful departure from the baseline.

To better understand the relationship between registered COVID-19 deaths and excess mortality across Italy, I developed a model to estimate excess mortality by age group, sex, and week across the country's 107 provinces. This approach estimated counter-factual baseline mortality rates for each of these groups from March through August 2020, based on mortality rates and predictive covariates observed from January 2015 through February 2020. The baseline mortality model combined elements from a widely-used Poisson generalised linear modelling (GLM) strategy for mortality estimation[201]; a structured province-year-age random effect that draws power from local correlations

in mortality across those three dimensions, based on disease mapping principles[57]; and a Fourier curve-fitting method to capture seasonal trends in age-specific all-cause mortality[200]. Once baseline mortality was estimated for each week and sub-population in the study period, I compared these estimates to observed counts from vital records over the same period while preserving uncertainty in the expected baseline mortality. Here, I present major findings on excess deaths (the difference between observed and baseline death counts) as well as SMRs (the ratio between observed and expected deaths) across these sub-populations of Italy.

5.2 Methods

5.2.1 Overview

I fitted a spatially-explicit hierarchical model with fixed effects by age group and for seven covariates, correlated age-province-year structured random effects, and harmonic curves capturing seasonal variation for each age group and province. Separate models were fit for each sex. I fit this model using mortality and population data from 1 January 2015 through 25 February 2020, then generated 1,000 predictive samples of the baseline mortality rate for each sex, age group, and province for the weeks of 26 February through 31 August 2020. For each of the 1,000 sampled draws, I compared the baseline mortality rate with the observed mortality rate to estimate a Standardised Mortality Ratio, and compared the predicted baseline deaths with observed deaths to estimate excess deaths. Figure 5.2, below, summarizes the analyses described in this chapter.

5.2.2 Data

Mortality and population data were downloaded from Istat, the Italian National Institute of Statistics. As of 22 October 2020, complete mortality data covering all provinces and municipalities of Italy over the time period 1 January 2015 through 31 August 2020 was available for download from Istat[208]. The number of deaths over this time period were recorded by year, month, day, Italian municipality, sex, and five-year age group. For the purposes of analysis, these observations were aggregated by sex, Italian province, age group, and week of the year. The five age groups used in this analysis were 0-59 years, 60-69 years, 70-79 years, 80-89 years, and 90+ years of age. These age groups

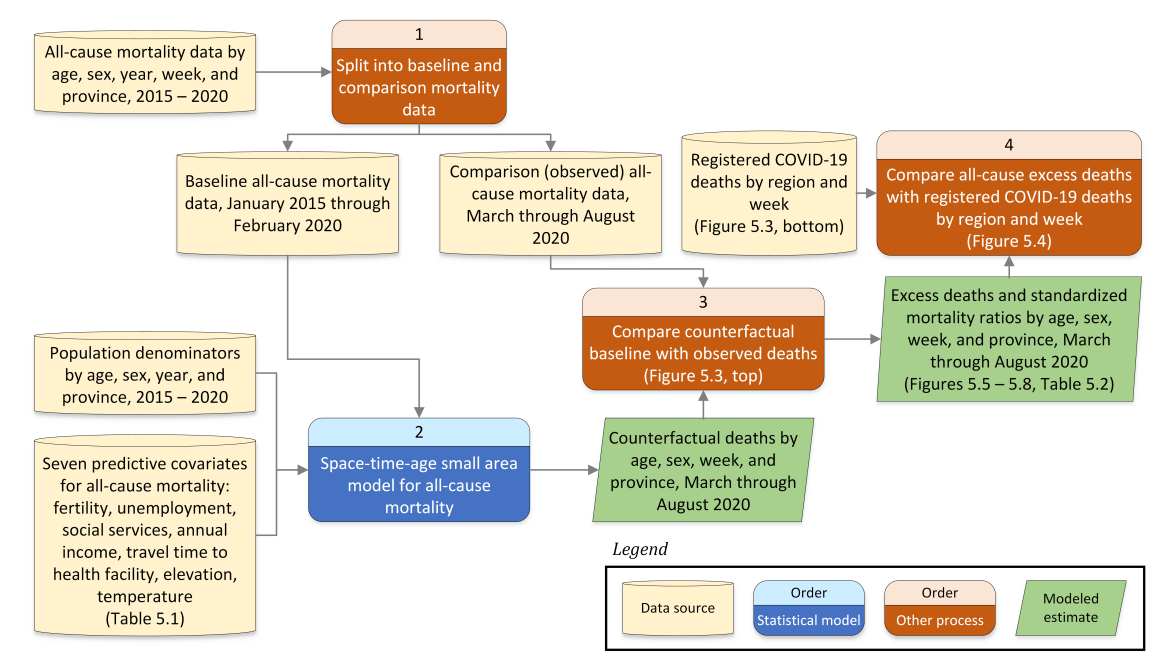


Figure 5.2: Summary of data sources, analyses, and results described in Chapter 5.

were chosen based on the prior knowledge that the large majority of both all-cause mortality and registered COVID-19 deaths occurred among adults aged 60 and above, meaning that relatively few deaths would be observed among the 0-59 age group. Weeks of the year were assigned based on the numeric day of the year, where January 1st of each year was assigned as the first day of the first week. The 365th and 366th days of the year were assigned to week 52, with the hierarchical model adjusting for observed weeks with more than seven days.

Population data by sex, age, and province for the years 2015 through 2020 was downloaded from the Istat web data portal[209]. Population counts were aggregated by sex, province, year, and the five age groups listed above.

I downloaded or extracted data for each of seven covariates, listed below in Table 5.1. After extraction, all covariates were normalised and rescaled to have a mean of zero and a standard deviation of 1 across all data observations.

Table 5.1: Covariates used to estimate baseline mortality by age group, sex, province, and week across Italy from January 2015 through August 2020. The source and space-time resolution of each covariate is listed.

Covariate	Source	Varies Over
Total fertility rate	Istat	Province, Year
First quarter unemployment	Istat	Province, Sex, Year
Proportion of eligible households receiving at-home social services	Istat	Province, Year
Proportion of households with taxable annual income below 10,000 Euros	Istat	Province, Year
Average driving time to the nearest health facility	Malaria Atlas Project	Province
Average elevation of residence	US Geological Survey	Province
Temperature	MeteoStat	Province, Year, Week

5.2.3 Space-time model

To construct a mortality baseline for the weeks of 26 February through 31 August 2020 that incorporated multiple sources of uncertainty, I fit a small area model with age and covariate fixed effects, correlated province-year-age errors, and harmonic terms to capture seasonality within each age grouping and province. Previous studies have demonstrated that the age structure of all-cause mortality varies by sex, and COVID-19 excess mortality may vary by both age and sex.[89,210] To better understand differences in age-specific mortality by sex, two models were fit for males and females. For a particular sex, the number of deaths in a given province p, age group a, year t, and week of the year w was assumed to follow a Poisson distribution:

$$D_{p,a,t,w} \sim Poisson(N_{p,a,t,w} r_{p,a,t,w})$$

In the formulation above, D is the number of observed deaths, N is the population, and r is the underlying mortality rate per person-week. The quantity r is then fit in log space to a space-time surface which varies by province, age, year, and week:

$$log(r_{p,a,t,w}) \sim \sum_{k=1}^{5} [I_\alpha \alpha_k] + \beta X_{p,a,t,w} + Z_{p,a,t} + f_{p,a}(w)$$

The first three terms on the right-hand side of this equation capture age and covariate fixed effects, corresponding to a discrete-time proportional hazards model where the baseline hazard varies by age group[2,211]. In this specification, α_k is the weekly baseline hazard for each of the five age

groups, while I_α is a boolean variable that is 1 when the age group index of an observation is equal to k and zero otherwise. Fixed effects for the covariate design matrix $X_{p,a,t,w}$ are denoted by β, a vector of length seven. Together, these terms correspond with a multivariate regression approach to estimating baseline mortality[212].

The term $Z_{p,a,t}$ is a structured random effect that accounts for residual variation across provinces, age groups, and years that is not captured by the age or covariate fixed effects. Z is structured as a Gaussian process with mean zero and covariance matrix K, where K is a separable process across the dimensions of space, age, and time: $K = \Sigma_p \otimes \Sigma_a \otimes \Sigma_t$. The spatial covariance structure Σ_p corresponds to a conditional autoregressive (CAR) process in space[43], while the age and temporal covariance structures both correspond to discrete autoregressive processes of order 1. Separable covariance structures have been widely used in the fields of ecology and public health to construct models across space, time, and other dimensions[108,213], and have been found to fit a wide variety of space-time covariance structures[181].

The term $f_{p,a}(w)$ refers to a set of harmonic functions that are fit to account for weekly variation in mortality not captured by covariates. A separate function is fit for each age group and province to account for the fact that seasonal variation in mortality may be driven by different factors across space and by age group. Each function is tuned to fit the parameters A and B to the following harmonics:

$$f(w) = \sum_{j=1}^{2} [A_j sin(\frac{2\pi jw}{52}) + B_j cos(\frac{2\pi jw}{52})]$$

The harmonic series described in the equation above, which adapts principles from Fourier analysis, is the basis for a classic model for predicting seasonality in flu mortality developed by Robert Serfling[200]. In Serfling's original formulation as well as more recent excess mortality papers, seasonality was fit using two Fourier terms[70,214]. I performed five-fold cross-validation to determine the best grouping variables and harmonic terms for seasonal curve fits. Based on the metrics of out-of-sample mean squared error and coverage, the model performed best when seasonal curves were fit separately by province, sex, and age group, using two Fourier terms.

I assigned priors to all model parameters and then fit the model using the Laplace approximation for mixed-effect parameter estimation[118,119]. The model was fit in R v.4.0.3 using the package Template Model Builder v.1.7.18[118,120].

5.2.4 Compiling and interpreting results

Using the maximum a posteriori predictions and joint precision matrix for all parameters, I generated 1,000 samples for all model parameters using a multivariate-normal approximation of the posterior predictive distribution. These parameter samples were then entered into the original model to construct 1,000 draws or "candidate maps" estimating the mortality rate across all provinces, age groups, and weeks in the study period[71]. Although the model was fit to data from 1 January 2015 through 25 February 2020, the fitted parameter fixed effects, random effects, and seasonality terms could all be applied forward to estimate 1,000 draws of predicted baseline mortality from 26 February through 31 August 2020. All subsequent calculations were performed across draws to preserve the correlation structure within draws as well as the model uncertainty across draws.

I compared the distribution of predicted mortality rates with observed mortality rates, calculated as observed deaths divided by population, to calculate 1,000 draws of standardised mortality ratios for each province-age-sex-year-week grouping g using the following formula:

$$SMR_{g,draw} = \frac{Observed\ Deaths_g}{Predicted\ Deaths_{g,draw}}$$

I also multiplied the predicted mortality rates by the population in each province-age-sex-year grouping to calculate predicted baseline death counts for each draw. I then calculated 1,000 draws of excess deaths for each grouping:

$$Excess\ Deaths_{g,draw} = Observed\ Deaths_g - Predicted\ Deaths_{g,draw}$$

In the results section below, draws for predicted mortality, SMRs, and excess deaths are summarised using the mean and 95% uncertainty interval bounds. The 95% uncertainty interval is reported as the 2.5th percentile and 97.5th percentile of values across 1,000 draws.

5.3 Model validation

I used five-fold cross validation to compare predictive performance across multiple model specifications and to compare predictive performance with simpler models for calculating excess mortality. Each fold was created by fitting the model without data from the weeks in March through December for each of the years 2015 through 2019, then comparing predicted values for the held out weeks with the observed values. This holdout strategy mirrors the baseline deaths scenario for the months of March through August 2020 in the counter-factual where COVID-19 did not change the pattern of mortality across Italy.

Because the expected number of deaths in a given province-age-sex-year-week groupings can be very low, particularly in lower age groups, I aggregated all out-of-sample observations across four-week intervals while preserving the other groupings. I then calculated the difference between the out-of-sample recorded deaths and the modelled mortality, and calculated summary metrics: root mean squared error, coverage of the 95% uncertainty intervals, and relative squared error when compared to a simpler model that uses the average mortality rate across all other years.

I found that the out-of-sample root mean squared error for the best-performing model was 2.32E-5, compared to an average weekly mortality rate of 2.05E-4 across all age groups, suggesting a reasonably good fit for the model's mean estimates. The out-of-sample relative squared error was 0.330 compared to the simple method of averaging weekly values across other years, suggesting that this predictive model substantially outperformed the simpler alternative for the years 2015-2019 even when an entire year of data was held out. The in-sample relative squared error compared to the simpler averaging method was 0.273, a much lower ratio of error, which indicates that the model provides a more flexible fit to the data than the simpler averaging strategy. The out-of-sample coverage of the 95% uncertainty interval was 99.1%, indicating that the predicted uncertainty bounds are conservative.

5.3.1 Data availability

All data sources used in this analysis are publicly available on-line. The code repository accompanying this chapter, described in Appendix E, contains detailed instructions for downloading and formatting each data source.

5.4 Results

5.4.1 Excess mortality exceeds registered COVID-19 deaths

At the national level, the aggregated results from the excess mortality model generally agreed with previous national studies on the timing and magnitude of excess deaths associated with the COVID-19 pandemic across Italy. Figure 5.3, below, shows the estimated weekly death count (in brown) as well as the observed weekly death count (in black) from 26 February through 31 August 2020 across Italy. In line with other studies, the model results suggest that excess mortality peaked on the week of March 25, on the same week as the peak in registered COVID-19 deaths, and then consistently fell until approximately returning to baseline by the end of May 2020, coinciding with the lifting of most lock-down measures across Italy[193]. At no point between June and August 2020 did the observed death count exceed the upper bound of the 95% uncertainty interval for baseline deaths: as a result, the remainder of this section will focus on excess mortality during the 13-week period from 26 February through 26 May 2020. During these 13 weeks, an estimated 53,200 excess deaths (95% uncertainty interval 26,500 to 79,700) were associated with the COVID-19 pandemic, compared with 35,500 deaths registered with COVID-19 as the underlying cause during the same period.

Figure 5.4, below, displays the ratio between the number of registered COVID-19 deaths and the number of estimated excess deaths during the weeks between 26 March and 26 May 2020 for nine northern regions with the highest counts of registered COVID-19 deaths. Across all regions nationwide, the ratio of registered COVID-19 deaths to estimated excess mortality during this time varied from 58% to 95%. Figure 5.4 also displays the ratio between registered COVID-19 mortality and excess deaths in select northern regions of Italy. Across all regions, COVID-19 mortality reporting completeness varied widely by week, including weeks with registration proportions approaching both zero and one. The larger regions of Emilia-Romagna, Liguria, Lombardy, Piedmont, and Tuscany all display a similar trend of increasing registration completeness over time.

5.4.2 Excess deaths and sex

These findings can be aggregated across sexes and provinces to understand the age structure of excess mortality across Italy. Figure 5.5 plots weekly excess deaths across the five modelled age

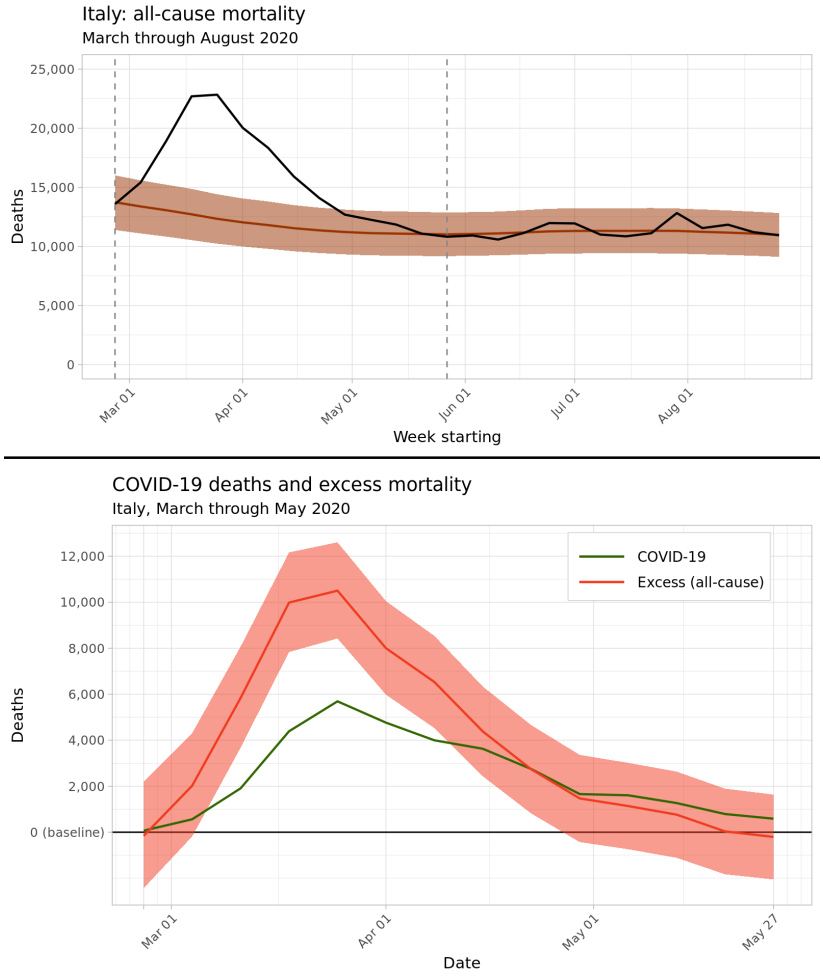


Figure 5.3: *Top* : Predicted baseline and observed deaths aggregated nationwide across Italy, March through August 2020. Bounds of the 95% uncertainty interval for predicted baseline deaths are shaded in light brown. *Bottom* : Estimated weekly excess deaths across Italy from March through May 2020, in red, compared with medically-certified deaths from COVID-19, in green.

categories, while Table 5.2 lists estimated excess deaths by age and sex grouping at the national level. Excess deaths were overwhelmingly concentrated in older age groups: of the estimated 53,200 excess deaths in Italy from March through May 2020, an estimated 51,600, or 97.0%, of these excess deaths occurred among adults aged 60 and above, while 36,400, or 68%, occurred in adults older than age 80. Among adults aged 90 and above, women experienced 11,000 (5,800 to 16,400) excess deaths, more than double the 4,400 (2,000 to 6,900) excess deaths among men in the same age group; this reflects the sex composition of the oldest age group, where women made up 72.6% of the over-90 population in January 2020.

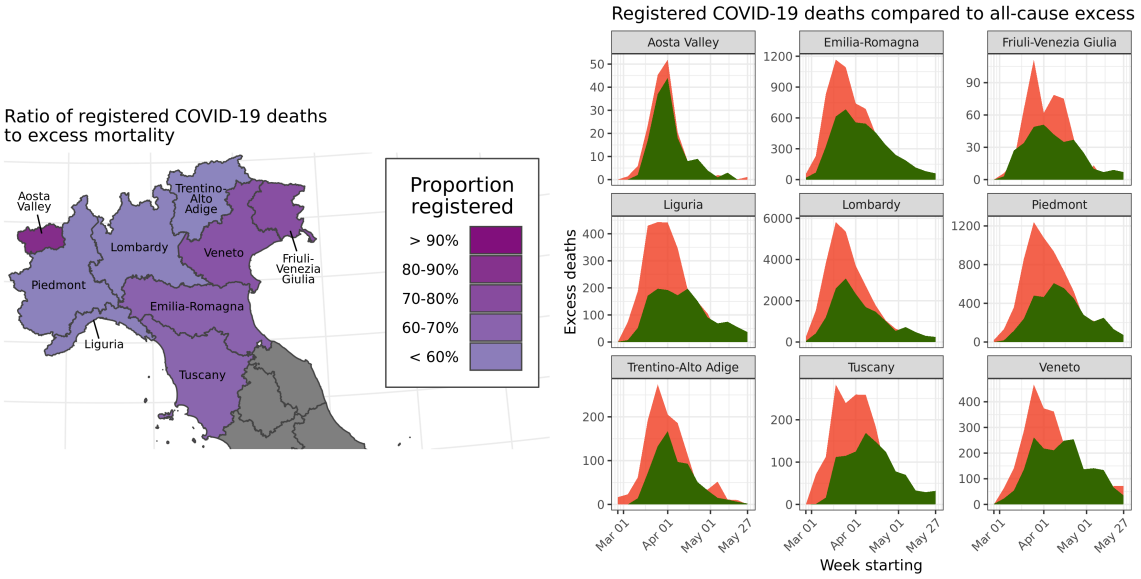


Figure 5.4: *Left* : Map showing the ratio between the number of registered COVID-19 deaths in the northern regions of Italy and the total excess deaths in these regions, 26 February through 26 May 2020. *Right* : Line charts showing the total estimated excess deaths, in red, and the number of registered COVID-19 deaths, in green, by region and week.

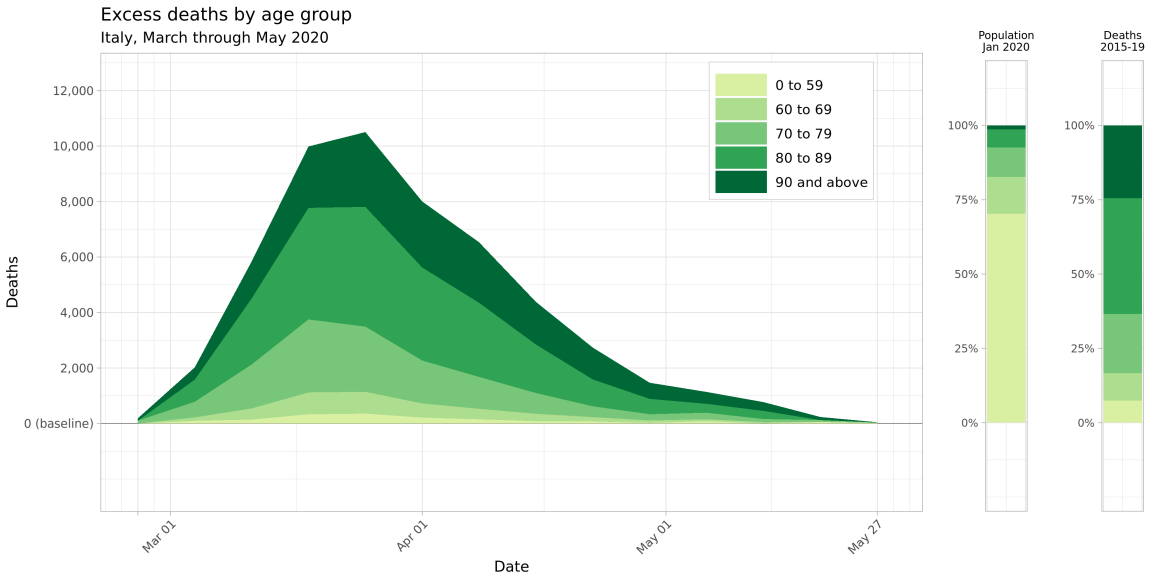


Figure 5.5: Mean estimates for weekly excess deaths across Italy by age group, March through May 2020. Two nationwide age group breakdowns are shown to the right for comparison: the national age distribution as of 1 January 2020, and the proportion of nationwide deaths that occurred in each age group during the years 2015 through 2019.

The outsize proportion of excess deaths observed within older age categories reflects two aspects of age-structured mortality across Italy during this period: baseline mortality is highest in older age

Table 5.2: Estimated excess mortality by age and sex grouping, aggregated to the national level for the weeks of 26 February through 26 May 2020. Nationally, all age-sex groupings experienced significantly elevated mortality compared to the expected baseline during this period.

Sex	Age	Excess deaths (95% UI)	Standardized Mortality Ratio (95% UI)	Population on 1 Jan 2020, millions (% total)
Males	0-59	1,081 (918 to 1,240)	1.16 (1.13 to 1.19)	21.37 (35.5%)
Females	0-59	517 (401 to 627)	1.13 (1.1 to 1.16)	21.00 (34.9%)
Males	60-69	2,600 (1,152 to 3,990)	1.29 (1.11 to 1.52)	3.55 (5.9%)
Females	60-69	947 (109 to 1,750)	1.19 (1.02 to 1.41)	3.87 (6.4%)
Males	70-79	7,703 (4,639 to 10,730)	1.42 (1.22 to 1.71)	2.75 (4.6%)
Females	70-79	3,991 (1,885 to 6,048)	1.33 (1.13 to 1.6)	3.25 (5.4%)
Males	80-89	10,649 (4,574 to 16,946)	1.36 (1.13 to 1.72)	1.45 (2.4%)
Females	80-89	10,308 (3,675 to 17,110)	1.32 (1.09 to 1.67)	2.20 (3.7%)
Males	90	4,405 (1,979 to 6,879)	1.36 (1.14 to 1.71)	0.22 (0.4%)
Females	90	11,043 (5,784 to 16,432)	1.41 (1.18 to 1.76)	0.58 (1.0%)

groups even in normal years, and these age groups also experienced a larger relative increase in mortality during the study period. Figure 5.6 maps the standardised mortality ratio, or the ratio between observed and baseline mortality rates, by age group and province across the entire 13-week period from 26 February through 26 May 2020. Note that the standardised mortality ratio uniformly increases across the northern provinces of Italy when moving from the 0-59 age group to the 60-69 and 70-79 age groups, and remain heightened in the oldest age groups, with some provinces in Lombardy and Emilia Romagna experiencing over three times the expected baseline mortality in older age groups. The greater increase in excess mortality among older age groups can also be expressed by comparing the proportion of excess mortality with each age group with that age group's share of all-cause mortality in the years 2015 through 2019. Adults over age 60 accounted for 92.6% of all deaths in the years 2015-2019 (2.99 million deaths out of 3.23 million total), but accounted for 97.0% of all excess deaths (51,600 of 53,200), a significantly higher proportion, in March through May of 2020.

Similarly, adults over age 80 experienced 63.4% of all deaths in previous years (2.04 million deaths out of 3.23 million), but 68.4% of excess mortality in these 13 weeks (36,400 of 53,200 excess deaths).

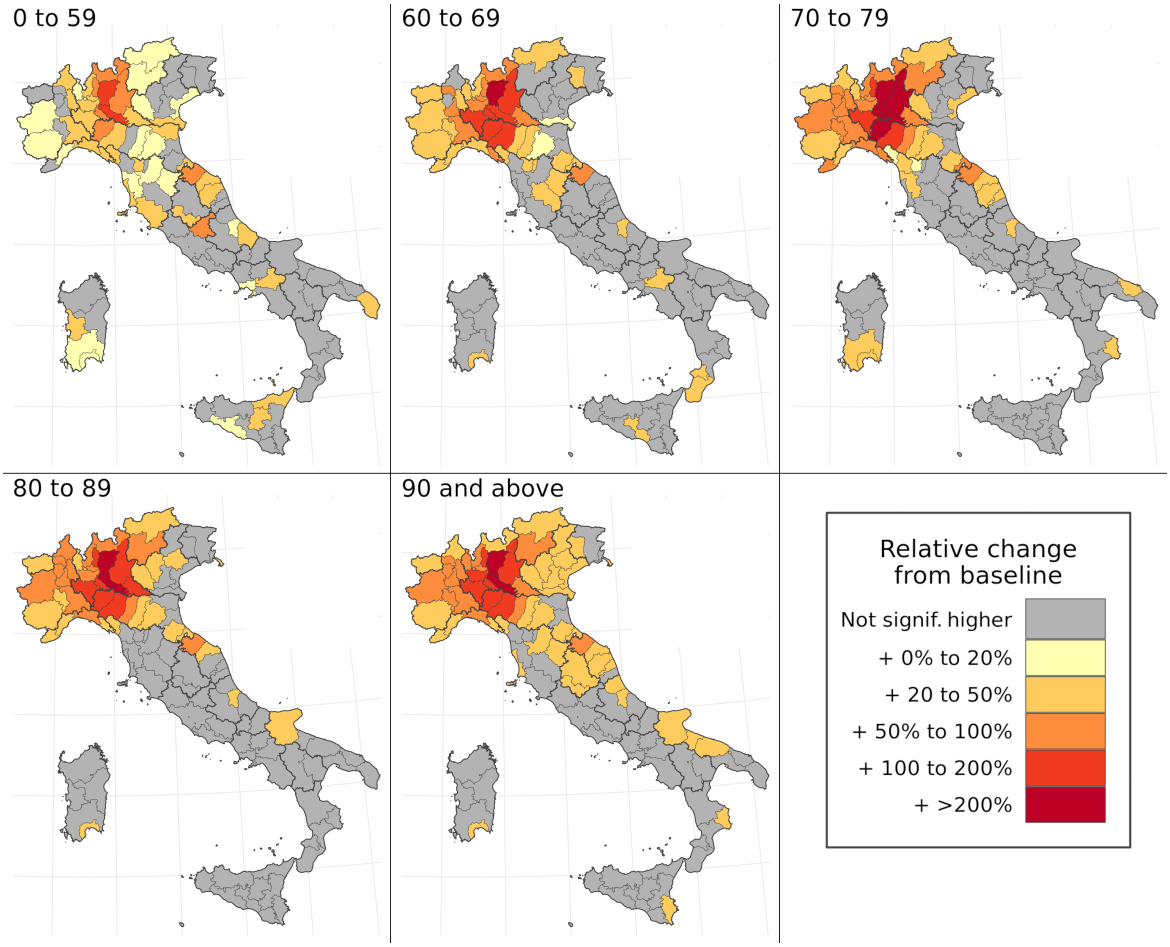


Figure 5.6: Map of standardised mortality ratios by age group and province over the entire period of March through May 2020. In these figures, significance is defined as falling outside of the 95% uncertainty interval for baseline mortality for a given age group and province during this period.

5.4.3 Spatial concentration

This model is also able to characterise spatial variation in excess mortality during the first wave of the COVID-19 epidemic in Italy. From March through May 2020, just three provinces—Milan, Bergamo, and Brescia, all in the Lombardy region—accounted for 32% of all excess deaths in Italy, with just 9% of the country's total population. By including just four other provinces across the Piedmont, Emilia-Romagna, and Liguria regions in northern Italy, nearly one-half of all excess deaths are captured in

a region that makes up just 16% of Italy's population. Expanding further, the 22 provinces with the highest number of excess deaths account for three-quarters of all excess mortality, but just 30% of Italy's national population. Figure 5.7 shows the locations of these province groupings and each group's marginal contribution to the total excess mortality curve from March through May 2020.

5.4.4 Diverse experiences of the first wave

Comparing Figures 5.1, 5.6, and 5.7, both the timing and geographical distribution of excess mortality during March through May 2020 seem to follow the time pattern of registered COVID-19 mortality by region in that same period: both peaked during the week of 25 March and were concentrated in the northern regions of the country, particularly Lombardy. Figure 5.8 complicates this analysis by demonstrating two features of excess mortality that are apparent at the province level, but not the regional or national levels. The first feature identifies the geographic centre of the peak excess mortality during the COVID-19 first wave. While most news outlets and previous studies have identified Lombardy region as the centre of the COVID-19 outbreak, the top-right panel of Figure 5.8 suggests that the mortality rate increased most in provinces on the border between south-west Lombardy and north-west Emilia-Romagna, adjacent to the Piedmont and Liguria regions. This border effect suggests that by analysing excess mortality across the 107 provinces of Italy, national public health agencies can identify mortality hot spots that would be obscured at the regional level.

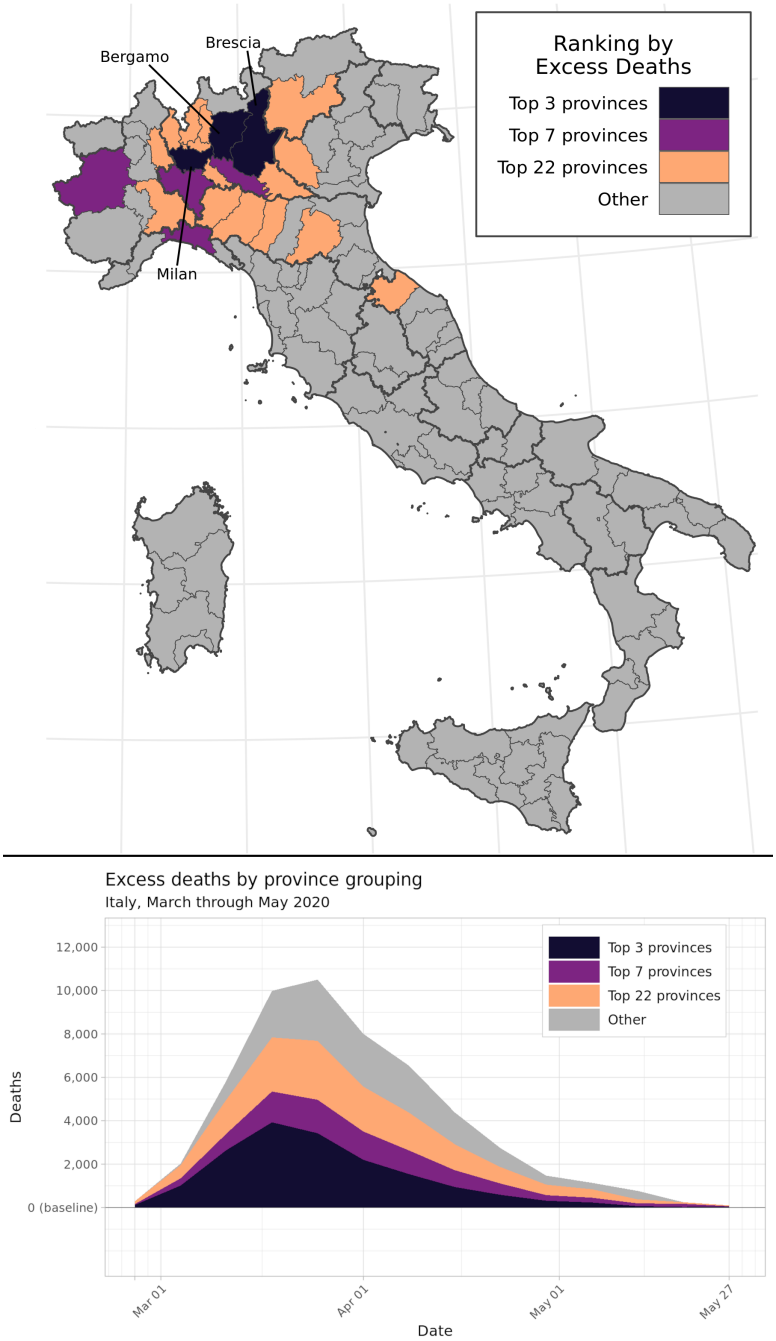


Figure 5.7: Concentration of excess mortality burden among the 3, 7, and 22 provinces with the greatest number of excess deaths, out of 107 provinces total. These province groupings respectively accounted for over 25%, 50%, and 75% of the national excess mortality burden between March and May 2020. The graph on the bottom left displays mean estimated excess deaths by province grouping and week during this time period.

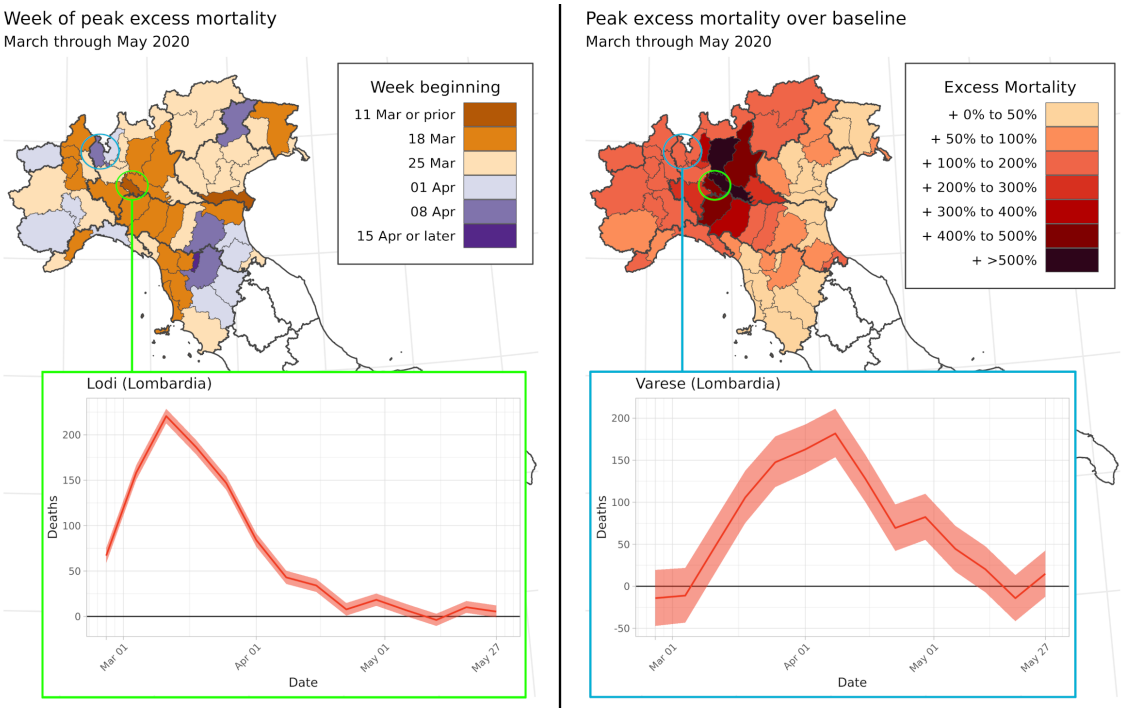


Figure 5.8: Excess deaths vary across provinces within the Lombardy region. *Top left* : Map of Northern Italy showing the week when province's excess mortality was estimated to peak during the first wave of COVID-19 in March through May 2020. *Top right* : Map of the peak weekly standardised mortality ratio (SMR) for each province across Northern Italy across all weeks in March through May 2020. *Bottom left* : Estimated excess deaths by week for Lodi province. *Bottom right* : Estimated excess deaths by week for Varese province.

A province-level analysis also demonstrates that within high-burden regions, neighbouring provinces experienced varying patterns of excess mortality during the first wave of COVID-19. The bottom two panels of Figure 5.8 show excess mortality curves for Lodi and Varese, two provinces that both share a border with Milan. Excess mortality in Lodi peaked on the week of March 11, with an SMR of over 6 times the baseline mortality rate; meanwhile, excess mortality in Varese peaked a month later on the week of 8 April, with mortality for that week peaking at over double the baseline rate.

5.5 Discussion

In this study, I explored whether an application of small-area methods to excess mortality analysis could identify previously unreported trends in the pattern of excess mortality across Italy during the first wave of the COVID-19 pandemic. The findings from this analysis suggest that a small-area

model yields estimates of excess mortality that are consistent with alternative calculation strategies at the national level, while offering new insights into the uneven distribution of excess mortality by age group, sex, province, and week across the country. Excess mortality estimates generated by this model suggest that a disproportionate majority of excess deaths occurred in adults age 60 and older, due to both the higher level of baseline mortality in these age groups and higher elevation of mortality above baseline during the first three months of the COVID-19 epidemic. This analysis also revealed a highly uneven spatial distribution of excess mortality: half of excess deaths were contained within just seven of Italy's 107 provinces, accounting for less than 16% of the population. While the general nationwide pattern of excess mortality reflected the timing and geographical concentration of registered COVID-19 deaths, regional analyses obscure meaningfully different excess mortality trends across neighbouring provinces within a region, highlighting the necessity of geospatial analysis as a component of the response to public health crises.

Previous analyses of excess mortality during the COVID-19 pandemic have compared registered COVID-19 deaths with excess deaths as a measure of cause-of-death misclassification during the pandemic.[70,193,214] My analysis found that between 26 March and 26 May 2020, 1 out of every 3 excess deaths nationally was not assigned to COVID-19. While this ratio is approximately consistent with previous analyses of excess mortality in Italy,[193] it conceals diversity across regions of the country and over time. Figure 5.4 demonstrates that COVID-19 death registration completeness varied nearly two-fold by region of the country; additionally, many of the regions with the greatest number of registered COVID-19 deaths experienced increasing registration accuracy over time, suggesting that medical professionals may have been increasingly likely to assign COVID-19 as an underlying cause of respiratory deaths as the first wave of COVID-19 progressed across northern Italy. Wide variation in diagnostic accuracy over time and across the country challenges the assumption of complete and accurate vital records that underpin many epidemiologic studies in high-income countries. Additional investigation is needed to understand the factors that may reduce completeness or diagnostic accuracy of death registration in these settings.

This study extends regression-based methods for estimating age-structured baseline mortality by incorporating location, year, and week structured random effects within a Bayesian hierarchical

framework.[200,212] This method was found to substantially outperform a simpler approach, in which average death counts across past years are used as the baseline, for predicting weekly excess mortality. It appears that a structured space-time approach stabilises stochastic variation across relatively small death counts by province and week, producing a smoother mortality risk surface while still accounting for meaningful trends captured by covariates. Because this approach structures uncertainty in a way that allows for principled aggregation, the results can indicate high-risk subgroups by age or location, and identify local variation in excess mortality that might be masked at a less detailed level, without overstating the confidence of findings for individual sub-populations.

This method for measuring excess mortality also has several limitations that should be noted. Because the process for estimating baseline deaths is more complicated and requires additional inputs compared to a simpler averaging method, it is less accessible to a wide range of users. The model for estimating baseline mortality assumes the same relationship between each covariate and mortality across age groups. In reality, some covariates may have a differential effect by age—for example, temperature may have more of an impact on mortality in older age groups due to the greater prevalence of risk factors that inhibit the body's thermoregulatory response.[215,216] This limitation is partly addressed by the separate harmonic seasonality fits for each age group. This study is also limited to the set of covariates which can be estimated by province and year: other covariates that may be predictive of all-cause mortality, such as the prevalence of environmental and occupational risk factors, were excluded due to lack of availability at the province level. While the population groupings reported in this study could be divided into even more granular units, small-area analyses must also adhere to ethical requirements to protect the privacy rights of individuals.[217] The Istat mortality data underlying this analysis did not include descriptors for ethnicity, so it was not possible to assess potential inequality in baseline and excess mortality by ethnicity as part of this analysis. Finally, as described in the Introduction, findings from excess mortality analyses must be carefully interpreted due to the many possible sources for changing mortality which are not accounted for in the modelling strategy.

Additional mechanisms for public health monitoring are needed to catch future resurgences of COVID-19 and future epidemics. In the context of high-income countries such as Italy, where high-quality mortality data has been rapidly prepared and cleaned for public use, this approach to

small-area excess mortality analysis could be employed as a routine surveillance tool, allowing health officials to identify high-mortality subgroups in a population and to introduce intervention measures in a timely manner. This approach could also be applied in many countries across Latin America which maintain high-quality mortality registration systems,[84] but where capacity to diagnose cause-specific mortality from COVID-19 may have been limited early in the pandemic.[218] Combining this excess mortality data with cause-of-death information by province would also reveal new insights about the local drivers of excess mortality. This study provides a new avenue to convert excess mortality analysis into a tool for decision-making in public health.

5.6 References

2. Burstein, R. *et al.* Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature* **574,** 353–358. http://www.nature.com/articles/s41586-019-1545-0 (Oct. 2019).
43. Riebler, A. *et al.* An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research* **25,** 1145–1165. arXiv: 1601.01180 (2016).
57. Banerjee, S., Carlin, B. P. & Gelfand, A. E. *Hierarchical modeling and analysis for spatial data* 2nd Editio (CRC Press, 2014).
70. Weinberger, D. M. *et al.* Estimation of Excess Deaths Associated with the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Internal Medicine* **06520,** E1–E9 (2020).
71. Patil, A. P., Gething, P. W., Piel, F. B. & Hay, S. I. Bayesian geostatistics in health cartography: The perspective of malaria. *Trends in Parasitology* **27,** 246–253 (2011).
84. Mikkelsen, L. *et al.* A global assessment of civil registration and vital statistics systems: Monitoring data quality and progress. *The Lancet* **386,** 1395–1406. http://dx.doi.org/10.1016/S0140-6736(15)60171-4 (2015).
89. Dicker, D. *et al.* Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **392,** 1684–1735. https://linkinghub.elsevier.com/retrieve/pii/S0140673618318919 (Nov. 2018).
108. Wakefield, J. *et al.* Estimating under-five mortality in space and time in a developing world context. *Statistical Methods in Medical Research* **28,** 2614–2634 (2019).
115. Dickman, P. W., Sloggett, A., Hills, M. & Hakulinen, T. Regression models for relative survival. *Statistics in Medicine* **23,** 51–64 (2004).
118. Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H. & Bell, B. M. TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* **70.** arXiv: 1509.00660 (2016).
119. Thorson, J. T. & Kristensen, K. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research* **175,** 66–74 (2016).
120. R Core Team. *R: A Language and Environment for Statistical Computing* Vienna, Austria, 2018. https://www.r-project.org/.
181. Huang, H. C., Martinez, F., Mateu, J. & Montes, F. Model comparison and selection for stationary space-time models. *Computational Statistics and Data Analysis* **51,** 4577–4596 (2007).
190. Thacker, S. B. & Berkelman, R. L. Public Health Surveillance in the United States. *Epidemiologic Reviews* **10,** 164–190 (1988).

192. Sebastiani, G., Massa, M. & Riboli, E. Covid-19 epidemic in Italy: evolution, projections and impact of government measures. *European Journal of Epidemiology* **35,** 341–345. https://doi.org/10.1007/s10654-020-00631-6 (2020).
193. Alicandro, G., Remuzzi, G. & La Vecchia, C. Italy's first wave of the COVID-19 pandemic has ended: no excess mortality in May, 2020. *The Lancet* **396,** e27–e28. http://dx.doi.org/10.1016/S0140-6736(20)31865-1 (2020).
194. Institute for Health Metrics and Evaluation (IHME). *COVID-19 Mortality, Infection, Testing, Hospital Resource Use, and Social Distancing Projections* Seattle, WA, 2020. http://www.healthdata.org/covid/data-downloads.
195. La Maestra, S., Abbondandolo, A. & De Flora, S. Epidemiological trends of COVID-19 epidemic in Italy over March 2020: From 1,000 to 100,000 cases. *Journal of Medical Virology* **92,** 1956–1961 (2020).
196. Albitar, O., Ballouze, R., Ooi, J. P. & Sheikh Ghadzi, S. M. Risk factors for mortality among COVID-19 patients. *Diabetes Research and Clinical Practice* **166,** 108293. https://doi.org/10.1016/j.diabres.2020.108293 (2020).
197. Vestergaard, L. S. *et al.* Excess all-cause mortality during the COVID-19 pandemic in Europe - preliminary pooled estimates from the EuroMOMO network, March to April 2020. *Eurosurveillance* **25.** http://dx.doi.org/10.2807/1560-7917.ES.2020.25.26.2001214 (2020).
198. Pasquariello, P. & Stranges, S. Excess mortality from COVID-19: a commentary on the Italian experience. *International Journal of Public Health* **65,** 529–531. https://doi.org/10.1007/s00038-020-01399-y (2020).
199. Mannucci, E., Nreu, B. & Monami, M. Factors associated with increased all-cause mortality during the COVID-19 pandemic in Italy. *International Journal of Infectious Diseases* **98,** 121–124. https://doi.org/10.1016/j.ijid.2020.06.077 (2020).
200. Serfling, R. E. Methods for Current Statistical Analysis of Excess Pneumonia-Influenza Deaths. *Public Health Reports (1896-1970)* **78,** 494 (1963).
201. Noufaily, A. *et al.* An improved algorithm for outbreak detection in multiple surveillance systems. *Statistics in Medicine* **32,** 1206–1222 (2013).
202. Michelozzi, P. *et al.* Temporal dynamics in total excess mortality and COVID-19 deaths in Italian cities. *BMC Public Health* **20,** 1238. https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-020-09335-8 (Dec. 2020).
203. Lambert, P. C., Smith, L. K., Jones, D. R. & Botha, J. L. Additive and multiplicative covariate regression models for relative survival incorporating fractional polynomials for time-dependent effects. *Statistics in Medicine* **24,** 3871–3885 (2005).
204. Wu, J., McCann, A., Katz, J., Peltier, E. & Singh, K. D. *Tracking the True Toll of the Coronavirus Outbreak* 2020. https://nyti.ms/2WbTKbS (2020).
205. The Economist. *COVID-19 Data: Tracking COVID-19 excess deaths across countries* 2020. https://econ.st/3oNlNL0 (2020).
206. FT Visual & Data Journalism team. *FT Coronavirus Tracker* 2020. https://on.ft.com/37ie1Dj (2020).
207. U.S. National Center for Health Statistics. *Excess Deaths Associated with COVID-19* 2021. https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm (2021).
208. Italian National Institute of Statistics (Istat). *All-cause mortality by Italian municipality, January 2015 through August 2020 (release: 22 October 2020).* Rome, 2020. https://www.istat.it/it/archivio/240401.
209. Italian National Institute of Statistics (Istat). *Resident municipal population by age, sex and marital status* Rome, 2020. https://dati.istat.it.
210. Dowd, J. B. *et al.* Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences of the United States of America* **117,** 9696–9698 (2020).

211. Cox, D. R. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* **34,** 187–220 (1972).
212. Ederer, F., Axtell, L. M. & Cutler, S. J. The Relative Survival Rate: A Statistical Methodology. *National Cancer Institute Monographs* **6,** 101–121. https://pubmed.ncbi.nlm.nih.gov/13889176/ (1961).
213. Thorson, J. T. & Barnett, L. A. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science* **74,** 1311–1321 (2017).
214. Woolf, S. H., Chapman, D. A., Sabo, R. T., Weinberger, D. M. & Hill, L. Excess Deaths From COVID-19 and Other Causes, March-April 2020. *JAMA* **324,** 510. https://jamanetwork.com/journals/jama/fullarticle/2768086 (Aug. 2020).
215. Yu, W., Vaneckova, P., Mengersen, K., Pan, X. & Tong, S. Is the association between temperature and mortality modified by age, gender and socio-economic status? *Science of the Total Environment* **408,** 3513–3518. http://dx.doi.org/10.1016/j.scitotenv.2010.04.058 (2010).
216. Stafoggia, M. *et al.* Factors affecting in-hospital heat-related mortality: A multi-city case-crossover analysis. *Journal of Epidemiology and Community Health* **62,** 209–215 (2008).
217. Bayer, R. & Fairchild, A. L. Surveillance and privacy. *Science* **290,** 1898–1899. https://www.sciencemag.org/lookup/doi/10.1126/science.290.5498.1898 (Dec. 2000).
218. León Cabrera, J. M. & Kurmanaev, A. *Ecuador's Death Toll During Outbreak Is Among the Worst in the World* New York, Apr. 2020. https://nyti.ms/2WbTKbS.

Chapter 6

Discussion

This study develops a class of methods to estimate health outcomes at the local level using health surveillance data. To do so, I extend traditional spatial statistical modelling approaches to integrate health surveillance with survey data sources, compensating for the limitations of each data type in the process. In four national case studies, I tailor this framework to accommodate country-specific data gaps, disease contexts, and programmatic needs. Because the resulting models identify local variation in surveillance completeness across a country, they can be used to correct raw surveillance data among small sub-populations, increasing the utility of health surveillance data for public health policy. Data completeness is also a measure of health system performance in its own right, and these model estimates can be further applied to target future health surveillance improvements nationwide. Previous research has assumed that health surveillance data in high-income countries is complete; I demonstrate that surveillance gaps remain even in high-income contexts, and the methods presented in this thesis can be applied to understand and address these gaps.

This discussion chapter summarises key findings from each chapter of the thesis, draws connections between the methods used in each country context, and links this research to active debates in global health. I also discuss the strengths and limitations of this research, and describe considerations for communicating this class of models in a policy-oriented context. In light of these considerations, I suggest future research and applications that can further improve our understanding of health surveillance data.

6.1 Chapter summary

The thesis begins with a description of data types commonly used to estimate disease burden and their known limitations for estimating health at the local level. Figure 6.1 summarises the country case studies presented in the subsequent four chapters.

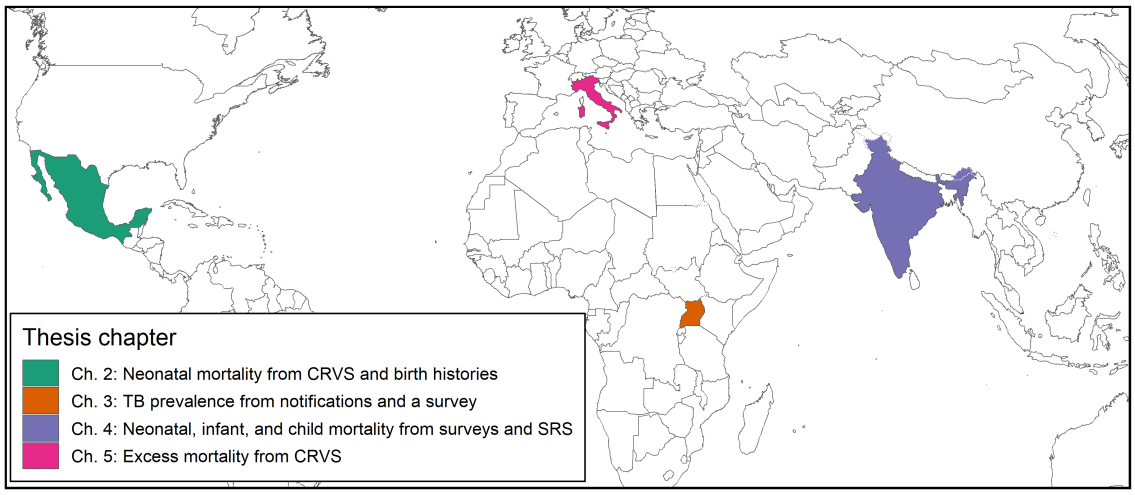


Figure 6.1: World map summarising the settings and themes of the four case studies presented in this thesis.

Chapter 2 identifies the core issue of latent, spatially-varying bias as a barrier to utilisation of health surveillance data at the local level. This chapter describes past approaches used to estimate bias in Civil Registration and Vital Statistics (CRVS) systems[85,92,97,100,101], as well as the limitations of these approaches for spatial analysis.[102,103,106] I then introduce a new approach to jointly estimate age-specific mortality and CRVS incompleteness at the district level. This approach extends previous research on Bayesian small area estimation for mortality[106,108] by estimating the ratio between survey-based and routinely-collected data sources, subject to priors about the relationship between these sources. I implement this model to estimate the neonatal mortality rate (NMR) by municipality in Mexico. I apply context-specific priors for CRVS bias by municipality based on previous research that demonstrated birth and death under-reporting in indigenous and socially-marginalised communities across Mexico.[16,95,112,113] Compared to a spatial model that treats all CRVS data as unbiased, this joint model alters municipality infant mortality rankings in the southernmost states

of Mexico, suggesting that funding for municipality health programs may be distorted without accounting for bias in CRVS data. The joint model also identifies locations where a large proportion of births or neonatal deaths have been omitted from CRVS records, offering a new tool to improve the quality of health statistics in these municipalities.

While strong priors are appropriate to a data context where health surveillance is nearly complete nationwide, the same class of model can be applied in an infectious disease context where case notification completeness is known to be largely incomplete and variable across the country. In Chapter 3, I develop a similar model in the context of global tuberculosis (TB) control, where disaggregated estimates of TB prevalence and incidence are highly desirable for national TB control programs as a tool for targeting interventions and identifying treatment gaps.[136] I explore the TB data context in Uganda, which is representative of the situation in many high-burden, low-income countries: while TB case notifications from routine diagnosis are collected and reported annually by the National TB and Leprosy Control Programme (NTLP), these notifications are known to vary in completeness due to gaps in the reporting cascade, making unadjusted notifications an inappropriate proxy for subnational variation in TB incidence.[26,139] A national TB prevalence survey conducted in 2014-2015 does not suffer from the same under-reporting bias, but this survey was not powered for subnational estimation.[145] I present a spatial model that combines these data types to generate small-area estimates of both TB prevalence and case notification completeness by district. The resulting model reports subnational prevalence with greater precision than a model based on survey data alone, allowing for more precise tracking of high-burden regions within the country. Additionally, the Uganda NTLP can use estimates of notification completeness at the district level to track progress towards their strategic goal of complete TB case reporting nationwide.[148]

While reporting completeness is a crucial measure of health data quality, routine surveillance data must meet several other criteria before it can be useful for policymaking at the local level. The United Nations *Principles and Recommendations for a Vital Statistics System* proposes that CRVS (and other health surveillance data) should also meet stringent standards for timeliness, accuracy in reporting outcomes, and confidentiality.[14] In addition to their utility for policy, CRVS birth and death certification enable legal rights for covered individuals.[13,219] In Chapter 4, I evaluate three

overlapping mortality surveillance programs in India against these standards of quality and utility, as well as their alignment with the goals of India's National Health Plan (NHP) 2017. The household survey series conducted by the Indian government have an unparalleled sample size worldwide, allowing for disaggregated analysis of neonatal, infant, and child mortality using a novel space-time-age modelling framework.[34,179] This analysis demonstrates the value of spatial modelling to monitor progress towards NHP child survival goals; however, it also reveals the limitations of survey-based spatial models even in highly data-rich contexts. Because retrospective survey data on child mortality is most sparse in later years, uncertainty is greatest in the most recent year of analysis and in future projections. India's Sample Registration System, the second surveillance program for child health, is collected annually and includes larger sample sizes, making it a more timely and precise tool for health policy; also, unlike household surveys, the SRS collects data on mortality across adult age groups.[175] Previous analyses have underlined the variable completeness of the SRS at the state level.[87,176] In this chapter, I demonstrate how SRS estimates can be adjusted for incompleteness at the district grouping level through comparison to survey-based estimates. If the SRS reported estimated mortality across more disaggregated sub-populations of India, these would enable an equity-focused evaluation of India's progress towards NHP child mortality goals and allow researchers to quickly identify emerging health crises. The third program, India's Civil Registration System (CRS), is the only system designed to provide birth and death certification for all Indian citizens.[173,178] Ultimately, estimates from India's household survey series and SRS should be used to evaluate progress towards complete certification in the CRS, as this is the only program that can guarantee the rights associated with birth and death registration.[177] By comparing subnational infant mortality estimates across data sources, I identify regions where SRS data needs to be strengthened before it can be adopted as the primary data source for mortality surveillance nationwide.

Previous geospatial and small-area research has drawn a distinction between incomplete and complete data systems. Research in low- and middle-income countries (LMICs) has often assumed incompleteness in health surveillance systems, requiring epidemiologists to develop corrections for reporting or otherwise limit their analyses.[37,82,139] Conversely, in high-income countries, spatial analyses often treat health surveillance as complete, needing correction only for stochastic fluctuations

associated with small sample sizes.[68,220] In Chapter 5, I challenge this assumption by investigating spatial variation in excess mortality across Italy during the first months of the COVID-19 pandemic. I use all-cause mortality data from 2015-2019 to estimate baseline mortality by age group, province, and week of the year from March through May 2020; I then compare estimated excess deaths across these groupings to recorded deaths assigned to COVID-19. I find that medically-certified cause of death reporting for COVID-19 missed 30% of excess mortality nationwide, and 42% of excess mortality burden in the provinces where excess mortality was highest. This finding challenges the assumption that health surveillance systems in high-income countries can be interpreted as uniformly complete and accurate. Epidemiologists in high-income countries should adopt excess mortality analyses as well as completeness audits developed by middle-income countries to better understand the quality of their data sources in the context of an emerging infectious disease.

While these investigations explore diverse health contexts in four different countries, they all use spatial analysis as a lens to understand the middle ground between unbiased and unusable health data sources. Subnational analyses can provide insight into health data bias because these biases may be related to local health system or socioeconomic differences that inhibit data quality at the point of collection. Across the investigations in this thesis, I demonstrate how biases in health surveillance data can be identified and corrected by (1) comparing multiple health data sources with different structural biases, and (2) directly estimating health surveillance bias in spatially-structured statistical models. In the contexts of Uganda and India, I demonstrate how data sources with known biases can still contribute to robust subnational estimates of health burden. The approach applied in these chapters could be applied more widely to integrate incomplete data sources into health decision-making processes and to advocate for health surveillance strengthening. In the contexts of Mexico and Italy, I use the same approach to highlight local biases in CRVS data sources that are often treated as unbiased and complete in epidemiological studies. The approaches used in these chapters could be applied to identify areas for improvement in health surveillance systems that are nearly complete. Across all chapters, spatial models are deployed not as replacements for health surveillance data, but rather as tools for understanding and strengthening that data.

6.2 Methodological strengths

Small area estimation and geostatistical methods can measure mortality and disease burden at the scale that is most programmatically relevant for health system management. In recent years, national and international health policy documents have stressed the importance of equity in delivering and sustaining health.[148,168,221] Measurement facilitates management: tracking health status across social and geographical groupings can identify inequity, prompting responsive health policy and reform.[8,90] Small area estimation can also offer metrics for understanding program effectiveness at the same scale where money is disbursed and programs implemented, revealing patterns that might be obscured at the national level. Because these models are designed to accommodate data with small sample sizes,[11,108] they can be extended to measure differences across other dimensions such as age group, as demonstrated in Chapters 4 and 5.

Country-specific models of health can be parametrised so that model terms directly reflect quantities of programmatic interest, as demonstrated by the case studies in this thesis. In Mexico, both birth and neonatal death reporting are estimated to be incomplete among some municipalities,[16,112] so the geostatistical model was implemented with a strictly-positive bias term to show the relative completeness between these two data sources. In Uganda, incompleteness in case reporting was a key metric for the NTLP,[135] and was expected to outweigh bias in modelled estimates of population denominators in most districts: to best accommodate program needs, under-reporting was directly estimated as a quantity varying between 0 and 1 by district, allowing for the estimation of notification completeness as well as change over time. In India, a model was implemented to track mortality across age groups to track progress towards separate goals for neonatal, infant, and child survival. Within this framework, priors for each country-specific model incorporated expert knowledge and past evidence on the distribution of disease aetiology and surveillance reporting across the country. The flexibility of these models allows them to be adapted for country-specific needs, increasing interpretability and utility for developing policy.

In LMICs, small-area modelling methods can supplement existing approaches that have been used to estimate health surveillance data completeness. Local audits have been developed to esti-

mate the completeness of CRVS data and infectious disease notifications;[16,92,93,104,105] While these studies offer a high-quality estimate of completeness in a small number of communities, they are prohibitively expensive to implement nationwide. Completeness audits can easily be integrated into the framework described in this thesis to offer a more representative estimate of completeness nationwide that also incorporates survey and CRVS or notification data. This data synthesis approach has advantages over the previous capture-recapture method, which makes unrealistically strong assumptions about independence between data sources.[102,103]

Finally, the Bayesian hierarchical modelling approach deployed in all of these models propagates uncertainty from input data and parameter fits in a principled way, allowing for nuanced analysis about the confidence of results.[41] By saving the structured uncertainty in the resulting model fits, these estimates can be used as principled prior estimates for future analyses. The confidence of district-level estimates can also be visualised and evaluated against programmatically-relevant goals: for example, the uncertainty "draws" generated by this class of models can be compared to a disease reduction target to determine which districts fall below a target threshold in 95% of potential model realisations.[71]

6.3 Methodological limitations

Despite their flexibility, this class of semi-parametric estimation techniques relies on assumptions that may bias model results. All models drew predictive power from space-time covariates that, based on past evidence, were expected to correlate to disease burden or reporting completeness. If the parametric relationships between covariates and these outcomes are incorrect, it may skew results in areas with extreme covariate values or little observed data. Each model also included spatial latent surfaces, which formalise the assumption that unobserved predictors are likely to vary according to an underlying spatial structure: in other words, two neighbouring observations in space are more likely to be similar than two distant observations, conditional on all observed fixed effects. This assumption, while reasonable for many ecological and disease phenomena,[11] can also over-smooth discontinuities across space, under-stating the true inequality between groups. Furthermore, these models are sensitive to priors for space-time variation.[108] If strong priors are inaccurate, they can

bias the model in the direction of existing expert opinion, potentially obscuring or misrepresenting important patterns in the data. For this reason, priors for highly uncertain quantities, such as those used to set reasonable bounds of CRVS bias in Chapter 2, should be thoughtfully considered and closely inspected before a model is implemented.

The results presented in this thesis are also subject to data limitations that are difficult to correct for in a statistical model. Although the models presented in this thesis correct for bias present by data type and by survey, bias can also vary within a survey: for example, one survey data collection team may rephrase a question in a way that favours one response over another. Because the outcomes estimated in this thesis are relatively rare, these types of systematic biases in survey data would be difficult to distinguish from stochastic noise or true variation in an outcome. Covariates used to predict health burden or surveillance completeness may themselves be uncertain: the covariate measuring travel time to health facilities, for example, is sensitive to health facilities that may be missing from master facility lists and on-line mapping databases.[149] Finally, the analytical frameworks described in this thesis require obtaining data maintained by multiple institutions, including national ministries of health, country disease control programs, and the maintainers of international survey programs such as Demographic and Health Survey Program or the UNICEF Multiple Indicator Cluster Surveys. Compiling data from these diverse sources presents a challenge for any public health analyst, potentially reducing the timeliness of model results.

Translating model uncertainty into a policy setting can be challenging. Many policy processes are not developed with uncertainty in mind, and the prospect of simply interpreting summary estimates without uncertainty may be tempting from a decision-making perspective. However, reporting only mean estimates can introduce spurious trends in the results that are not supported when uncertainty is considered. In fact, propagating uncertainty is arguably one of the most useful functions of Bayesian hierarchical models, and should be highlighted as part of the decision-making process. New communication strategies are needed to intuitively demonstrate variability across many possible realisations of the spatial models presented in this thesis.[71]

6.4 Future directions

In light of the methodological strengths and limitations described above, this final section discusses how the findings of this thesis can be applied to align health surveillance data, modelling, and policymaking.

The surveillance-based spatial models presented in this thesis offer a new avenue for integrating relevant findings from health surveillance data into policymaking. While this thesis has described approaches to improve the completeness and accuracy of health surveillance data, models must also be updated with new data in a timely manner to maximise their usefulness for policy. The pace of integrating new data and methods depends on the policy context. For example, in high-burden settings for tuberculosis, many National TB Control Programs operate on an annual cycle for disbursing funding and evaluating program effectiveness; in this situation, new case notifications may be integrated into estimates of TB burden on an annual or quarterly basis. Conversely, excess mortality surveillance must be updated as often as possible to allow public health programs to respond to emerging crises. For this use case, a weekly reporting structure akin to the European Mortality Monitoring Activity (EuroMoMo) may be most appropriate. Near-continuous model updates require additional data engineering to automate the integration of new mortality data into regularly-updated statistical models. Automated dashboards and other interactive visualisations can also facilitate the uptake of surveillance-based disease maps as policy tools: importing modelled small area estimates into existing health information tools, such as DHIS2 dashboards in many low-income countries, should be explored.

All results presented in this thesis have included summaries of the uncertainty surrounding small area estimates. Knowledge about the relative certainty of modelled estimates can inform cautious and data-driven policy; moreover, public health agencies can incorporate knowledge about relative uncertainty in disease burden or surveillance completeness across a country to target future data gathering activities. For example, in Latin American countries with moderately complete CRVS, birth and mortality registration audits could be planned in those districts where estimates of reporting

completeness are the most uncertain. Because tuberculosis prevalence surveys are conducted infrequently in high-burden countries, the surveys cluster sampling design can influence knowledge about the disease distribution for years to come; future survey designs could consider relative uncertainty in modelled prevalence in order to place survey clusters in a configuration that maximises precision. Across a variety of diseases and mortality contexts, targeted surveys in districts with the greatest uncertainty could improve the certainty of models, increasing their value as a decision-making tool.

In Chapters 3 and 5 of this thesis, I expanded concepts from space-time modelling to estimate variation across the additional dimension of age. Future research should expand further on this concept, disaggregating data and estimates across relevant social categories to better understand health inequality over space and time. Social research has extensively explored how health disparities are expressed over space; by formalising this geographical knowledge into bespoke spatial statistical models, public health researchers can reveal the scope of health inequalities at the local level.

6.5 Conclusions

This thesis redefines the important role modelling plays in countries that are developing health surveillance systems. CRVS advocates in public health have previously argued that models have a deleterious crowding-out effect on CRVS improvement efforts: in this conception, policy-makers interpret models as a superior alternative to imperfect health surveillance data, reducing the momentum to improve CRVS quality.[13,73] In this thesis, I have described a set of models that are explicitly designed to integrate health surveillance data into health decision-making processes at the country and local levels. These models provide a new feedback mechanism by which improved health surveillance leads to more certain model estimates and better health policy locally, strengthening the pragmatic argument for investing in health data systems. Additionally, these models can be designed to directly measure reporting completeness, offering a new tool for tracking surveillance completeness over a country. Throughout this thesis, I have emphasised that health surveillance systems are more than a data source: by legally documenting illness and vital events, they guarantee legal and civil rights as well as appropriate health care to a country's citizens.[9]

Let me conclude this thesis with three takeaways for those who maintain and interpret health surveillance data. First, for health policymakers: develop heuristics that are keenly sensitive to local variation in health conditions and data quality. National statistics can obscure important inequalities in health outcomes, and maps that fail to account for data biases may understate regional health differences. Second, for national statistical agencies: diversify data sources for key health outcomes such as mortality and notifiable infectious diseases. Local data audits, originally developed to assess vital statistics completeness in Latin America,[16,92,104] could be adapted to check data quality across many countries and health contexts. Third, for health researchers and statisticians: take advantage of spatial relationships as well as expert consensus to better understand the generative process underlying health surveillance data. Incorporating both spatial dependency and qualitative knowledge into statistical models can offer new insights about health outcomes and yield more consensus-based tools for health policy.

High-quality health surveillance is the necessary foundation for long-lasting, sustainable improvements in global public health. As the primary data tool for a national health system, part of a health surveillance system's utility rests on its capacity to produce high-quality, timely, and disaggregated estimates of health status across a country's sub-populations. This thesis strengthens the link between developing national health surveillance systems and improved health outcomes, leading to more equitable policies and a strong basis for the universal right to health.

6.6 References

8. Roberts, M. J., Hsiao, W., Berman, P. & Reich, M. R. *Getting health reform right: a guide to improving performance and equity* (2008).
9. AbouZahr, C. *et al.* Towards universal civil registration and vital statistics systems: The time is now. *The Lancet* **386,** 1407–1418. http://dx.doi.org/10.1016/S0140-6736(15)60170-2 (2015).
11. Diggle, P. J. & Giorgi, E. Model-based geostatistics for prevalence mapping in low-resource settings. *Journal of the American Statistical Association* **111,** 1096–1120. arXiv: 1505.06891. http://dx.doi.org/10.1080/01621459.2015.1123158 (2016).
13. Setel, P. W. *et al.* A scandal of invisibility: making everyone count by counting everyone. *Lancet* **370,** 1569–1577 (2007).
14. United Nations Statistics Division. *Principles and Recommendations for a Vital Statistics System Revision 3* **19.** http://unstats.un.org/unsd/Demographic/standmeth/principles/M19Rev3en.pdf (2014).
16. Hernández, B. *et al.* Subregistro de defunciones de menores y certificación de nacimiento en una muestra representativa de los 101 municipios con más bajo índice de desarrollo humano en México. *Salud Pública de México* **54,** 393–400 (2012).

26. Rood, E. *et al.* A spatial analysis framework to monitor and accelerate progress towards SDG 3 to end TB in Bangladesh. *ISPRS International Journal of Geo-Information* **8,** 1–11 (2019).
34. Dandona, R., Pandey, A. & Dandona, L. A review of national health surveys in India. *Bulletin of the World Health Organization* **94,** 286–296A (2016).
37. Zeng, X. *et al.* Measuring the completeness of death registration in 2844 Chinese counties in 2018. *BMC medicine* **18,** 176 (2020).
41. McElreath, R. *Statistical Rethinking* (Taylor & Francis, Boca Raton, FL, 2016).
68. Papoila, A. L. *et al.* Stomach cancer incidence in Southern Portugal 1998-2006: A spatio-temporal analysis. *Biometrical Journal* **56,** 403–415 (2014).
71. Patil, A. P., Gething, P. W., Piel, F. B. & Hay, S. I. Bayesian geostatistics in health cartography: The perspective of malaria. *Trends in Parasitology* **27,** 246–253 (2011).
73. Tichenor, M. & Sridhar, D. Metric partnerships: Global burden of disease estimates within the World Bank, the World Health Organisation and the Institute for Health Metrics and Evaluation. *Wellcome Open Research* **4** (2020).
82. Adair, T. & Lopez, A. D. Estimating the completeness of death registration: An empirical method. *PLoS ONE* **13,** 1–19 (2018).
85. Murray, C. J., Rajaratnam, J. K., Marcus, J., Laakso, T. & Lopez, A. D. What can we conclude from death registration? Improved methods for evaluating completeness. *PLoS Medicine* **7** (2010).
87. Bhat, P. N. Completeness of India's sample registration system: An assessment using the general growth balance method. *Population Studies* **56,** 119–134 (2002).
90. Frenk, J. Bridging the divide: global lessons from evidence-based health policy in Mexico. *Lancet* **368,** 954–961 (2006).
92. De Frias, P. G., Szwarcwald, C. L., de Souza, P. R. B., da Silva de Almeida, W. & Lira, P. I. C. Correcting vital information: Estimating infant mortality, Brazil, 2000-2009. *Revista de Saude Publica* **47,** 1048–1058 (2013).
93. Szwarcwald, C. L., de Frias, P. G., deSouza Júnior, P. R. B., da Silva de Almeida, W. & de Morais Neto, O. L. Correction of vital statistics based on a proactive search of deaths and live births: Evidence from a study of the North and Northeast regions of Brazil. *Population Health Metrics* **12,** 1–10 (2014).
95. Ribotta, B. S., Acosta, L. M. S. & Bertone, C. L. Evaluaciones subnacionales de la cobertura de las estadísticas vitales. Estudios recientes en América Latina [Evaluations of the Vital Statistics Coverage at a Subnational Level. Recent Studies in Latin America]. *Revista Gerencia y Políticas de Salud* **18** (2019).
97. Chandra Sekar, C. & Deming, E. On a Method of Estimating Birth and Death Rates and the Extent of Registration. *Journal of the American Statistical Association* **44,** 101–115 (1949).
100. Yip, S. F. *et al.* Capture-recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology* **142,** 1047–1058 (1995).
101. Becker, S., Waheeb, Y., El-deeb, B., Khallaf, N. & Black, R. Estimating the Completeness of Under-5 Death Registration in Egypt. *Demograph* **33,** 329–339. https://www.jstor.org/stable/2061765 (1996).
102. Tilling, K. Capture-recapture methods - Useful or misleading? *International Journal of Epidemiology* **30,** 12–14 (2001).
103. Cormack, R. M. Problems with using capture-recapture in epidemiology: An example of a measles epidemic. *Journal of Clinical Epidemiology* **52,** 909–914 (1999).
104. De Almeida, W. D. S. *et al.* Captação de óbitos não informados ao ministério da saúde: Pesquisa de busca ativa de óbitos em municípios brasileiros [Capturing deaths not informed to the Ministry of Health: proactive search of deaths in Brazilian municipalities]. *Revista Brasileira de Epidemiologia* **20,** 200–211 (2017).
105. National Adminstrative Department of Statistics (DANE). *La Mortalidad Materna y Perinatal en Colombia en los albores del siglo XXI. Estimación del subregistro de nacimientos*

y defunciones y estimaciones ajustadas de nacimientos, mortalidad materna y perinatal por departamentos tech. rep. (2006), 46. http://docplayer.es/40107048-Estudio-la-mortalidad-materna-y-perinatal-en-colombia-en-los-albores-del-siglo-xxi.html.

106. Schmertmann, C. P. & Gonzaga, M. R. Bayesian Estimation of Age-Specific Mortality and Life Expectancy for Small Areas With Defective Vital Records. *Demography* **55,** 1363–1388 (2018).
108. Wakefield, J. *et al.* Estimating under-five mortality in space and time in a developing world context. *Statistical Methods in Medical Research* **28,** 2614–2634 (2019).
112. Enciso, G. F., Del Pilar Ochoa Torres, M. & Hernández, J. A. M. The subsystem of information on births. Case study of an indigenous region of Chiapas, Mexico. *Estudios Demográficos y Urbanos* **32,** 451–486 (2017).
113. Paulino, N. A., Vázquez, M. S. & Bolúmar, F. Indigenous language and inequitable maternal health care, Guatemala, Mexico, Peru and the Plurinational State of Bolivia. *Bulletin of the World Health Organization* **97,** 59–67. http://www.who.int/entity/bulletin/volumes/97/1/18-216184.pdf (Jan. 2019).
135. Uganda National Tuberculosis and Leprosy Programme. *Uganda National TB and Leprosy Program: July 2019-June 2020 report* tech. rep. October 2020 (Uganda Ministry of Health, Kampala, 2020), 1–72.
136. Glaziou, P. & Floyd, K. *Latest developments in WHO estimates of TB disease burden* tech. rep. (WHO Global Task Force on TB Impact Measurement, Glion, 2018), 1–13.
139. Shaweno, D. *et al.* Methods used in the spatial analysis of tuberculosis epidemiology: A systematic review. *BMC Medicine* **16,** 1–18 (2018).
145. Uganda Ministry of Health. *The Uganda National Tuberculosis Prevalence Survey, 2014-2015: survey report* tech. rep. (Uganda Ministry of Health, Kampala, 2015), 1–162. https://www.health.go.ug/cause/the-uganda-national-tuberculosis-prevalence-survey-2014-2015-survey-report/.
148. Uganda National Tuberculosis and Leprosy Programme. *National strategic plan for tuberculosis and leprosy control: 2020/21-2024/25* tech. rep. November 2017 (Uganda Ministry of Health, Kampala, 2020), 1–106.
149. Weiss, D. J. *et al.* Global maps of travel time to healthcare facilities. *Nature Medicine* **26,** 1835–1838 (2020).
168. Government of India Ministry of Health and Family Welfare. *National Health Policy 2017* tech. rep. (Government of India, New Delhi, 2017), 28. https://www.nhp.gov.in/nhpfiles/national_health_policy_2017.pdf.
173. Parliament of the Republic of India. *The Registration of Births and Deaths Act, 1969* New Delhi, 1969.
175. Office of the Registrar General & Census Commissioner. *Sample Registration System Statistical Report 2017* tech. rep. (Ministry of Home Affairs, Government of India, New Delhi, 2017), 337. http://www.censusindia.gov.in/vital_statistics/SRS_Report/9Chap%202%20-%202011.pdf.
176. Mahapatra, P. *An Overview of the Sample Registration System in India* in *Prince Mahidol Award Conference & Global Health Information Forum* (Institute of Health Systems, Bangkok, 2010), 1–13. http://www.ihs.org.in/IHS/PMahapatra-SRSInIndiaAnOverview.pdf.
177. Abouzahr, C. *et al.* The way forward. *Lancet Health Metrics Network, World Health Organization* **370,** 1791–99 (2007).
178. Mohanty, I. & Gebremedhin, T. A. Maternal autonomy and birth registration in India: Who gets counted? *PLoS ONE* **13,** 1–19 (2018).
179. Dandona, R. *et al.* Subnational mapping of under-5 and neonatal mortality trends in India: the Global Burden of Disease Study 2000-17. *The Lancet* **395,** 1640–1658 (2020).
219. Duryea, S., Olgiati, A. & Stone, L. *The Under-Registration of Births in Latin America* 2006.
220. Boing, A. F., Boing, A. C., Cordes, J., Kim, R. & Subramanian, S. V. Quantifying and explaining variation in life expectancy at census tract, county, and state levels in the United

States. *Proceedings of the National Academy of Sciences of the United States of America* **117,** 17688–17694 (2020).

221. Büyüm, A. M., Kenney, C., Koris, A., Mkumba, L. & Raveendran, Y. Decolonising global health: If not now, when? *BMJ Global Health* **5,** 1–4 (2020).

Appendix A

Appendix to Chapter Two

This section contains additional tables related to Chapter Two: Joint estimation of neonatal mortality and vital registration completeness across Mexico.

Marginalisation groupings

As described in Section 2.2.1, priors for VR bias were set by municipality based on an index for estimated social marginalisation. Lists of municipalities in the moderately marginalised grouping (N=437/2,441) and the severely marginalised grouping (N=152/2,441) are included below in Table A.1 and Table A.2, respectively. Values for the four indicators used to define the marginalisation groupings are listed for each municipality.

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators.

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
El Bosque	Chiapas	75.6	34.9	0	98.1
Chapultenango	Chiapas	88.6	30.8	0	93.7
Francisco Leon	Chiapas	80.3	35.6	0	84.3
Huixtan	Chiapas	81.0	35.2	0	96.9
Huitiupan	Chiapas	77.2	36.2	0	92.3
Ixhuatan	Chiapas	83.2	39.9	0	61.6
Pantepec	Chiapas	75.2	35.4	0	66.8
Rayon	Chiapas	82.1	38.6	0	82.2
Sabanilla	Chiapas	75.8	34.7	0	98.3
Tapalapa	Chiapas	88.5	39.6	0	98.9
Morelos	Chihuahua	83.1	30.2	0	51.3
Nonoava	Chihuahua	92.4	39.4	0	53.6
Tierra Blanca	Guanajuato	93.4	33.5	0	76.9
Cuetzala del Progreso	Guerrero	90.7	28.1	0	58.1

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Martir de Cuilapan	Guerrero	75.4	49.7	0	69.7
Tlalixtaquilla de Maldonado	Guerrero	82.6	38.2	0	60.6
Alfajayucan	Hidalgo	97.0	38.9	0	68.9
Calnali	Hidalgo	81.9	36.7	0	97.5
Cardonal	Hidalgo	96.1	35.7	0	96.5
Chilcuautla	Hidalgo	95.8	45.3	0	92.0
Huautla	Hidalgo	95.7	36.5	0	90.9
Huazalingo	Hidalgo	87.6	30.7	0	99.6
Jaltocan	Hidalgo	79.9	41.2	0	98.3
Nicolas Flores	Hidalgo	96.9	27.4	0	94.1
San Felipe Orizatlan	Hidalgo	83.3	36.6	0	87.8
San Salvador	Hidalgo	98.6	46.7	0	66.1
Tianguistengo	Hidalgo	82.5	27.9	0	63.2
Xochiatipan	Hidalgo	79.5	32.9	0	97.1
Yahualica	Hidalgo	81.1	31.9	0	90.3
Bolanos	Jalisco	79.3	28.1	0	64.0
Cuautitlan de Garcia Barragan	Jalisco	90.6	31.9	0	69.2
Tuxpan	Jalisco	97.0	48.7	0	57.5
Zapotitlan de Vadillo	Jalisco	98.1	44.6	0	84.7
Chapa de Mota	Mexico	94.7	44.5	0	55.5
Morelos	Mexico	95.6	41.0	0	76.8
Timilpan	Mexico	96.1	40.7	0	57.7
Charapan	Michoacan	83.6	46.2	0	95.3
Chilchota	Michoacan	89.4	48.8	0	87.3
Erongaricuaro	Michoacan	94.3	46.7	0	79.4
Nahuatzen	Michoacan	87.8	37.6	0	78.1
Tingambato	Michoacan	96.0	45.4	0	87.3
Tinguindin	Michoacan	94.4	47.1	0	55.4
Tzintzuntzan	Michoacan	95.0	46.7	0	82.0
Ziracuaretiro	Michoacan	96.1	47.8	0	53.3
Asuncion Cacalotepec	Oaxaca	87.6	37.2	0	100.0
Asuncion Cuyotepeji	Oaxaca	92.0	43.3	0	53.6
Asuncion Ocotlan	Oaxaca	87.0	44.9	0	94.7
Asuncion Tlacolulita	Oaxaca	98.3	34.0	0	89.5
Ayotzintepec	Oaxaca	91.3	31.1	0	91.8
Candelaria Loxicha	Oaxaca	77.6	36.7	0	94.1
Concepcion Papalo	Oaxaca	88.4	40.1	0	86.9
Cosoltepec	Oaxaca	97.1	44.2	0	59.3
Cuyamecalco Villa de Zaragoza	Oaxaca	75.6	37.3	0	78.4
Chiquihuitlan de Benito Juarez	Oaxaca	76.4	40.6	0	92.6
El Espinal	Oaxaca	97.9	48.0	0	75.3
Guevea de Humboldt	Oaxaca	82.9	29.0	0	92.7
Mesones Hidalgo	Oaxaca	80.2	28.1	0	66.1
Villa Hidalgo	Oaxaca	86.1	44.8	0	97.2
Magdalena Jaltepec	Oaxaca	96.6	25.4	0	64.4
Magdalena Mixtepec	Oaxaca	81.2	27.2	0	71.6
Magdalena Tequisistlan	Oaxaca	97.4	39.1	0	75.3
Magdalena Tlacotepec	Oaxaca	96.5	41.8	0	70.0
Magdalena Zahuatlan	Oaxaca	97.8	28.9	0	60.9

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Mazatlan Villa de Flores	Oaxaca	75.4	34.9	0	98.3
Mixistlan de la Reforma	Oaxaca	88.8	43.2	0	100.0
Natividad	Oaxaca	99.4	40.6	0	90.5
Pinotepa de Don Luis	Oaxaca	90.3	46.5	0	92.3
San Agustin Atenango	Oaxaca	89.0	29.9	0	92.0
San Agustin Chayuco	Oaxaca	92.2	38.4	0	77.3
San Agustin Loxicha	Oaxaca	75.6	32.0	0	97.7
San Andres Dinicuiti	Oaxaca	93.5	33.8	0	56.0
San Andres Ixtlahuaca	Oaxaca	95.9	36.3	0	59.1
San Andres Solaga	Oaxaca	97.2	39.8	0	98.6
San Andres Teotilalpam	Oaxaca	85.3	33.0	0	94.0
San Andres Yaa	Oaxaca	94.2	44.9	0	94.2
San Antonino El Alto	Oaxaca	91.2	30.4	0	91.1
San Baltazar Chichicapam	Oaxaca	92.8	33.8	0	91.8
San Baltazar Yatzachi El Bajo	Oaxaca	94.2	36.0	0	92.5
San Bartolome Loxicha	Oaxaca	84.0	25.7	0	95.5
San Bartolome Quialana	Oaxaca	83.5	36.5	0	90.5
San Bartolome Yucuane	Oaxaca	95.8	27.7	0	100.0
San Bartolome Zoogocho	Oaxaca	98.5	33.5	0	98.5
San Bartolo Soyaltepec	Oaxaca	96.2	34.4	0	88.5
San Bartolo Yautepec	Oaxaca	98.6	39.7	0	93.7
San Carlos Yautepec	Oaxaca	91.8	38.3	0	94.8
San Cristobal Lachirioag	Oaxaca	94.3	43.5	0	95.3
San Cristobal Suchixtlahuaca	Oaxaca	100.0	43.6	0	66.2
San Dionisio del Mar	Oaxaca	83.6	34.9	0	85.4
San Dionisio Ocotepec	Oaxaca	82.8	33.3	0	96.1
San Felipe Jalapa de Diaz	Oaxaca	78.0	36.3	0	98.4
San Felipe Tejalapam	Oaxaca	97.0	46.8	0	52.1
San Felipe Usila	Oaxaca	88.1	35.1	0	99.6
San Francisco Cajonos	Oaxaca	95.2	33.2	0	98.8
San Francisco Chapulapa	Oaxaca	78.7	30.4	0	53.6
San Francisco del Mar	Oaxaca	90.4	42.7	0	79.1
San Francisco Huehuetlan	Oaxaca	77.2	33.7	0	97.7
San Francisco Jaltepetongo	Oaxaca	97.1	33.4	0	88.1
San Francisco Logueche	Oaxaca	77.0	35.0	0	98.8
San Francisco Nuxano	Oaxaca	98.8	40.7	0	79.8
San Francisco Ozolotepec	Oaxaca	77.8	26.8	0	90.4
San Ildefonso Amatlan	Oaxaca	91.9	35.7	0	62.5
San Jeronimo Silacayoapilla	Oaxaca	97.5	44.1	0	59.3
San Jeronimo Taviche	Oaxaca	93.4	41.1	0	68.4
San Jeronimo Tecoatl	Oaxaca	79.2	41.3	0	94.1
San Jose Chiltepec	Oaxaca	90.9	44.4	0	84.6
San Jose Independencia	Oaxaca	85.1	26.6	0	99.2
San Juan Achiutla	Oaxaca	100.0	30.7	0	91.4
San Juan Bautista Atatlahuca	Oaxaca	93.6	36.4	0	58.4
San Juan Bautista Coixtlahuaca	Oaxaca	97.0	32.0	0	64.1
San Juan Bautista Suchitepec	Oaxaca	93.4	40.5	0	92.3
San Juan Bautista Tlacoatzintepec	Oaxaca	79.7	28.4	0	98.2
San Juan Coatzospam	Oaxaca	82.9	42.3	0	98.8

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
San Juan Colorado	Oaxaca	89.1	32.9	0	82.4
San Juan Comaltepec	Oaxaca	84.8	29.8	0	98.5
San Juan Chicomezuchil	Oaxaca	100.0	38.6	0	97.3
San Juan del Rio	Oaxaca	92.5	48.2	0	93.7
San Juan Evangelista Analco	Oaxaca	98.7	25.5	0	90.9
San Juan Guelavia	Oaxaca	93.6	39.0	0	90.0
San Juan Guichicovi	Oaxaca	77.3	43.6	0	83.6
San Juan Juquila Mixes	Oaxaca	79.9	27.7	0	99.6
San Juan Lachao	Oaxaca	82.9	37.3	0	79.0
San Juan Lajarcia	Oaxaca	97.1	34.6	0	50.6
San Juan Lalana	Oaxaca	76.0	33.7	0	87.8
San Juan de los Cues	Oaxaca	92.6	43.5	0	52.9
San Juan Mazatlan	Oaxaca	89.6	36.6	0	78.9
San Juan Mixtepec - Distr. 08 -	Oaxaca	88.2	31.3	0	98.0
San Juan Ozolotepec	Oaxaca	80.5	37.3	0	80.0
San Juan Quiahije	Oaxaca	78.3	28.0	0	99.0
San Juan Quiotepec	Oaxaca	85.2	28.3	0	98.1
San Juan Tabaa	Oaxaca	95.9	39.7	0	100.0
San Juan Teita	Oaxaca	90.5	37.3	0	98.5
San Juan Teitipac	Oaxaca	91.3	43.5	0	52.4
San Juan Tepeuxila	Oaxaca	92.0	39.5	0	90.2
San Juan Yatzona	Oaxaca	97.3	35.5	0	99.1
San Juan Yucuita	Oaxaca	97.7	36.5	0	50.6
San Lorenzo	Oaxaca	80.9	33.2	0	89.3
San Lorenzo Albarradas	Oaxaca	93.8	37.9	0	74.2
San Lorenzo Cuaunecuiltitla	Oaxaca	80.0	32.4	0	98.7
San Lorenzo Victoria	Oaxaca	95.0	32.4	0	79.9
San Lucas Ojitlan	Oaxaca	88.3	39.1	0	98.3
San Lucas Quiavini	Oaxaca	78.0	36.1	0	92.9
San Marcial Ozolotepec	Oaxaca	77.3	30.7	0	98.6
San Marcos Arteaga	Oaxaca	95.0	38.3	0	54.1
San Martin Huamelulpam	Oaxaca	98.9	44.7	0	81.9
San Martin Tilcajete	Oaxaca	98.3	43.7	0	63.8
San Mateo Cajonos	Oaxaca	85.7	26.5	0	98.1
Capulalpam de Mendez	Oaxaca	98.8	45.4	0	79.7
San Mateo del Mar	Oaxaca	84.5	42.2	0	99.8
San Mateo Yoloxochitlan	Oaxaca	85.0	40.3	0	97.6
San Mateo Etlatongo	Oaxaca	98.6	34.5	0	67.9
San Mateo Penasco	Oaxaca	77.1	32.6	0	96.5
San Melchor Betaza	Oaxaca	85.7	37.8	0	99.3
San Miguel Achiutla	Oaxaca	98.1	26.8	0	76.4
San Miguel Aloapam	Oaxaca	79.0	39.1	0	98.6
San Miguel Amatlan	Oaxaca	98.8	42.2	0	91.1
San Miguel Coatlan	Oaxaca	82.5	32.3	0	61.0
San Miguel Chimalapa	Oaxaca	88.2	36.7	0	92.5
San Miguel del Rio	Oaxaca	98.7	27.3	0	94.9
San Miguel El Grande	Oaxaca	97.0	35.1	0	96.1
San Miguel Peras	Oaxaca	90.7	29.8	0	81.4
San Miguel Suchixtepec	Oaxaca	89.3	39.3	0	96.0

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
San Miguel Tecomatlan	Oaxaca	94.8	37.5	0	70.7
San Miguel Tenango	Oaxaca	94.7	40.4	0	61.1
San Miguel Tlacamama	Oaxaca	91.7	46.4	0	58.8
San Miguel Tlacotepec	Oaxaca	89.7	31.3	0	87.6
San Miguel Tulancingo	Oaxaca	98.4	34.0	0	54.7
San Miguel Yotao	Oaxaca	93.8	49.9	0	100.0
San Pablo Macuiltianguis	Oaxaca	94.2	27.1	0	87.9
San Pablo Villa de Mitla	Oaxaca	94.5	49.2	0	70.7
San Pablo Yaganiza	Oaxaca	97.9	34.1	0	99.3
San Pedro Amuzgos	Oaxaca	83.4	43.4	0	93.2
San Pedro Atoyac	Oaxaca	82.2	30.4	0	83.2
San Pedro Cajonos	Oaxaca	97.9	44.2	0	98.6
San Pedro Coxcaltepec Cantaros	Oaxaca	97.5	31.2	0	83.6
San Pedro Comitancillo	Oaxaca	97.6	44.1	0	70.2
San Pedro El Alto	Oaxaca	77.8	32.3	0	83.5
San Pedro Huilotepec	Oaxaca	90.6	46.1	0	91.0
San Pedro Ixcatlan	Oaxaca	75.8	35.6	0	97.9
San Pedro Jaltepetongo	Oaxaca	92.0	32.7	0	97.4
San Pedro Jicayan	Oaxaca	91.9	25.6	0	96.2
San Pedro Jocotipac	Oaxaca	85.6	43.9	0	90.0
San Pedro Martir	Oaxaca	95.0	31.7	0	91.0
San Pedro Martir Quiechapa	Oaxaca	100.0	25.2	0	56.0
San Pedro Molinos	Oaxaca	95.1	37.3	0	96.2
San Pedro Ocotepec	Oaxaca	85.6	29.7	0	99.2
San Pedro Quiatoni	Oaxaca	75.2	33.1	0	96.7
San Pedro Sochiapam	Oaxaca	83.4	31.9	0	96.6
San Pedro Teutila	Oaxaca	88.9	34.6	0	81.7
San Pedro Tidaa	Oaxaca	97.3	32.6	0	87.6
San Pedro Topiltepec	Oaxaca	100.0	33.2	0	71.2
San Pedro Yaneri	Oaxaca	90.6	46.4	0	99.1
San Pedro Yolox	Oaxaca	91.7	28.8	0	98.7
San Pedro Y San Pablo Teposcolula	Oaxaca	97.8	41.0	0	53.6
San Sebastian Abasolo	Oaxaca	98.5	45.3	0	80.9
San Sebastian Rio Hondo	Oaxaca	92.6	40.3	0	60.8
San Sebastian Tecomaxtlahuaca	Oaxaca	85.3	40.3	0	85.6
San Sebastian Teitipac	Oaxaca	91.5	45.7	0	66.0
Santa Ana Ateixtlahuaca	Oaxaca	75.5	36.7	0	100.0
Santa Ana Cuauhtemoc	Oaxaca	86.2	32.9	0	96.1
Santa Ana del Valle	Oaxaca	96.7	37.1	0	97.0
Santa Ana Tavela	Oaxaca	96.0	31.2	0	52.5
Santa Ana Yareni	Oaxaca	90.9	39.1	0	99.0
Santa Ana Zegache	Oaxaca	95.0	45.7	0	77.6
Santa Catalina Quieri	Oaxaca	94.1	39.0	0	100.0
Santa Catarina Cuixtla	Oaxaca	95.9	38.8	0	76.7
Santa Catarina Ixtepeji	Oaxaca	98.3	41.3	0	87.5
Santa Catarina Lachatao	Oaxaca	97.2	42.8	0	93.5
Santa Catarina Loxicha	Oaxaca	90.6	27.5	0	61.4
Santa Catarina Mechoacan	Oaxaca	87.7	37.4	0	98.9
Santa Catarina Tayata	Oaxaca	98.1	39.3	0	66.7

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Santa Catarina Ticua	Oaxaca	95.1	41.8	0	90.3
Santa Catarina Yosonotu	Oaxaca	89.8	26.0	0	98.9
Santa Cruz Itundujia	Oaxaca	94.3	26.0	0	61.2
Santa Cruz Mixtepec	Oaxaca	92.3	41.8	0	58.1
Santa Cruz Nundaco	Oaxaca	94.7	40.4	0	97.9
Santa Cruz Xitla	Oaxaca	81.9	36.4	0	77.1
Santa Ines Yatzeche	Oaxaca	76.5	32.6	0	98.9
Santa Lucia Ocotlan	Oaxaca	86.7	37.8	0	97.2
Santa Maria Alotepec	Oaxaca	91.6	31.6	0	97.9
Villa de Chilapa de Diaz	Oaxaca	96.0	37.8	0	56.5
Santa Maria Ecatepec	Oaxaca	96.6	39.0	0	91.3
Santa Maria Guienagati	Oaxaca	86.3	32.7	0	95.9
Santa Maria Ixcatlan	Oaxaca	95.7	27.3	0	61.2
Santa Maria Jacatepec	Oaxaca	88.1	34.4	0	85.3
Santa Maria Lachixio	Oaxaca	77.5	26.0	0	96.1
Santa Maria Mixtequilla	Oaxaca	96.1	38.7	0	60.2
Santa Maria Nativitas	Oaxaca	98.0	41.9	0	85.9
Santa Maria Nduayaco	Oaxaca	99.0	32.6	0	59.0
Santa Maria Ozolotepec	Oaxaca	83.8	34.5	0	90.5
Santa Maria Papalo	Oaxaca	86.0	33.3	0	98.6
Santa Maria Quiegolani	Oaxaca	79.5	34.2	0	98.0
Santa Maria Tataltepec	Oaxaca	98.0	26.0	0	100.0
Santa Maria Temaxcalapa	Oaxaca	97.0	41.6	0	99.6
Santa Maria Teopoxco	Oaxaca	81.1	29.2	0	97.3
Santa Maria Texcatitlan	Oaxaca	85.7	50.5	0	97.3
Santa Maria Tlahuitoltepec	Oaxaca	86.7	41.9	0	99.1
Santa Maria Tonameca	Oaxaca	81.3	35.4	0	54.1
Santa Maria Totolapilla	Oaxaca	90.4	35.2	0	88.1
Santa Maria Yalina	Oaxaca	96.7	32.0	0	96.7
Santa Maria Yavesia	Oaxaca	100.0	27.8	0	99.0
Santiago Apoala	Oaxaca	90.8	27.6	0	94.3
Santiago Apostol	Oaxaca	82.2	40.7	0	94.3
Santiago Atitlan	Oaxaca	82.4	38.1	0	98.9
Santiago Cacaloxtepec	Oaxaca	85.0	44.9	0	86.7
Santiago Camotlan	Oaxaca	90.6	38.5	0	98.3
Santiago Comaltepec	Oaxaca	97.1	34.2	0	94.2
Santiago Choapam	Oaxaca	83.0	29.3	0	88.8
Santiago del Rio	Oaxaca	78.8	30.7	0	69.2
Santiago Huauclilla	Oaxaca	95.8	28.5	0	81.0
Santiago Ixcuintepec	Oaxaca	91.4	34.4	0	98.7
Santiago Jocotepec	Oaxaca	83.9	30.6	0	95.2
Santiago Lachiguiri	Oaxaca	90.5	39.7	0	84.1
Santiago Lalopa	Oaxaca	94.7	47.3	0	97.7
Santiago Laxopa	Oaxaca	98.6	39.5	0	98.8
Santiago Miltepec	Oaxaca	97.1	47.5	0	67.0
Santiago Nejapilla	Oaxaca	100.0	29.5	0	61.5
Santiago Nundiche	Oaxaca	97.0	30.2	0	99.1
Santiago Tetepec	Oaxaca	89.8	33.1	0	56.4
Santiago Tillo	Oaxaca	97.2	47.0	0	58.0

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Santiago Xanica	Oaxaca	78.2	29.8	0	90.4
Santiago Xiacui	Oaxaca	98.8	37.1	0	85.9
Santiago Yaveo	Oaxaca	91.1	37.1	0	81.7
Santiago Yosondua	Oaxaca	94.1	34.6	0	87.1
Santiago Zoochila	Oaxaca	98.9	36.0	0	100.0
Nuevo Zoquiapam	Oaxaca	96.3	39.4	0	92.2
Santo Domingo Albarradas	Oaxaca	95.4	28.6	0	99.5
Santo Domingo Chihuitan	Oaxaca	96.0	41.4	0	55.0
Santo Domingo Ixcatlan	Oaxaca	98.7	33.2	0	93.1
Santo Domingo Nuxaa	Oaxaca	98.4	30.8	0	99.2
Santo Domingo Petapa	Oaxaca	83.3	44.3	0	88.4
Santo Domingo Roayaga	Oaxaca	79.8	40.8	0	99.2
Santo Domingo Tlatayapam	Oaxaca	96.8	41.1	0	63.6
Santo Domingo Tomaltepec	Oaxaca	96.8	50.0	0	77.7
Santo Domingo Tonaltepec	Oaxaca	97.1	29.1	0	90.0
Santo Domingo Xagacia	Oaxaca	97.1	43.0	0	99.6
Santo Domingo Yanhuitlan	Oaxaca	98.0	38.0	0	75.5
Santos Reyes Papalo	Oaxaca	86.3	29.6	0	89.4
San Vicente Nunu	Oaxaca	97.7	40.9	0	85.9
Silacayoapam	Oaxaca	93.5	32.3	0	67.0
Tataltepec de Valdes	Oaxaca	80.5	30.9	0	90.8
Teococuilco de Marcos Perez	Oaxaca	98.8	38.0	0	95.0
Teotitlan del Valle	Oaxaca	95.3	49.0	0	96.6
Teotongo	Oaxaca	96.7	32.2	0	73.7
Tezoatlan de Segura y Luna	Oaxaca	93.1	29.7	0	77.5
Tlacotepec Plumas	Oaxaca	98.0	27.9	0	75.2
Totontepec Villa de Morelos	Oaxaca	85.3	26.7	0	98.3
La Trinidad Vista Hermosa	Oaxaca	98.5	31.1	0	60.0
Union Hidalgo	Oaxaca	97.6	40.8	0	89.4
Villa Diaz Ordaz	Oaxaca	94.1	44.6	0	96.1
Magdalena Yodocono de Porfirio Diaz	Oaxaca	96.7	43.8	0	71.1
Yutanduchi de Guerrero	Oaxaca	94.1	35.7	0	82.6
Zapotitlan Palmas	Oaxaca	93.1	32.0	0	76.7
Acteopan	Puebla	79.5	62.8	0	82.1
Altepexi	Puebla	87.6	58.9	0	75.4
Amixtlan	Puebla	77.4	34.2	0	96.7
Atempan	Puebla	86.0	41.2	0	86.5
Atexcal	Puebla	93.7	33.5	0	55.5
Caltepec	Puebla	91.6	36.1	0	55.6
Camocuautla	Puebla	76.2	34.2	0	99.7
Caxhuacan	Puebla	88.2	38.1	0	93.7
Coatepec	Puebla	91.7	37.6	0	100.0
Cohuecan	Puebla	96.8	46.0	0	63.4
Cuautempan	Puebla	91.5	45.4	0	72.3
Chigmecatitlan	Puebla	96.2	46.9	0	95.5
Chignautla	Puebla	88.3	50.5	0	69.2
Huatlatlauca	Puebla	86.2	40.1	0	88.2
Hueyapan	Puebla	90.7	42.3	0	97.1
Atlequizayan	Puebla	83.4	32.0	0	99.7

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Jalpan	Puebla	92.0	35.8	0	53.7
Jonotla	Puebla	89.9	40.2	0	83.3
Jopala	Puebla	79.8	33.1	0	93.2
Naupan	Puebla	86.9	44.5	0	96.1
Nauzontla	Puebla	93.1	39.3	0	82.1
Nealtican	Puebla	95.0	47.0	0	57.8
San Antonio Canada	Puebla	77.6	41.2	0	65.9
San Felipe Tepatlan	Puebla	79.3	38.8	0	80.1
San Gabriel Chilac	Puebla	85.9	53.6	0	85.6
San Jeronimo Xayacatlan	Puebla	96.2	37.1	0	90.9
San Jose Miahuatlan	Puebla	90.0	47.5	0	85.5
Santa Catarina Tlaltempan	Puebla	91.8	37.4	0	89.3
Santa Ines Ahuatempan	Puebla	88.5	28.8	0	80.2
Huehuetlan El Grande	Puebla	87.0	30.3	0	54.9
Tenampulco	Puebla	92.3	38.3	0	50.6
Teopantlan	Puebla	81.9	43.9	0	84.0
Teteles de Avila Castillo	Puebla	96.5	46.2	0	53.2
Tlacuilotepec	Puebla	91.4	35.8	0	69.2
Tlapacoya	Puebla	82.3	35.3	0	72.1
Tlaxco	Puebla	88.5	40.9	0	74.6
Totoltepec de Guerrero	Puebla	97.6	33.6	0	58.7
Tuzamapan de Galeana	Puebla	91.3	42.6	0	91.8
Xayacatlan de Bravo	Puebla	94.9	37.9	0	93.2
Xochiapulco	Puebla	94.6	36.6	0	89.4
Xochitlan de Vicente Suarez	Puebla	81.3	37.0	0	94.6
Yaonahuac	Puebla	96.6	46.3	0	91.2
Zautla	Puebla	91.9	36.4	0	83.1
Zihuateutla	Puebla	85.4	39.9	0	66.6
Zongozotla	Puebla	82.6	45.1	0	97.0
Zoquiapan	Puebla	86.7	39.4	0	89.8
Toliman	Queretaro	95.8	42.8	0	82.2
Tancanhuitz de Santos	San Luis Potosi	94.8	33.2	0	89.6
Coxcatlan	San Luis Potosi	94.0	35.1	0	97.7
Huehuetlan	San Luis Potosi	92.1	34.5	0	93.2
San Antonio	San Luis Potosi	94.4	31.0	0	97.7
San Martin Chalchicuautla	San Luis Potosi	93.7	36.0	0	66.0
Tampacan	San Luis Potosi	96.0	39.9	0	85.3
Tampamolon Corona	San Luis Potosi	92.9	38.5	0	93.6
Tanlajas	San Luis Potosi	93.2	30.7	0	92.2
Tanquian de Escobedo	San Luis Potosi	95.2	43.3	0	58.6
Matlapa	San Luis Potosi	88.3	33.8	0	86.8
Etchojoa	Sonora	98.2	44.3	0	79.6
Benito Juarez	Sonora	98.5	44.4	0	50.9
Benito Juarez	Veracruz	90.0	32.6	0	91.0
Coxquihui	Veracruz	85.7	35.0	0	95.8
Coyutla	Veracruz	79.7	40.5	0	92.6
Chiconamel	Veracruz	79.4	33.0	0	57.2
Chontla	Veracruz	92.7	39.9	0	78.2
Chumatlan	Veracruz	79.8	32.3	0	99.0

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Ixcatepec	Veracruz	93.0	38.1	0	75.0
Magdalena	Veracruz	79.9	38.8	0	98.6
Oteapan	Veracruz	89.9	48.4	0	56.7
Rafael Delgado	Veracruz	91.3	47.8	0	83.4
San Andres Tenejapan	Veracruz	82.0	48.6	0	93.9
Tepetzintla	Veracruz	96.2	42.3	0	56.5
Texhuacan	Veracruz	80.5	33.2	0	97.7
Tlachichilco	Veracruz	81.7	30.1	0	63.0
Tlilapan	Veracruz	88.1	46.4	0	73.8
Zaragoza	Veracruz	92.3	44.2	0	91.1
Zozocolco de Hidalgo	Veracruz	83.5	33.5	0	87.6
Tatahuicapan de Juarez	Veracruz	75.2	34.6	0	88.2
Abala	Yucatan	87.9	49.5	0	92.8
Akil	Yucatan	90.7	49.4	0	84.9
Buctzotz	Yucatan	93.5	48.5	0	53.1
Calotmul	Yucatan	94.0	42.4	0	95.9
Cansahcab	Yucatan	95.0	44.5	0	82.0
Cantamayec	Yucatan	81.9	46.5	0	99.0
Cenotillo	Yucatan	97.1	42.4	0	61.1
Cuncunul	Yucatan	90.2	43.4	0	98.7
Chacsinkin	Yucatan	83.5	45.3	0	99.4
Chankom	Yucatan	87.1	44.4	0	99.6
Chapab	Yucatan	93.3	45.8	0	95.1
Chemax	Yucatan	78.3	41.1	0	98.7
Chichimila	Yucatan	84.4	40.6	0	98.3
Chikindzonot	Yucatan	85.4	41.5	0	98.7
Chochola	Yucatan	95.6	48.0	0	81.5
Chumayel	Yucatan	88.7	59.9	0	97.2
Dzan	Yucatan	95.5	48.1	0	95.6
Dzemul	Yucatan	93.4	49.6	0	81.9
Dzoncauich	Yucatan	91.6	33.5	0	76.9
Espita	Yucatan	87.8	41.9	0	93.8
Halacho	Yucatan	88.6	47.0	0	94.4
Hocaba	Yucatan	88.6	47.6	0	82.8
Hoctun	Yucatan	81.7	44.5	0	92.5
Huhi	Yucatan	92.8	48.0	0	86.2
Kantunil	Yucatan	91.8	36.9	0	82.4
Kaua	Yucatan	87.8	44.1	0	95.9
Kopoma	Yucatan	96.5	47.2	0	79.0
Mama	Yucatan	89.3	49.2	0	93.0
Mani	Yucatan	91.8	49.1	0	97.5
Mayapan	Yucatan	75.5	45.0	0	87.0
Muna	Yucatan	93.2	48.2	0	52.2
Muxupip	Yucatan	94.2	47.4	0	83.2
Opichen	Yucatan	93.2	48.7	0	95.3
Panaba	Yucatan	92.4	45.8	0	78.7
Quintana Roo	Yucatan	86.6	39.3	0	67.1
Rio Lagartos	Yucatan	96.4	44.0	0	72.3
Samahil	Yucatan	91.9	49.1	0	90.8

Table A.1: List of municipalities included in the moderately marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Sanahcat	Yucatan	93.8	43.8	0	92.2
San Felipe	Yucatan	97.5	49.7	0	66.4
Santa Elena	Yucatan	90.0	43.7	0	96.1
Sinanche	Yucatan	95.1	43.2	0	59.1
Sotuta	Yucatan	92.5	46.4	0	90.6
Sucila	Yucatan	96.4	49.4	0	77.0
Sudzal	Yucatan	91.5	43.1	0	78.4
Tahdziu	Yucatan	76.0	54.0	0	99.3
Tahmek	Yucatan	95.0	48.5	0	67.1
Teabo	Yucatan	81.2	61.7	0	96.3
Tecoh	Yucatan	89.9	55.0	0	91.6
Tekal de Venegas	Yucatan	90.9	39.0	0	83.4
Tekanto	Yucatan	91.9	47.8	0	84.9
Tekom	Yucatan	86.1	41.7	0	98.4
Telchac Pueblo	Yucatan	96.2	49.4	0	71.6
Telchac Puerto	Yucatan	94.3	44.6	0	71.9
Temax	Yucatan	93.3	43.9	0	70.4
Temozon	Yucatan	90.0	44.1	0	99.3
Tepakan	Yucatan	89.3	40.4	0	91.0
Tetiz	Yucatan	85.5	52.8	0	87.5
Teya	Yucatan	92.5	42.8	0	52.0
Timucuy	Yucatan	82.0	54.3	0	94.6
Tinum	Yucatan	91.1	46.6	0	96.8
Tixcacalcupul	Yucatan	82.7	38.3	0	99.1
Tixmehuac	Yucatan	83.4	43.1	0	97.8
Tunkas	Yucatan	92.2	41.6	0	87.0
Tzucacab	Yucatan	90.1	43.3	0	98.4
Uayma	Yucatan	86.0	43.3	0	92.8
Xocchel	Yucatan	91.8	46.1	0	87.7
Yaxcaba	Yucatan	86.1	42.9	0	99.3
Yobain	Yucatan	97.1	44.1	0	60.4

Table A.2: List of municipalities included in the severely marginalised category, along with relevant indicators.

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Amatenango del Valle	Chiapas	67.1	50.0	0	90.4
Chanal	Chiapas	71.5	32.4	0	99.9
Chenalho	Chiapas	63.0	39.1	0	98.8
Chilon	Chiapas	63.2	33.2	0	96.7
Jitotol	Chiapas	72.2	39.1	0	78.6
Mitontic	Chiapas	52.2	37.6	0	99.7
Ocotepec	Chiapas	66.0	31.0	0	97.1
Pantelho	Chiapas	53.3	32.8	0	89.7
Pueblo Nuevo Solistahuacan	Chiapas	66.7	40.8	0	55.0
Sitala	Chiapas	48.2	34.1	0	94.5

Table A.2: List of municipalities included in the severely marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Tenejapa	Chiapas	74.6	39.7	0	99.5
Tumbala	Chiapas	66.3	35.0	0	99.2
Zinacantan	Chiapas	52.9	41.9	0	99.1
San Juan Cancuc	Chiapas	58.7	33.6	0	99.8
Aldama	Chiapas	72.6	33.0	0	99.4
Maravilla Tenejapa	Chiapas	74.5	30.4	0	78.5
Marques de Comillas	Chiapas	73.7	41.1	0	54.9
San Andres Duraznal	Chiapas	62.8	32.6	0	98.9
Santiago El Pinar	Chiapas	53.8	35.2	0	99.7
Urique	Chihuahua	80.6	0.0	0	52.7
Ahuacuotzingo	Guerrero	73.3	28.7	0	50.9
Atlamajalcingo del Monte	Guerrero	74.1	13.6	0	99.2
Atlixtac	Guerrero	62.5	35.2	0	75.5
Copalillo	Guerrero	69.7	27.2	0	97.3
Copanatoyac	Guerrero	62.4	30.3	0	97.4
Tlacoachistlahuaca	Guerrero	56.6	34.4	0	92.6
Xalpatlahuac	Guerrero	68.4	23.4	0	97.9
Zitlala	Guerrero	73.9	37.5	0	82.2
Mezquitic	Jalisco	74.1	24.5	0	83.5
Abejones	Oaxaca	82.5	10.8	0	95.2
Calihuala	Oaxaca	74.9	24.1	0	61.5
Coatecas Altas	Oaxaca	66.3	21.4	0	61.2
Coicoyan de las Flores	Oaxaca	42.9	16.3	0	99.1
Constancia del Rosario	Oaxaca	68.2	30.5	0	78.9
Eloxochitlan de Flores Magon	Oaxaca	66.0	39.9	0	98.4
Huautepec	Oaxaca	58.8	38.2	0	98.8
Santa Magdalena Jicotlan	Oaxaca	100.0	14.1	0	73.9
Magdalena Penasco	Oaxaca	79.0	17.8	0	97.9
Magdalena Teitipac	Oaxaca	72.6	44.7	0	97.6
Ixpantepec Nieves	Oaxaca	90.6	24.9	0	96.5
Santa Catarina Quioquitani	Oaxaca	97.3	21.6	0	99.1
San Agustin Tlacotepec	Oaxaca	95.4	23.8	0	97.0
San Andres Nuxino	Oaxaca	98.0	18.6	0	81.0
San Andres Paxtlan	Oaxaca	72.7	40.2	0	95.7
San Antonino Monte Verde	Oaxaca	92.9	14.7	0	97.5
San Antonio Acutla	Oaxaca	94.3	19.9	0	64.3
San Antonio Huitepec	Oaxaca	97.7	20.6	0	94.7
San Antonio Sinicahua	Oaxaca	81.2	11.9	0	100.0
San Antonio Tepetlapa	Oaxaca	84.0	14.1	0	95.0
San Bartolome Ayautla	Oaxaca	65.5	30.1	0	99.2
San Blas Atempa	Oaxaca	68.9	50.1	0	85.9
San Cristobal Amatlan	Oaxaca	70.3	42.9	0	96.5
San Cristobal Amoltepec	Oaxaca	89.0	24.0	0	99.7
San Esteban Atatlahuca	Oaxaca	92.7	24.8	0	98.2
San Francisco Cahuacua	Oaxaca	95.5	22.4	0	61.0
San Francisco Tlapancingo	Oaxaca	73.0	22.7	0	71.7
San Jose Lachiguiri	Oaxaca	68.7	12.1	0	99.7
San Jose Tenango	Oaxaca	67.7	37.2	0	98.8
San Juan Atepec	Oaxaca	96.2	24.9	0	96.8

Table A.2: List of municipalities included in the severely marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
San Juan Bautista Tlachichilco	Oaxaca	88.2	19.3	0	58.1
San Juan Diuxi	Oaxaca	78.2	9.3	0	97.5
San Juan Mixtepec - Distr. 26 -	Oaxaca	89.0	14.1	0	98.3
San Juan Numi	Oaxaca	95.3	21.2	0	98.6
San Juan Petlapa	Oaxaca	70.2	36.7	0	100.0
San Juan Tamazola	Oaxaca	96.0	23.2	0	96.9
San Lorenzo Texmelucan	Oaxaca	71.0	22.3	0	98.4
San Lucas Camotlan	Oaxaca	76.0	20.6	0	99.6
San Lucas Zoquiapam	Oaxaca	72.8	38.4	0	99.0
San Martin Itunyoso	Oaxaca	66.3	16.6	0	99.5
San Martin Peras	Oaxaca	48.9	10.2	0	99.3
San Mateo Nejapam	Oaxaca	86.7	13.2	0	68.3
San Miguel Ahuehuetitlan	Oaxaca	68.3	35.3	0	80.4
San Miguel Chicahua	Oaxaca	91.7	21.7	0	93.2
San Miguel Huautla	Oaxaca	89.8	3.5	0	84.6
San Miguel Mixtepec	Oaxaca	77.8	23.0	0	97.0
San Miguel Panixtlahuaca	Oaxaca	76.6	24.6	0	98.4
San Miguel Piedras	Oaxaca	96.9	23.5	0	84.7
San Miguel Quetzaltepec	Oaxaca	68.7	35.4	0	98.8
San Miguel Santa Flor	Oaxaca	74.5	32.1	0	93.4
San Miguel Tilquiapam	Oaxaca	72.9	33.0	0	96.0
San Pablo Tijaltepec	Oaxaca	80.9	20.2	0	97.5
San Pedro Mixtepec - Distr. 26 -	Oaxaca	94.6	14.7	0	97.0
San Pedro Ocopetatillo	Oaxaca	74.2	40.7	0	98.4
San Pedro Y San Pablo Ayutla	Oaxaca	74.8	41.4	0	98.5
San Sebastian Nicananduta	Oaxaca	93.8	18.1	0	97.6
San Simon Zahuatlan	Oaxaca	44.0	10.6	0	97.0
Santa Cruz Acatepec	Oaxaca	73.3	40.7	0	98.5
Santa Cruz Tacahua	Oaxaca	95.9	11.2	0	83.6
Santa Cruz Zenzontepec	Oaxaca	71.2	11.7	0	95.8
Santa Lucia Miahuatlan	Oaxaca	61.7	26.4	0	98.6
Santa Lucia Monteverde	Oaxaca	88.9	15.4	0	97.6
Santa Maria Apazco	Oaxaca	79.1	9.7	0	93.1
Santa Maria La Asuncion	Oaxaca	49.8	36.6	0	99.2
Santa Maria Chilchotla	Oaxaca	74.5	33.1	0	98.4
Santa Maria Chimalapa	Oaxaca	81.1	0.0	0	87.6
Santa Maria Jaltianguis	Oaxaca	97.2	19.6	0	93.0
Santa Maria Penoles	Oaxaca	83.6	15.1	0	97.9
Santa Maria Petapa	Oaxaca	92.4	0.0	0	78.5
Santa Maria Temaxcaltepec	Oaxaca	67.8	36.6	0	96.9
Santa Maria Tepantlali	Oaxaca	69.2	50.4	0	99.2
Santa Maria Tlalixtac	Oaxaca	77.7	12.3	0	74.2
Santa Maria Xadani	Oaxaca	72.1	52.3	0	94.5
Santa Maria Yolotepec	Oaxaca	99.0	24.4	0	99.0
Santa Maria Yosoyua	Oaxaca	89.6	19.4	0	98.3
Santa Maria Yucuhiti	Oaxaca	95.5	14.5	0	99.4
Santa Maria Zaniza	Oaxaca	80.0	18.4	0	73.2
Santiago Amoltepec	Oaxaca	73.6	21.9	0	93.7
Santiago Ixtayutla	Oaxaca	62.2	29.0	0	87.3

Table A.2: List of municipalities included in the severely marginalised category, along with relevant indicators. *(continued)*

Municipality	State	Maternal Literacy (%)	Proportion in Formal Economy (%)	Clinics per capita	Proportion Identifying as Indigenous (%)
Santiago Nuyoo	Oaxaca	98.1	24.1	0	99.4
Santiago Texcalcingo	Oaxaca	73.8	30.8	0	97.9
Santiago Textitlan	Oaxaca	88.9	13.0	0	61.7
Santiago Tilantongo	Oaxaca	93.5	24.6	0	87.3
Santiago Tlazoyaltepec	Oaxaca	78.6	13.9	0	99.0
Santiago Yaitepec	Oaxaca	55.1	39.5	0	98.5
Santiago Zacatepec	Oaxaca	59.7	34.8	0	99.4
Santo Domingo de Morelos	Oaxaca	73.2	21.1	0	96.5
Santo Domingo Tepuxtepec	Oaxaca	70.9	23.1	0	98.9
Santos Reyes Tepejillo	Oaxaca	88.1	24.9	0	95.5
Santos Reyes Yucuna	Oaxaca	68.0	12.4	0	98.6
Santo Tomas Ocotepec	Oaxaca	89.3	20.6	0	98.6
San Vicente Coatlan	Oaxaca	85.6	20.4	0	95.0
San Vicente Lachixio	Oaxaca	85.7	15.6	0	99.6
Santa Ines de Zaragoza	Oaxaca	98.0	20.8	0	96.0
Coyomeapan	Puebla	74.8	28.8	0	99.0
Chiconcuautla	Puebla	64.8	34.4	0	84.8
Chichiquila	Puebla	73.7	33.7	0	77.1
Chilchotla	Puebla	73.5	31.4	0	68.2
Eloxochitlan	Puebla	69.5	32.1	0	94.3
Hermenegildo Galeana	Puebla	74.4	32.7	0	96.1
Hueytlalpan	Puebla	73.1	36.4	0	95.5
Huitzilan de Serdan	Puebla	69.6	42.8	0	92.7
Olintla	Puebla	73.4	33.8	0	97.2
Tepango de Rodriguez	Puebla	73.5	36.1	0	96.8
Tepetzintla	Puebla	72.2	37.2	0	95.1
Santa Catarina	San Luis Potosi	81.2	24.7	0	71.2
Astacinga	Veracruz	64.7	31.6	0	98.7
Atlahuilco	Veracruz	65.2	31.6	0	98.5
Coahuitlan	Veracruz	69.5	30.5	0	96.3
Filomeno Mata	Veracruz	61.9	31.9	0	98.3
Ilamatlan	Veracruz	68.4	30.8	0	92.9
Mecatlan	Veracruz	71.6	28.0	0	98.7
Mixtla de Altamirano	Veracruz	53.6	36.5	0	99.6
Pajapan	Veracruz	73.6	41.7	0	90.8
Los Reyes	Veracruz	74.2	33.4	0	98.3
Soledad Atzompa	Veracruz	66.7	32.8	0	98.2
Soteapan	Veracruz	62.3	33.4	0	93.1
Tehuipango	Veracruz	40.5	27.8	0	99.6
Tequila	Veracruz	74.5	38.7	0	98.4
Texcatepec	Veracruz	68.4	28.0	0	84.8
Xoxocotla	Veracruz	74.2	36.0	0	91.8
Zontecomatlan de Lopez y Fuentes	Veracruz	74.8	22.2	0	90.9
Metlatonoc, Cochoapa el Grande	Guerrero	42.5	19.7	0	99.6

Appendix B

Appendix to Chapter Three

This section contains additional figures and tables related to Chapter Three: Mapping the relationship between tuberculosis burden and case notifications in Uganda.

Scenarios: duration as a function of completeness

The relationship between between data sources describing TB prevalence and incidence is mediated by the expected (mean) duration of active TB disease: $Prevalence = Incidence * E[Duration]$. As described in Section 3.2.2, IHME and the WHO use differing methods to estimate active TB duration. I extended the WHO approach, which estimates duration among adults across four population groupings depending on whether an individual is co-infected with HIV and whether they are receiving treatment.[158] I applied several simplifying assumptions to estimate average duration as a function of case notification completeness:

1. The probabilities of receiving treatment and HIV status are independent: for example, the probability of having a TB/HIV co-infection and being treated is equal to the probability of having a TB/HIV co-infection times the probability of receiving treatment.
2. The proportion of all individuals with TB who also have an HIV co-infection is 41.3% across all districts of Uganda.
3. The proportion of individuals receiving treatment is fixed at 71% of the case detection rate, where 71% is the estimated nationwide treatment success rate in Uganda.

4. For individuals who receive TB treatment, the average delay between the onset of active TB and the beginning of treatment is one month (0.077 years).

These assumptions allow for the calculation of a linear relationship between expected duration and TB case notification completeness, π: $E[D] = 1.88 - .67 * \pi$. However, if any of these assumptions are mis-specified, it may bias the expected duration and alter the true relationship between the incidence and prevalence of active TB disease. To test the effect of these assumptions on the expected duration of TB disease, I varied each of the parameters defining the completeness-duration equation and examined differences in the functional relationship as well as changes in estimated duration values at varying levels of notification completeness.

Table B.1 shows the results of this analysis. These results suggest that the greatest possible biasing factor that might alter the expected duration is a large change in the expected duration of disease for individual sub-groupings. Increasing the TB/HIV co-infection rate shortens the average duration of infection and reduces the influence of notification completeness on duration; decreasing the co-infection rate has the opposite effect. Predictably, changing treatment lags and the treatment success rate changes only the slope of the duration-completeness equation, as the intercept can be interpreted as the expected duration when notification completeness is zero. Under these scenarios, the WHO approach to estimating duration only overlaps the duration estimated by IHME, 2.15 years, under two sets of circumstances (a) the expected disease duration for all population subgroups increases, or (b) TB/HIV co-infection rates are lowered and TB notification completeness is zero.

Table B.1: Estimates of active TB duration as a function of case notification completeness, contingent on various assumptions about other factors influencing disease duration. The default scenario corresponds to the duration function applied in Chapter 3. All following scenarios change one parameter other than completeness. The duration equation is listed for each scenario, along with estimated values of duration at varying levels of case notification completeness (denoted as PI).

Scenario	Duration formula (years)	PI=0%	PI=20%	PI=40%	PI=60%	PI=80%	PI=100%
Equation used in chapter	1.88 - 0.68*PI	1.88	1.75	1.61	1.48	1.34	1.21
TB/HIV coinfection rate=20.7%	2.19 - 0.81*PI	2.19	2.03	1.87	1.71	1.54	1.38
TB/HIV coinfection rate=82.6%	1.27 - 0.41*PI	1.27	1.18	1.10	1.02	0.94	0.85
Treatment lag=0	1.88 - 0.73*PI	1.88	1.74	1.59	1.44	1.30	1.15
Treatment lag=0.25 years	1.88 - 0.55*PI	1.88	1.77	1.66	1.55	1.44	1.33
Treatment success rate=1	1.88 - 0.95*PI	1.88	1.69	1.50	1.31	1.12	0.93
Treatment success rate=.5	1.88 - 0.48*PI	1.88	1.79	1.69	1.60	1.50	1.41
All expected durations at lower end of range	0.59 - 0.28*PI	0.59	0.54	0.48	0.42	0.37	0.31
All expected durations at upper end of range	3.17 - 1.07*PI	3.17	2.96	2.75	2.53	2.32	2.10

Model sensitivity analysis at fixed durations

The relationship between expected duration, model estimates of TB prevalence, and model estimates of TB notification completeness also deserves further attention. To test the effect of varying duration on model estimates, I ran a sensitivity analysis using two experimental models with fixed durations. The first experimental model fixed TB duration nationwide at 2.15 years, the expected duration of TB in Uganda as estimated by IHME.[156] The second experimental model fixed TB duration nationwide at 1 year; this value was chosen both as a relatively low estimate for duration based on the results in Table B.1, and because it makes conversion between estimates of prevalence and incidence trivial: $Prevalence = Incidence * 1$.

Table B.2 summarises differences in estimated prevalence and notification completeness across the three models. This sensitivity analysis finds that fixing TB duration at 1 year does not have a significant effect on prevalence nationwide or in the districts with the highest and lowest estimated prevalence values. However, estimated notification completeness drops significantly both at the national level and in Kiruhura, the district with the lowest estimated completeness.

When duration is fixed at 2.15 years, estimated prevalence increases significantly nationwide and in the lowest-prevalence district of Butaleja. The estimated completeness of notifications also rises significantly at the national level and across districts, with an estimated notification completeness of nearly 100% in Kampala. This rise in both estimated prevalence and notification completeness is logical given the known relationship of $Prevalence = Incidence * E[Duration]$. When expected duration increases, a fixed number of incidence cases corresponds to an increasingly high disease prevalence.

Table B.2: Results of sensitivity testing based on duration. Summaries for estimated TB prevalence and notification completeness are provided for each of the three model specifications.

Duration (years)	Nationwide prevalence per 100,000	Lowest prevalence	Highest prevalence	Nationwide completeness (2019)	Lowest completeness (2019)	Highest completeness (2019)
1.88-.67*PI	498 (446-554)	Butaleja: 161 (132-193)	Moroto: 2300 (1645-3186)	54.9% (50.7%-59.1%)	Kiruhura: 36.2% (31.7%-40.8%)	Kampala: 79.6% (68.4%-88.0%)
1	537 (485-594)	Butaleja: 172 (142-209)	Moroto: 2495 (1767-3455)	44.3% (40.0%-48.5%)	Kiruhura: 25.4% (21.5%-29.6%)	Wakiso: 73.2% (62.7%-82.6%)
2.15	776 (720-835)	Butaleja: 259 (216-304)	Napak: 3411 (2625-4346)	64.9% (60.5%-69.2%)	Kiruhura: 33.7% (28.3%-39.7%)	Kampala: 99.6% (96.6%-99.9%)

Maps of estimated TB prevalence and case reporting completeness for each of the experimental models can help to clarify the effects of a changing incidence-prevalence relationship on model predictions. Figure B.1 shows the predicted prevalence, and Figure B.2 shows estimated notification completeness, when duration is fixed at 1 year. These figures use the same colour scheme and plotting bounds as Figures 3.5 and 3.6 to allow for easy comparison. When duration is reduced from an average of approximately 1.5 years to a fixed value of 1 year, the resulting gap between case notifications and estimated prevalent cases is estimated as lower case reporting completeness across all districts nationwide, while prevalence estimates are remarkably similar between the baseline and experimental models. This change suggests that notification data informs estimates of prevalence in a consistent manner as long as the estimated case reporting completeness remains below 1.

Figures B.3 and B.4 show estimates of prevalence and case notification completeness generated by the experimental model with a fixed duration of 2.15 years. In this model, estimated case notification completeness approaches 100% in Kampala and the surrounding Wakiso district. Because notification completeness cannot exceed 100%, case notifications drive up prevalence estimates across the country, although the rank order of districts in terms of prevalence remains similar to the baseline model. These results suggest that when notification completeness approaches 100% in some districts, case notifications inform not only the spatial pattern but also the overall level of TB prevalence in a country.

The results of these sensitivity analyses, and the relationship of each spatial model to the experiences of program managers in the NTLP, can help to refine the range of plausible duration estimates by district. Further work is needed to explore whether case notification reporting can plausibly exceed 100% due to overlapping diagnoses and travel between districts for treatment.

References

156. Ledesma, J. R. *et al.* Global, regional, and national sex differences in the global burden of tuberculosis by HIV status, 1990-2019: results from the Global Burden of Disease Study 2019. *The Lancet Infectious Diseases* **3099,** 1–20 (2021).

158. Glaziou, P., Dodd, P. J., Dean, A. S. & Floyd, K. *Methods used by WHO to estimate the global burden of TB disease: global tuberculosis report 2020* tech. rep. October (World Health Organization, Geneva, 2020). https://cdn.who.int/media/docs/default-source/hq-tuberculosis/global-tuberculosis-report-2020/tb2020_technical_appendix_20201014.pdf.

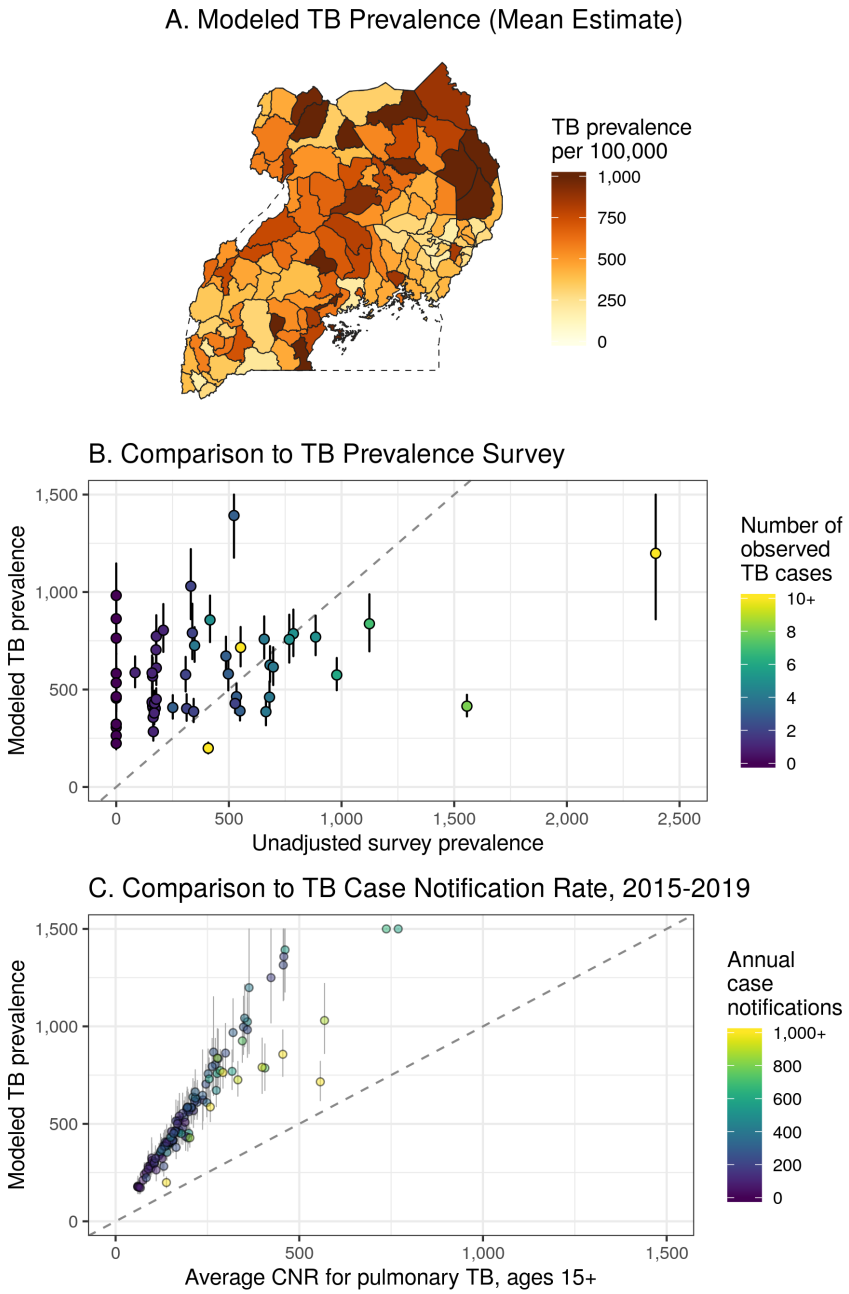


Figure B.1: Reproduction of Figure 3.4 for a model with a fixed disease duration of 1 year.

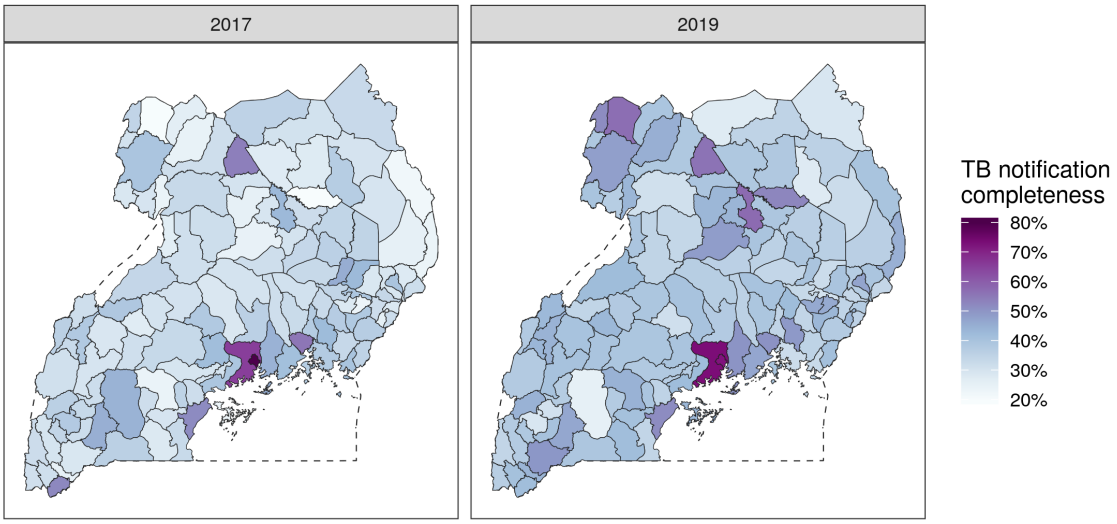


Figure B.2: Reproduction of Figure 3.5 for a model with a fixed disease duration of 1 year.

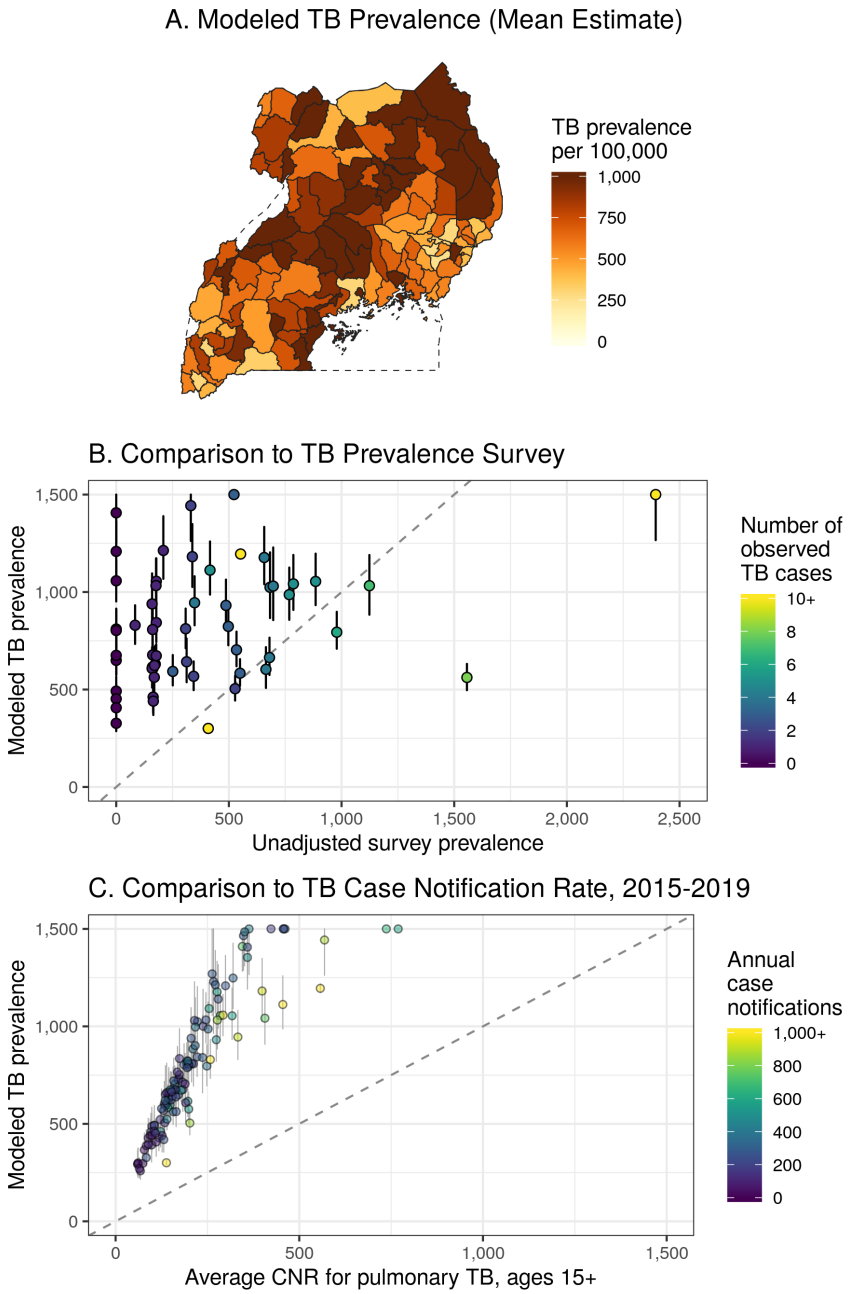


Figure B.3: Reproduction of Figure 3.4 for a model with a fixed disease duration of 2.15 years.

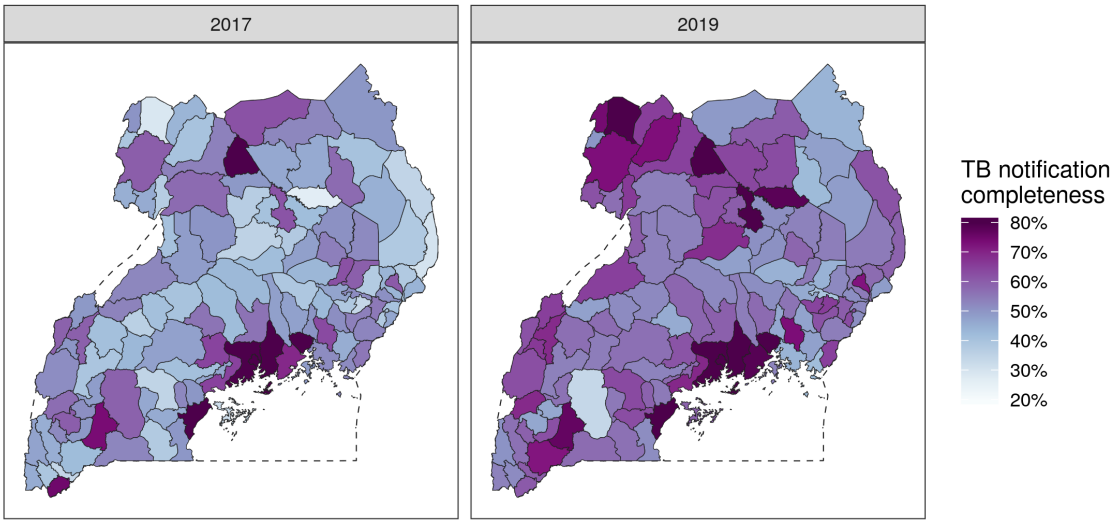


Figure B.4: Reproduction of Figure 3.5 for a model with a fixed disease duration of 2.15 years.

Appendix C

Appendix to Chapter Four

This section contains additional figures related to Chapter Four: A space-time-age model for subnational child mortality estimation in India.

Spatial covariates

The space-time-age mortality estimation model outlined in Section 4.2.2 drew predictive power from ten gridded covariates which varied by pixel and year. Each survey data point was associated with the covariate values in the same pixel-year. After fitting the mortality estimation model, fixed effects were projected across all pixels and years of India using the gridded covariate surfaces, then aggregated to the district level using a population-weighted mean. Covariates were "trimmed" before projecting to remove any values that fell outside of the range observed in the fit data; this process prevents covariate extrapolation that might yield extreme estimates for mortality. All covariates were also rescaled to have mean zero and standard deviation one before fitting to improve comparability of model fixed effects.

Each of the ten gridded covariate surfaces are visualised below for the years 2000 and 2017. Gridded estimates for travel time to settlements of greater than 50,000 people as well as *Plasmodium falciparum* incidence were provided by the Malaria Atlas Project.[60,149] Gridded estimates of population density and the population ratio of children under 5 to women 15-45 were both estimated using data from the WorldPop project.[55] A gridded surface of night-time light intensity was used

as a proxy for local variation in economic activity; this surface was provided by the Defence Meteorological Satellite Program.[66] A gridded land cover covariate was provided by the Global Human Settlement Layer project.[66] Gridded estimates of PM 2.5 air pollution concentration, average years of education among women 15-45, coverage of the diphtheria-tetanus toxoid-pertussis (DPT3) vaccine for children under 5, and child moderate and severe stunting were all estimated by the Institute for Health Metrics and Evaluation.[222,223] Fixed effects coefficients fit to these data should be interpreted cautiously due to collinearity between gridded estimate surfaces.

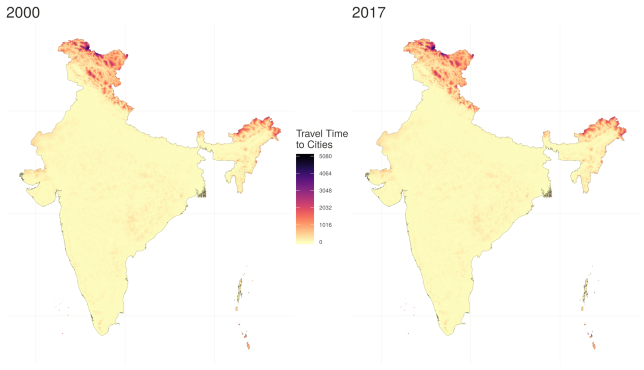


Figure C.1: Map of gridded covariate showing travel time by motorised vehicle to the nearest settlement of 50,000 or greater in 2000 and 2017. Source: Malaria Atlas Project

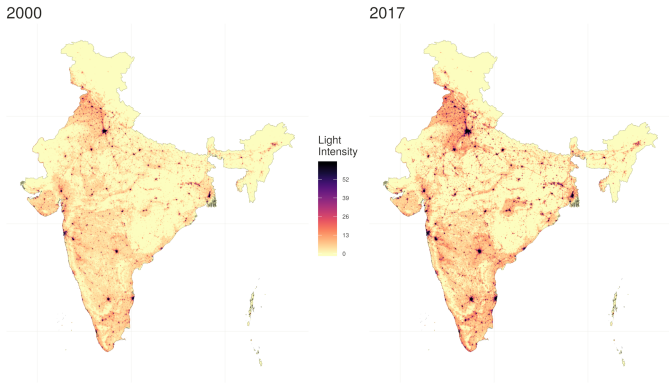


Figure C.2: Map of gridded covariate showing relative nighttime light intensity in 2000 and 2017. Source: Defense Meteorological Satellite Program

References

55. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* **4,** 2–5 (2017).

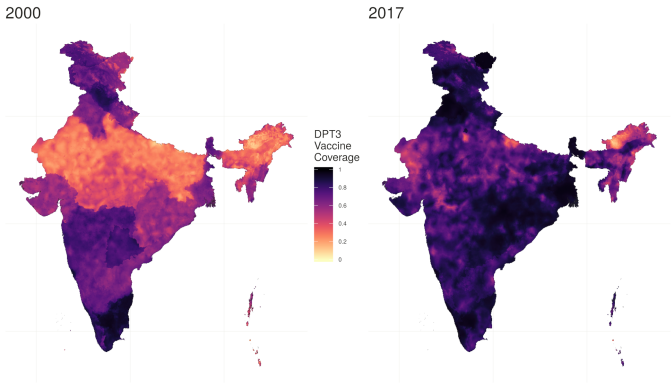


Figure C.3: Map of gridded covariate showing the estiamted proportion of children under age 5 who received the DPT3 vaccine in 2000 and 2017. Source: IHME

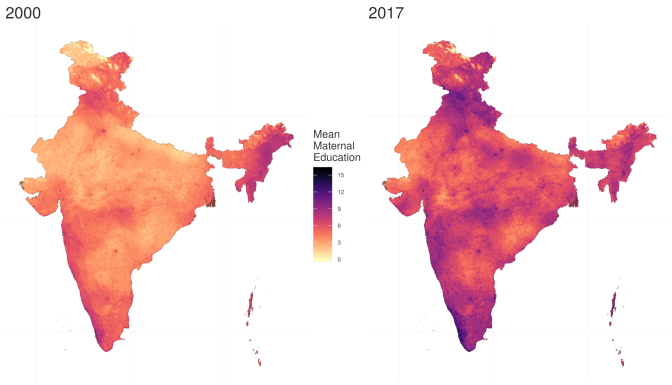


Figure C.4: Map of gridded covariate showing the estimated mean years of education attained for women aged 15-45 in 2000 and 2017. Source: IHME

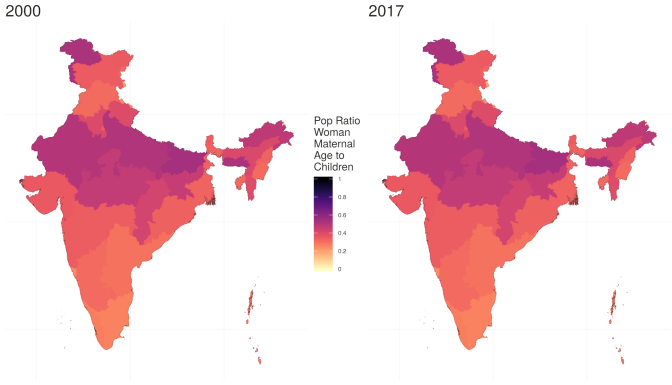


Figure C.5: Map of gridded covariate showing the estimated ratio between children below age 5 and women of childbearing age (15-45) in 2000 and 2017, used as an approximation of fertility. Source: Derived from WorldPop age-specific population estimates.

60. Weiss, D. J. *et al.* Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000-17: a spatial and temporal modelling study. *The Lancet* **394,** 322–331. http://dx.doi.org/10.1016/S0140-6736(19)31097-9 (2019).

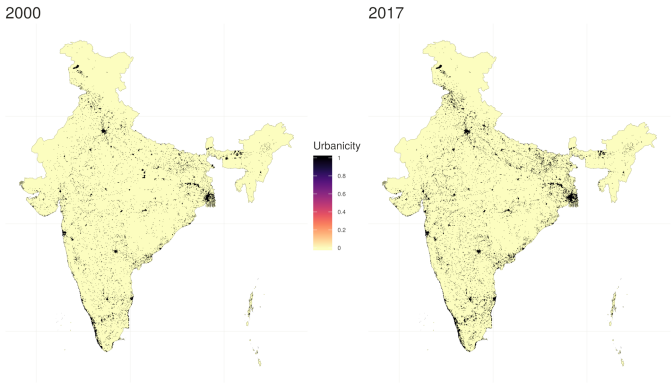


Figure C.6: Map of gridded covariate showing urban land cover extent in 2000 and 2017. Source: Global Human Settlement Layer

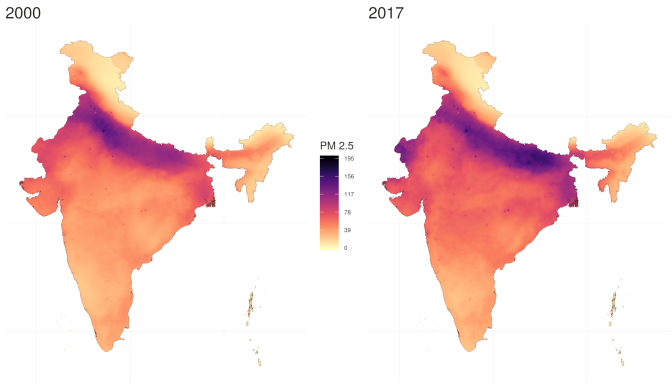


Figure C.7: Map of gridded covariate for estimated PM 2.5 air pollution concentration in 2000 and 2017. Source: IHME

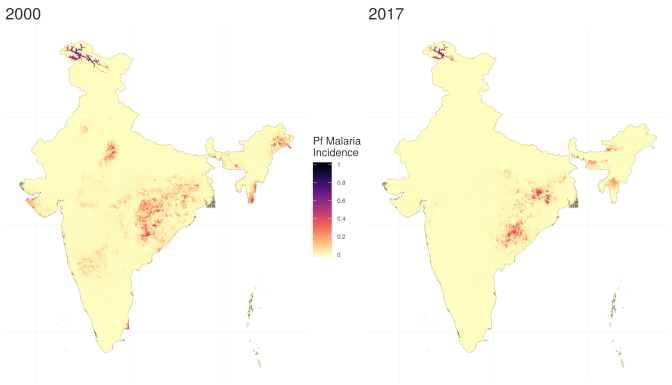


Figure C.8: Map of gridded covariate showing estimated *Plasmodium falciparum* incidence in 2000 and 2017. Source: Malaria Atlas Project

66. Thomson, D. R. *et al.* Extending Data for Urban Health Decision-Making: a Menu of New and Potential Neighborhood-Level Health Determinants Datasets in LMICs. *Journal of Urban Health* **96,** 514–536 (2019).

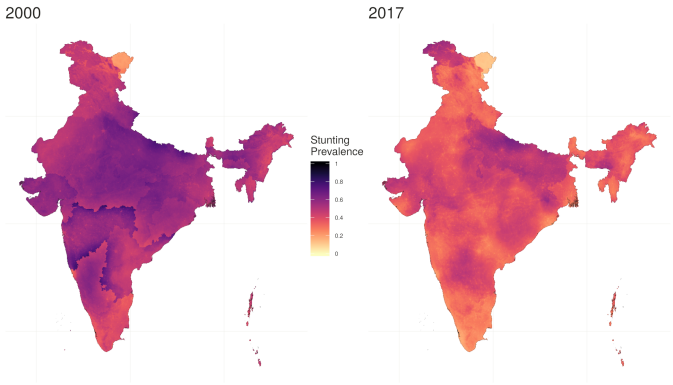


Figure C.9: Map of gridded covariate showing the estimated proportion of children under age 5 who are moderately or severely stunted (low height for age) in 2000 and 2017. Source: IHME

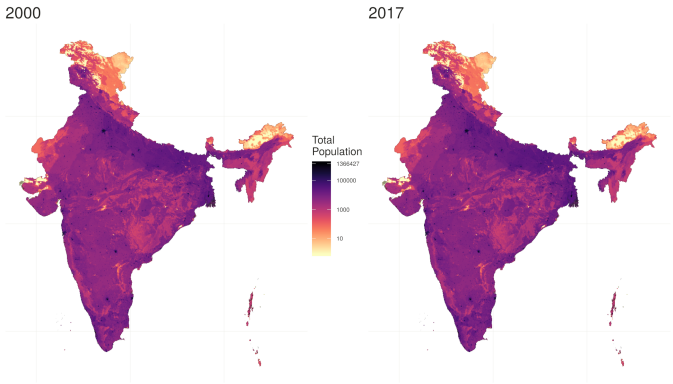


Figure C.10: Map of gridded covariate showing estimated population density in 2000 and 2017. Source: WorldPop

149. Weiss, D. J. *et al.* Global maps of travel time to healthcare facilities. *Nature Medicine* **26,** 1835–1838 (2020).
222. Graetz, N. *et al.* Mapping disparities in education across low- and middle-income countries. *Nature* **577,** 235–238 (2020).
223. Kinyoki, D. K. *et al.* Mapping child growth failure across low- and middle-income countries. *Nature* **577,** 231–234 (2020).

Appendix D

Appendix to Chapter Five

This section contains additional tables related to Chapter Five: Variation in COVID-19 excess mortality by age, sex, and province within Italy.

Covariate preparation

Table 5.1 lists the predictive covariates used to estimate baseline mortality over the study time period, 2015-2020. Synoptic and single-year covariates were used across all years, while time-varying covariates not available for the final years of the time series were projected forwards by replicating values from the final observed year of data. Covariates observed at a more detailed spatial resolution than provinces were aggregated to the province level using a population-weighted mean across units within a province.

Weekly temperature covariate

Point estimates of daily temperature were downloaded for the three most populous pixels in each province using the Meteostat API. Because temperature was not recorded for all time periods and locations, the following process was used to fill temperatures across all provinces and weeks in the time series:

1. Find average weekly temperatures across all observed days by year, week, and observed location.
2. Interpolate by week. In cases where 3 consecutive weeks or fewer are missing between observed temperature values in a location, estimate missing temperatures by interpolating between the nearest observed weeks

3. Aggregate by province. For each province, year, and week of the year, average available observation across the three sampled locations.
4. In rare instances where all observations for a province were missing for a given province, year, and week, temperature was filled with recorded observations from neighbouring provinces with a similar elevation and level of solar exposure.

Model validation

See Section 5.3 for an overview of the model validation process. I compared predictive validity metrics across six model specifications, denoted A through G. Each model specification was fit under two sets of conditions. For in-sample testing, I fit each model type using the full baseline mortality dataset from January 2015 through February 2020, and then compared estimates of underlying mortality rates with the observed data. However, the resulting in-sample metrics for goodness of fit can mask model over-fitting, allowing overly flexible models to follow spurious local trends in the data.

To generate more robust estimates of model predictive validity, I also conducted out-of-sample testing that mimicked the data generation process for all-cause mortality data under a hypothetical baseline where no mortality shock occurred in 2020. Five data holdouts were created, where each holdout was missing observations from March through December in one of the five baseline years 2015-2019. I then fit each model specification based on all remaining observations, and then generated predictive validity metrics by comparing mortality rate estimates and observed data from only the held-out weeks.

As described in Section 5.2, when the number of observed outcomes is low, it may be advantageous to aggregate the results across multiple dimensions of interest before generating predictive validity estimates. For this analysis, all predictive validity metrics were generated for individual observations (disaggregated by province, age group, year, and week) as well as grouped mortality rates across each province-year.

Supplementary Tables D.1 and D.2 report in-sample predictive validity metrics for the six model specifications when compared by province-age-year-week and by province-year groupings, respectively. Supplementary Tables D.3 and D.4 report out-of-sample predictive validity metrics for the six model

Table D.1: In-sample predictive validity metrics comparing model estimates to individual data observations. RMSE: root mean squared error. RSE: relative squared error.

Model	Fourier terms	Seasonality fit by:	Sex	RMSE	RSE	RSE vs. age avg	50% UI	80% UI	90% UI	95% UI	99% UI
A	3	Age, province	Male	0.000977	0.1571	0.8612	0.6174	0.8629	0.9351	0.9686	0.9929
A	3	Age, province	Female	0.000572	0.0913	0.8667	0.6335	0.8683	0.9372	0.9692	0.9924
A	3	Age, province	(Average)	0.000775	0.1242	0.8639	0.6255	0.8656	0.9362	0.9689	0.9927
B	2	Age, province	Male	0.000981	0.1581	0.8667	0.6182	0.8642	0.9365	0.9696	0.9932
B	2	Age, province	Female	0.000574	0.0920	0.8731	0.6344	0.8699	0.9392	0.9701	0.9928
B	2	Age, province	(Average)	0.000778	0.1250	0.8699	0.6263	0.8671	0.9379	0.9699	0.9930
C	1	Age, province	Male	0.000991	0.1614	0.8849	0.6080	0.8546	0.9281	0.9639	0.9910
C	1	Age, province	Female	0.000587	0.0960	0.9107	0.6223	0.8584	0.9295	0.9635	0.9895
C	1	Age, province	(Average)	0.000789	0.1287	0.8978	0.6151	0.8565	0.9288	0.9637	0.9903
D	3	Province	Male	0.000994	0.1625	0.8910	0.6085	0.8545	0.9298	0.9652	0.9917
D	3	Province	Female	0.000585	0.0954	0.9056	0.6263	0.8639	0.9335	0.9668	0.9916
D	3	Province	(Average)	0.000790	0.1290	0.8983	0.6174	0.8592	0.9317	0.9660	0.9917
E	2	Province	Male	0.000995	0.1626	0.8915	0.6067	0.8531	0.9283	0.9640	0.9914
E	2	Province	Female	0.000582	0.0944	0.8964	0.6300	0.8683	0.9375	0.9694	0.9927
E	2	Province	(Average)	0.000788	0.1285	0.8939	0.6184	0.8607	0.9329	0.9667	0.9921
F	1	Province	Male	0.001000	0.1643	0.9010	0.6028	0.8498	0.9250	0.9617	0.9906
F	1	Province	Female	0.000595	0.0985	0.9353	0.6246	0.8624	0.9327	0.9662	0.9913
F	1	Province	(Average)	0.000797	0.1314	0.9181	0.6137	0.8561	0.9288	0.9639	0.9909
G	None	N/A	Male	0.001010	0.1661	0.9105	0.5995	0.8479	0.9242	0.9610	0.9903
G	None	N/A	Female	0.000602	0.1011	0.9596	0.6151	0.8526	0.9264	0.9616	0.9890
G	None	N/A	(Average)	0.000804	0.1336	0.9350	0.6073	0.8502	0.9253	0.9613	0.9897

Table D.2: In-sample predictive validity metrics comparing model estimates to observed data by province-year grouping. RMSE: root mean squared error. RSE: relative squared error.

Model	Fourier terms	Seasonality fit by:	Sex	RMSE	RSE	50% UI	80% UI	90% UI	95% UI	99% UI
A	3	Age, province	Male	2.30e-05	0.2855	0.6426	0.9071	0.9577	0.9788	0.9943
A	3	Age, province	Female	2.55e-05	0.2729	0.6242	0.8929	0.9482	0.9725	0.9897
A	3	Age, province	(Average)	2.42e-05	0.2792	0.6334	0.9000	0.9530	0.9757	0.9920
B	2	Age, province	Male	2.27e-05	0.2778	0.6730	0.9210	0.9654	0.9831	0.9948
B	2	Age, province	Female	2.53e-05	0.2683	0.6532	0.9072	0.9590	0.9768	0.9920
B	2	Age, province	(Average)	2.40e-05	0.2730	0.6631	0.9141	0.9622	0.9800	0.9934
C	1	Age, province	Male	2.45e-05	0.3250	0.6064	0.8722	0.9363	0.9618	0.9848
C	1	Age, province	Female	2.77e-05	0.3225	0.5764	0.8487	0.9222	0.9527	0.9791
C	1	Age, province	(Average)	2.61e-05	0.3237	0.5914	0.8604	0.9292	0.9572	0.9819
D	3	Province	Male	2.33e-05	0.2937	0.6245	0.8980	0.9548	0.9780	0.9943
D	3	Province	Female	2.66e-05	0.2974	0.6203	0.8888	0.9501	0.9735	0.9916
D	3	Province	(Average)	2.50e-05	0.2956	0.6224	0.8934	0.9524	0.9757	0.9930
E	2	Province	Male	2.36e-05	0.3019	0.6117	0.8888	0.9489	0.9743	0.9925
E	2	Province	Female	2.58e-05	0.2809	0.6534	0.9131	0.9642	0.9813	0.9958
E	2	Province	(Average)	2.47e-05	0.2914	0.6325	0.9010	0.9565	0.9778	0.9941
F	1	Province	Male	2.44e-05	0.3223	0.5934	0.8625	0.9308	0.9626	0.9865
F	1	Province	Female	2.77e-05	0.3218	0.6027	0.8805	0.9463	0.9714	0.9884
F	1	Province	(Average)	2.60e-05	0.3220	0.5981	0.8715	0.9386	0.9670	0.9874
G	None	N/A	Male	2.49e-05	0.3351	0.5690	0.8431	0.9233	0.9589	0.9854
G	None	N/A	Female	2.86e-05	0.3446	0.5340	0.8234	0.9137	0.9491	0.9808
G	None	N/A	(Average)	2.68e-05	0.3398	0.5515	0.8332	0.9185	0.9540	0.9831

specifications when compared by province-age-year-week and by province-year groupings, respectively. For all comparisons, root mean squared error, squared error relative to a simple mean of all observations, and squared error relative to a simpler model that uses the mean observed mortality rate for each age group were generated. I also report empirical coverage of the 50%, 80%, 90%, 95%, and

99% uncertainty intervals of each model type. Because all models are run by sex, sex-specific results as well as averages across male and female models are reported.

Table D.3: Out-of-sample predictive validity metrics comparing model estimates to individual data observations. RMSE: root mean squared error. RSE: relative squared error.

Model	Fourier terms	Seasonality fit by:	Sex	RMSE	RSE	RSE vs. age avg	50% UI	80% UI	90% UI	95% UI	99% UI
A	3	Age, province	Male	0.000976	0.1727	0.9154	0.6172	0.8643	0.9364	0.9688	0.9931
A	3	Age, province	Female	0.000568	0.1004	0.9322	0.6370	0.8733	0.9409	0.9713	0.9937
A	3	Age, province	(Average)	0.000772	0.1365	0.9238	0.6271	0.8688	0.9387	0.9700	0.9934
B	2	Age, province	Male	0.000971	0.1710	0.9067	0.6257	0.8711	0.9406	0.9715	0.9942
B	2	Age, province	Female	0.000566	0.0996	0.9245	0.6441	0.8791	0.9458	0.9741	0.9940
B	2	Age, province	(Average)	0.000769	0.1353	0.9156	0.6349	0.8751	0.9432	0.9728	0.9941
C	1	Age, province	Male	0.000972	0.1715	0.9092	0.6195	0.8657	0.9370	0.9694	0.9931
C	1	Age, province	Female	0.000573	0.1021	0.9484	0.6402	0.8752	0.9416	0.9716	0.9931
C	1	Age, province	(Average)	0.000773	0.1368	0.9288	0.6299	0.8705	0.9393	0.9705	0.9931
D	3	Province	Male	0.000970	0.1705	0.9040	0.6320	0.8736	0.9413	0.9723	0.9939
D	3	Province	Female	0.000582	0.1054	0.9787	0.6521	0.8848	0.9479	0.9755	0.9946
D	3	Province	(Average)	0.000776	0.1380	0.9414	0.6420	0.8792	0.9446	0.9739	0.9942
E	2	Province	Male	0.000965	0.1688	0.8946	0.6269	0.8708	0.9404	0.9709	0.9936
E	2	Province	Female	0.000567	0.0998	0.9262	0.6506	0.8830	0.9467	0.9742	0.9944
E	2	Province	(Average)	0.000766	0.1343	0.9104	0.6388	0.8769	0.9436	0.9725	0.9940
F	1	Province	Male	0.000975	0.1723	0.9134	0.6266	0.8700	0.9391	0.9705	0.9933
F	1	Province	Female	0.000602	0.1127	1.0465	0.6492	0.8815	0.9451	0.9740	0.9939
F	1	Province	(Average)	0.000788	0.1425	0.9799	0.6379	0.8758	0.9421	0.9723	0.9936
G	None	N/A	Male	0.000991	0.1780	0.9437	0.6189	0.8652	0.9362	0.9686	0.9930
G	None	N/A	Female	0.000602	0.1128	1.0471	0.6384	0.8727	0.9406	0.9709	0.9932
G	None	N/A	(Average)	0.000797	0.1454	0.9954	0.6286	0.8689	0.9384	0.9697	0.9931

Table D.4: Out-of-sample predictive validity metrics comparing model estimates to observed data by province-year grouping. RMSE: root mean squared error. RSE: relative squared error.

Model	Fourier terms	Seasonality fit by:	Sex	RMSE	RSE	50% UI	80% UI	90% UI	95% UI	99% UI
A	3	Age, province	Male	2.31e-05	0.3643	0.6703	0.9239	0.9725	0.9874	0.9983
A	3	Age, province	Female	2.51e-05	0.3477	0.6656	0.9125	0.9648	0.9856	0.9969
A	3	Age, province	(Average)	2.41e-05	0.3560	0.6680	0.9182	0.9686	0.9865	0.9976
B	2	Age, province	Male	2.21e-05	0.3332	0.7201	0.9455	0.9798	0.9925	0.9990
B	2	Age, province	Female	2.43e-05	0.3259	0.6975	0.9325	0.9766	0.9900	0.9986
B	2	Age, province	(Average)	2.32e-05	0.3296	0.7088	0.9390	0.9782	0.9912	0.9988
C	1	Age, province	Male	2.30e-05	0.3618	0.6647	0.9188	0.9689	0.9835	0.9971
C	1	Age, province	Female	2.59e-05	0.3691	0.6503	0.9030	0.9602	0.9818	0.9947
C	1	Age, province	(Average)	2.44e-05	0.3654	0.6575	0.9109	0.9646	0.9827	0.9959
D	3	Province	Male	2.22e-05	0.3384	0.7303	0.9477	0.9832	0.9946	0.9993
D	3	Province	Female	2.68e-05	0.3953	0.7140	0.9455	0.9837	0.9934	0.9990
D	3	Province	(Average)	2.45e-05	0.3669	0.7222	0.9466	0.9834	0.9940	0.9992
E	2	Province	Male	2.25e-05	0.3479	0.7062	0.9410	0.9786	0.9922	0.9995
E	2	Province	Female	2.52e-05	0.3488	0.7001	0.9409	0.9815	0.9922	0.9992
E	2	Province	(Average)	2.39e-05	0.3484	0.7031	0.9410	0.9800	0.9922	0.9993
F	1	Province	Male	2.32e-05	0.3687	0.6892	0.9283	0.9715	0.9866	0.9971
F	1	Province	Female	2.88e-05	0.4553	0.6754	0.9201	0.9698	0.9873	0.9976
F	1	Province	(Average)	2.60e-05	0.4120	0.6823	0.9242	0.9706	0.9869	0.9974
G	None	N/A	Male	2.61e-05	0.4647	0.6112	0.8916	0.9545	0.9825	0.9952
G	None	N/A	Female	3.04e-05	0.5101	0.5934	0.8748	0.9456	0.9755	0.9946
G	None	N/A	(Average)	2.82e-05	0.4874	0.6023	0.8832	0.9500	0.9790	0.9949

Based on out-of-sample predictive validity metrics by province-year grouping, as summarised in Table D.4, I selected specification B as the model type that was used in the final analysis. This model

specification had the lowest root mean squared error across both male and female runs, with slightly conservative empirical coverage of the 95% and 99% uncertainty intervals. Out-of-sample predictive validity metrics did not strongly differentiate between the model specifications that included any seasonality terms. However, the model specification that included no seasonality term performed worst in terms of both in-sample and out-of-sample predictive validity.

Excess mortality time series estimation

When calculating a time series of excess mortality among small population groups, the cumulative effect of past mortality on the base population must be taken into account. This problem is most clearly illustrated under artificial conditions of a very high mortality rate. Imagine a population of 1,000 individuals that, under normal conditions, experiences a baseline mortality rate of 100 deaths per 1,000 person-weeks and predictably gains 100 new members at the end of each week. Under normal conditions, 100 individuals from the population would die and 100 new individuals would be added each week, resulting in a stable population week-to-week. Now, imagine an event that causes the mortality rate to increase to 900 deaths per 1,000 person-weeks. In the first week under these conditions, 900 individuals from the base population of 1,000 would die and 100 would be added, leading to a population of 200 entering the second week. In the second week, 180 of 200 individuals in the population would die. Although the standardised mortality ratio of this event compared to baseline is 9 across both weeks (900/100 and 180/20), an analysis that mistakenly assumed a starting population of 1,000 in the second week would dramatically underestimate this ratio as 1.8 (180/100).

This analysis took into account the effects of excess mortality on subsequent population denominators throughout the study weeks beginning on February 26, 2020. Population denominators listed by Istat for January 1, 2020 were used as the starting denominators for this time series. For each predictive posterior draw and week, calculated excess deaths were subtracted from the base population used as the denominator in the subsequent week. The cumulative effect of this correction is to avoid understating the toll of excess mortality over the time period due to a reduction in the base population.

Appendix E

Publications and conference proceedings

Publications related to this thesis

- **Henry, N. J.**, Elagali, A., Nguyen, M., Chipeta, M. G., & Moore, C. E. (2022). Variation in excess all-cause mortality by age, sex, and province during the first wave of the COVID-19 pandemic in Italy. Scientific Reports, 12(1), 1–12. https://doi.org/10.1038/s41598-022-04993-7
- Dandona, R., Kumar, G. A., **Henry, N.**, … Dandona, L. (2020). Subnational mapping of under-5 and neonatal mortality trends in India: the Global Burden of Disease Study 2000–17. *The Lancet*, 395(10237), 1640–1658. https://doi.org/10.1016/S0140-6736(20)30471-2

Publications alongside this thesis

- Alba, S., Rood, E., Mecatti, F., Ross, J. M., Dodd, P. J., Chang, S., Potgieter, M., Bertarelli, G., **Henry, N. J.**, LeGrand, K. E., Trouleau, W., Shaweno, D., MacPherson, P., Qin, Z. Z., Mergenthaler, C., Giardina, F., Augustijn, E.-W., Baloch, A. Q., & Latif, A. (2022). TB Hackathon: Development and Comparison of Five Models to Predict Subnational Tuberculosis Prevalence in Pakistan. Tropical Medicine and Infectious Disease, 7(1), 13. https://doi.org/10.3390/tropicalmed7010013
- Browne, A. J., Chipeta, M. G., Haines-Woodhouse, G., Kumaran, E. P. A., Hamadani, B. H. K., Zaraa, S., **Henry, N. J.**, Deshpande, A., Reiner, R. C., Day, N. P. J., Lopez, A. D., Dunachie, S., Moore, C. E., Stergachis, A., Hay, S. I., & Dolecek, C. (2021). Global antibiotic

consumption and usage in humans, 2000–18: a spatial modelling study. The Lancet Planetary Health, 5(12), e893–e904. https://doi.org/10.1016/s2542-5196(21)00280-1

- Cork, M., **Henry, N.**, Watson, S., … Dwyer-Lindgren, L. A. (2021). Mapping subnational HIV mortality in six Latin American countries with incomplete vital registration systems. *BMC Medicine*, 19(4), 1–25. https://doi.org/10.1186/s12916-020-01876-4
- **Henry, N.**, Burstein, R. (equal contribution), Collison, M. L., … Hay, S. I. (2019). Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature*, 574(7778), 353–358. https://doi.org/10.1038/s41586-019-1545-0
- Kyu, H. H., Maddison, E. R., **Henry, N.**, … Murray, C. J. L. (2018). Global, regional, and national burden of tuberculosis, 1990–2016: results from the Global Burden of Diseases, Injuries, and Risk Factors 2016 Study. *The Lancet Infectious Diseases*, 18(12), 1329–1349. https://doi.org/10.1016/S1473-3099(18)30625-X

Conference presentations

- *Variation in COVID-19 excess mortality by age, sex, and province within Italy.* Conference talk. Presented to the American Society of Tropical Medicine and Hygiene 2020 Conference, on-line (18 November 2020).
- *Subnational burden of tuberculosis across Pakistan: the KIT TB Hackathon* (with Ross, J., LeGrand, K.). Conference lightning talk. Presented at the 50th Union World Conference on Lung Health, Hyderabad, India (1 November 2019).
- *Modelling under-5 mortality across low- and middle-income countries: methods and results.* Invited talk. Presented at the Demographic and Health Surveys Program, Rockville, MD USA (18 September 2019).
- *Incorporating civil registration and vital statistics data into geospatial analyses of child mortality.* Conference poster. Presented at the Institute for Disease Modelling 2019 Symposium, Bellevue, WA USA (1-4 April 2019).

- *High resolution mapping of global child mortality.* Conference talk. Presented at the American Society of Tropical Medicine and Hygiene 2018 Conference, New Orleans, LA USA (31 October 2018).

Appendix F

Code repositories

All code used for the data cleaning, modelling, and production of figures is available on on-line in the following GitHub repositories:

- **Chapter Two**: https://github.com/njhenry/thesis_nmr_joint_model
- **Chapter Three**: https://github.com/njhenry/thesis_tb_joint_model
- **Chapter Four**: https://github.com/njhenry/thesis_u5mr
- **Chapter Five**: https://github.com/njhenry/covidemr
- Additionally, the code used to format this thesis can be found on-line at https://github.com/njhenry/thesis.

Bibliography

1. UN General Assembly. *Universal declaration of human rights* 1948.
2. Burstein, R. *et al.* Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature* **574,** 353–358. http://www.nature.com/articles/s41586-019-1545-0 (Oct. 2019).
3. Dwyer-Lindgren, L. *et al.* Variation in life expectancy and mortality by cause among neighbourhoods in King County, WA, USA, 1990-2014: a census tract-level analysis for the Global Burden of Disease Study 2015. *The Lancet Public Health* **2,** e400–e410. http://dx.doi.org/10.1016/S2468-2667(17)30165-2 (2017).
4. Ruger, J. P. Ethics and governance of global health inequalities. *Journal of Epidemiology and Community Health* **60,** 998–1003 (2006).
5. Brown, T. M., Cueto, M. & Fee, E. The World Health Organization and the transition from international to global public health. *American Journal of Public Health* **96,** 62–72 (2006).
6. World Health Organization. *Everybody's business: strengthening health systems to improve health outcomes: WHO's framework for action.* tech. rep. (World Health Organization, Geneva, 2007), 1–44.
7. World Health Organization. *Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies* tech. rep. 1 (2010), 1–92. http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.35.021103.105711.
8. Roberts, M. J., Hsiao, W., Berman, P. & Reich, M. R. *Getting health reform right: a guide to improving performance and equity* (2008).
9. AbouZahr, C. *et al.* Towards universal civil registration and vital statistics systems: The time is now. *The Lancet* **386,** 1407–1418. http://dx.doi.org/10.1016/S0140-6736(15)60170-2 (2015).
10. Rao, C. Elements of a strategic approach for strengthening national mortality statistics programmes. *BMJ Global Health* **4,** 1–10 (2019).
11. Diggle, P. J. & Giorgi, E. Model-based geostatistics for prevalence mapping in low-resource settings. *Journal of the American Statistical Association* **111,** 1096–1120. arXiv: 1505.06891. http://dx.doi.org/10.1080/01621459.2015.1123158 (2016).
12. Mikkelsen, L. & Lopez, A. D. *Improving the quality and use of birth, death and cause-of-death information* tech. rep. (2010), 1–92. http://www.who.int/healthinfo/tool_cod_2010.pdf.
13. Setel, P. W. *et al.* A scandal of invisibility: making everyone count by counting everyone. *Lancet* **370,** 1569–1577 (2007).
14. United Nations Statistics Division. *Principles and Recommendations for a Vital Statistics System Revision 3* **19.** http://unstats.un.org/unsd/Demographic/standmeth/principles/M19Rev3en.pdf (2014).
15. Fisker, A. B., Rodrigues, A. & Helleringer, S. Differences in barriers to birth and death registration in Guinea-Bissau: implications for monitoring national and global health objectives. *Tropical Medicine and International Health* **24,** 166–174 (2019).
16. Hernández, B. *et al.* Subregistro de defunciones de menores y certificación de nacimiento en una muestra representativa de los 101 municipios con más bajo índice de desarrollo humano en México. *Salud Pública de México* **54,** 393–400 (2012).

17. Mahapatra, P. *et al.* Civil registration systems and vital statistics: successes and missed opportunities. *Lancet* **370,** 1653–1663 (2007).
18. Roth, G. A. *et al.* Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **392,** 1736–1788. https://linkinghub.elsevier.com/retrieve/pii/S0140673618322037 (Nov. 2018).
19. Johnson, S. C. *et al.* Public health utility of cause of death data: applying empirical algorithms to improve data quality. *BMC Medical Informatics and Decision Making* **21,** 1–20 (2021).
20. Blake, J. B. The Early History of Vital Statistics in Massachusetts. *Bulletin of the History of Medicine* **29,** 46–68. https://www.jstor.org/stable/44449136 (1955).
21. AbouZahr, C. *et al.* The COVID-19 pandemic: Effects on civil registration of births and deaths and on availability and utility of vital events data. *American Journal of Public Health* **111,** 1123–1131 (2021).
22. Vlieg, W. L. *et al.* Comparing national infectious disease surveillance systems: China and the Netherlands. *BMC Public Health* **17,** 1–9 (2017).
23. Thacker, S. B., Berkelman, R. L. & Stroup, D. F. The Science of Public Health Surveillance. *Journal of Public Health Policy* **10,** 187–203. http://www.jstor.com/stable/3342679 (1989).
24. Mauch, V. *et al.* Structure and management of tuberculosis control programs in fragile states-Afghanistan, DR Congo, Haiti, Somalia. *Health Policy* **96,** 118–127. http://dx.doi.org/10.1016/j.healthpol.2010.01.003 (2010).
25. Uplekar, M. *et al.* Mandatory tuberculosis case notification in high tuberculosis-incidence countries: Policy and practice. *European Respiratory Journal* **48,** 1571–1581. http://dx.doi.org/10.1183/13993003.00956-2016 (2016).
26. Rood, E. *et al.* A spatial analysis framework to monitor and accelerate progress towards SDG 3 to end TB in Bangladesh. *ISPRS International Journal of Geo-Information* **8,** 1–11 (2019).
27. Johns Hopkins Center for Health Security. *Global Health Security Index: 2019 Report* tech. rep. (Johns Hopkins University, 2019), 270. www.ghsindex.org.
28. Bagherian, H., Farahbakhsh, M., Rabiei, R., Moghaddasi, H. & Asadi, F. National communicable disease surveillance system: A review on information and organizational structures in developed countries. *Acta Informatica Medica* **25,** 271–276 (2017).
29. Gold, J. A. *et al.* COVID-19 Case Surveillance: Trends in Person-Level Case Data Completeness, United States, April 5-September 30, 2020. *Public Health Reports* **136,** 466–474 (2021).
30. Krieger, N., Testa, C., Hanage, W. P. & Chen, J. T. US racial and ethnic data for COVID-19 cases: still missing in action. *The Lancet* **396,** e81. http://dx.doi.org/10.1016/S0140-6736(20)32220-0 (2020).
31. Kwak, N., Hwang, S. S. & Yima, A. J. Effect of COVID-19 on Tuberculosis Notification, South Korea. *Emerging Infectious Diseases* **26,** 2506–2508 (2020).
32. Corsi, D. J., Neuman, M., Finlay, J. E. & Subramanian, S. V. Demographic and health surveys: A profile. *International Journal of Epidemiology* **41,** 1602–1613 (2012).
33. Khan, S. & Hancioglu, A. Multiple Indicator Cluster Surveys: Delivering Robust Data on Children and Women across the Globe. *Studies in Family Planning* **50,** 279–286 (2019).
34. Dandona, R., Pandey, A. & Dandona, L. A review of national health surveys in India. *Bulletin of the World Health Organization* **94,** 286–296A (2016).
35. Kumar, G. A., Dandona, L. & Dandona, R. Completeness of death registration in the Civil Registration System, India (2005 to 2015). *Indian Journal of Medical Research* **149,** 740. http://www.ncbi.nlm.nih.gov/pubmed/23144490%20http://www.ijmr.org.in/text.asp?2019/149/6/740/265955 (2019).
36. Banister, J. & Hill, K. Mortality in China 1964-2000. *Population Studies* **58,** 55–75 (2004).
37. Zeng, X. *et al.* Measuring the completeness of death registration in 2844 Chinese counties in 2018. *BMC medicine* **18,** 176 (2020).

38. He, C. *et al.* National and subnational all-cause and cause-specific child mortality in China, 1996-2015: a systematic analysis with implications for the Sustainable Development Goals. *The Lancet Global Health* **5,** e186–e197 (2017).
39. World Health Organization. *SCORE for health data technical package: global report on health data systems and capacity, 2020* tech. rep. (World Health Organization, Geneva, 2021), 1–104. https://www.who.int/publications/global-report-on-health-data-systems-and-capacity-2020.
40. Makinde, O. A. *et al.* Death registration in Nigeria: a systematic literature review of its performance and challenges. *Global Health Action* **13.** https://doi.org/10.1080/16549716.2020.1811476 (2020).
41. McElreath, R. *Statistical Rethinking* (Taylor & Francis, Boca Raton, FL, 2016).
42. Tobler, W. R. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* **46,** 234–240 (1970).
43. Riebler, A. *et al.* An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research* **25,** 1145–1165. arXiv: 1601.01180 (2016).
44. Besag, J., York, J. & Mollié, A. A Bayesian image restoration with two applications in spatial statistics Ann Inst Statist Math 43: 1-59. *Find this article online* **43,** 1–20 (1991).
45. Lindgren, F. & Rue, H. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society. Series B* **73,** 423–498 (2011).
46. Fay, R. E. & Herriot, R. A. Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association* **74,** 269 (1979).
47. MacNab, Y. C. On Gaussian Markov random fields and Bayesian disease mapping. *Statistical Methods in Medical Research* **20,** 49–68 (2011).
48. Oliver, M. A. in *Handbook of Applied Spatial Analysis: Software Tools, Methods, and Applications* (eds Fischer, M. M. & Getis, A.) 319–352 (Springer, 2010). https://link.springer.com/book/10.1007/978-3-642-03647-7.
49. Miller, D. L., Glennie, R. & Seaton, A. E. Understanding the Stochastic Partial Differential Equation Approach to Smoothing. *Journal of Agricultural, Biological, and Environmental Statistics* **25,** 1–16. arXiv: 2001.07623. https://doi.org/10.1007/s13253-019-00377-z (2020).
50. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **71,** 319–392 (2009).
51. Krainski, E. *et al. Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA* **September** (2018).
52. Mercer, L. D. *et al.* Space-time smoothing of complex survey data: Small area estimation for child mortality. *Annals of Applied Statistics* **9,** 1889–1905 (2015).
53. Hay, S. I. *et al.* Global mapping of infectious disease. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368** (2013).
54. Bhatt, S. *et al.* Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *Journal of the Royal Society Interface* **14.** arXiv: 1612.03278 (2017).
55. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* **4,** 2–5 (2017).
56. Kulldorff, M. A spatial scan statistic. *Communications in Statistics - Theory and Methods* **26,** 1481–1496 (1997).
57. Banerjee, S., Carlin, B. P. & Gelfand, A. E. *Hierarchical modeling and analysis for spatial data* 2nd Editio (CRC Press, 2014).
58. Pigott, D. M. *et al.* Prioritising infectious disease mapping. *PLoS Neglected Tropical Diseases* **9,** 1–21 (2015).
59. Liang, J. *et al.* Maternal mortality ratios in 2852 Chinese counties, 1996-2015, and achievement of Millennium Development Goal 5 in China: a subnational analysis of the Global Burden of

Disease Study 2016. *The Lancet* **393,** 241–252. http://dx.doi.org/10.1016/S0140-6736(18)31712-4 (2019).

60. Weiss, D. J. *et al.* Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000-17: a spatial and temporal modelling study. *The Lancet* **394,** 322–331. http://dx.doi.org/10.1016/S0140-6736(19)31097-9 (2019).
61. Nguyen, M. *et al. Mapping malaria seasonality: a case study from Madagascar* 2019. arXiv: 1901.10782. http://arxiv.org/abs/1901.10782.
62. Osgood-Zimmerman, A. *et al.* Mapping child growth failure in Africa between 2000 and 2015. *Nature* **555,** 41–47. http://www.nature.com/articles/nature25760 (Mar. 2018).
63. Maina, J. *et al.* A spatial database of health facilities managed by the public health sector in sub Saharan Africa. *Scientific Data* **6,** 1–8 (2019).
64. Wakefield, J., Okonek, T. & Pedersen, J. Small Area Estimation for Disease Prevalence Mapping. *International Statistical Review,* insr.12400. https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12400 (July 2020).
65. Ericksen, P. *et al. Mapping hotspots of climate change and food insecurity in the global tropics* tech. rep. June (CGIAR Research Program on Climate Change, 2011). http://www.dfid.gov.uk/r4d/PDF/Outputs/CCAFS/ccafsreport5-climate_hotspots_final.pdf%5C%0Ahttp://cgspace.cgiar.org/bitstream/handle/10568/3826/ccafsreport5-climate_hotspots_final.pdf?sequence=13.
66. Thomson, D. R. *et al.* Extending Data for Urban Health Decision-Making: a Menu of New and Potential Neighborhood-Level Health Determinants Datasets in LMICs. *Journal of Urban Health* **96,** 514–536 (2019).
67. Zhang, X. *et al.* Multilevel regression and poststratification for small-area estimation of population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. *American Journal of Epidemiology* **179,** 1025–1033 (2014).
68. Papoila, A. L. *et al.* Stomach cancer incidence in Southern Portugal 1998-2006: A spatio-temporal analysis. *Biometrical Journal* **56,** 403–415 (2014).
69. Dwyer-Lindgren, L. *et al.* US county-level trends in mortality rates for major causes of death, 1980-2014. *JAMA - Journal of the American Medical Association* **316,** 2385–2401 (2016).
70. Weinberger, D. M. *et al.* Estimation of Excess Deaths Associated with the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Internal Medicine* **06520,** E1–E9 (2020).
71. Patil, A. P., Gething, P. W., Piel, F. B. & Hay, S. I. Bayesian geostatistics in health cartography: The perspective of malaria. *Trends in Parasitology* **27,** 246–253 (2011).
72. Elwood, S. Beyond cooptation or resistance: Urban spatial politics, community organizations, and GIS-based spatial narratives. *Annals of the Association of American Geographers* **96,** 323–341 (2006).
73. Tichenor, M. & Sridhar, D. Metric partnerships: Global burden of disease estimates within the World Bank, the World Health Organisation and the Institute for Health Metrics and Evaluation. *Wellcome Open Research* **4** (2020).
74. Cinnamon, J. Data inequalities and why they matter for development. *Information Technology for Development* **26,** 214–233 (2020).
75. Marchais, G., Bazuzi, P. & Amani Lameke, A. The data is gold, and we are the gold-diggers': whiteness, race and contemporary academic research in eastern DRC. *Critical African Studies* **12,** 372–394. https://doi.org/10.1080/21681392.2020.1724806 (2020).
76. Crampton, J. W. & Krygier, J. An introduction to critical cartography. *Acme* **4,** 11–33 (2006).
77. Tiffin, N., George, A. & Lefevre, A. E. How to use relevant data for maximal benefit with minimal risk: Digital health data governance to protect vulnerable populations in low-income and middle-income countries. *BMJ Global Health* **4,** 1–9 (2019).
78. Osgood-Zimmerman, A. & Wakefield, J. *A Statistical Introduction to Template Model Builder: A Flexible Tool for Spatial Modeling* 2021. arXiv: 2103.09929. http://arxiv.org/abs/2103.09929.

79. Schmertmann, C. P. & Gonzaga, M. R. Bayesian Estimation of Age-Specific Mortality and Life Expectancy for Small Areas With Defective Vital Records. *Demography* **55,** 1363–1388 (2018).
80. Shannon, J. & Walker, K. Opening GIScience: A process-based approach. *International Journal of Geographical Information Science* **32,** 1911–1926. https://doi.org/10.1080/13658816.2018.1464167 (2018).
81. Abouzahr, C. & Boerma, T. Health information systems: the foundations of public health. *Bulletin of the World Health Organization* **83,** 578–583. arXiv: arXiv:1011.1669v3 (2005).
82. Adair, T. & Lopez, A. D. Estimating the completeness of death registration: An empirical method. *PLoS ONE* **13,** 1–19 (2018).
83. Målqvist, M. *et al.* Unreported births and deaths, a severe obstacle for improved neonatal survival in low-income countries; a population based study. *BMC International Health and Human Rights* **8,** 1–7 (2008).
84. Mikkelsen, L. *et al.* A global assessment of civil registration and vital statistics systems: Monitoring data quality and progress. *The Lancet* **386,** 1395–1406. http://dx.doi.org/10.1016/S0140-6736(15)60171-4 (2015).
85. Murray, C. J., Rajaratnam, J. K., Marcus, J., Laakso, T. & Lopez, A. D. What can we conclude from death registration? Improved methods for evaluating completeness. *PLoS Medicine* **7** (2010).
86. Silva, R. *A Preliminary Evaluation of the Completeness and Quality of Death Registration Data in Selected Arab Countries* tech. rep. (Economic and Social Commission for Western Asia, United Nations, 2015), 1–20.
87. Bhat, P. N. Completeness of India's sample registration system: An assessment using the general growth balance method. *Population Studies* **56,** 119–134 (2002).
88. UN Inter-agency Group on Mortality Estimation (UN IGME). *Levels and trends in child mortality report 2020* tech. rep. (United Nations, New York, 2020). https://www.unicef.org/reports/levels-and-trends-child-mortality-report-2020.
89. Dicker, D. *et al.* Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **392,** 1684–1735. https://linkinghub.elsevier.com/retrieve/pii/S0140673618318919 (Nov. 2018).
90. Frenk, J. Bridging the divide: global lessons from evidence-based health policy in Mexico. *Lancet* **368,** 954–961 (2006).
91. Schmid, B. & da Silva, N. N. Estimation of live birth underreporting with a capturerecapture method, Sergipe, northeastern Brazil. *Revista de Saude Publica* **45,** 1088–1098 (2011).
92. De Frias, P. G., Szwarcwald, C. L., de Souza, P. R. B., da Silva de Almeida, W. & Lira, P. I. C. Correcting vital information: Estimating infant mortality, Brazil, 2000-2009. *Revista de Saude Publica* **47,** 1048–1058 (2013).
93. Szwarcwald, C. L., de Frias, P. G., deSouza Júnior, P. R. B., da Silva de Almeida, W. & de Morais Neto, O. L. Correction of vital statistics based on a proactive search of deaths and live births: Evidence from a study of the North and Northeast regions of Brazil. *Population Health Metrics* **12,** 1–10 (2014).
94. Lima, E. E., Queiroz, B. L. & Zeman, K. Completeness of birth registration in Brazil: an overview of methods and data sources. *Genus* **74** (2018).
95. Ribotta, B. S., Acosta, L. M. S. & Bertone, C. L. Evaluaciones subnacionales de la cobertura de las estadísticas vitales. Estudios recientes en América Latina [Evaluations of the Vital Statistics Coverage at a Subnational Level. Recent Studies in Latin America]. *Revista Gerencia y Políticas de Salud* **18** (2019).
96. Smith, P. J. Bayesian Methods for Multiple Capture-Recapture Surveys. *Biometrics* **44,** 1177–1189. http://www.jstor.com/stable/2531745 (1988).
97. Chandra Sekar, C. & Deming, E. On a Method of Estimating Birth and Death Rates and the Extent of Registration. *Journal of the American Statistical Association* **44,** 101–115 (1949).

98. Tilling, K., Sterne, J. A. & Wolfe, C. D. Estimation of the incidence of stroke using a capture-recapture model including covariates. *International Journal of Epidemiology* **30,** 1351–1359 (2001).
99. Hook, E. B. & Regal, R. R. Capture-recapture methods in epidemiology: Methods and limitations. *Epidemiologic Reviews* **17,** 243–264 (1995).
100. Yip, S. F. *et al.* Capture-recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology* **142,** 1047–1058 (1995).
101. Becker, S., Waheeb, Y., El-deeb, B., Khallaf, N. & Black, R. Estimating the Completeness of Under-5 Death Registration in Egypt. *Demograph* **33,** 329–339. https://www.jstor.org/stable/2061765 (1996).
102. Tilling, K. Capture-recapture methods - Useful or misleading? *International Journal of Epidemiology* **30,** 12–14 (2001).
103. Cormack, R. M. Problems with using capture-recapture in epidemiology: An example of a measles epidemic. *Journal of Clinical Epidemiology* **52,** 909–914 (1999).
104. De Almeida, W. D. S. *et al.* Captação de óbitos não informados ao ministério da saúde: Pesquisa de busca ativa de óbitos em municípios brasileiros [Capturing deaths not informed to the Ministry of Health: proactive search of deaths in Brazilian municipalities]. *Revista Brasileira de Epidemiologia* **20,** 200–211 (2017).
105. National Adminstrative Department of Statistics (DANE). *La Mortalidad Materna y Perinatal en Colombia en los albores del siglo XXI. Estimación del subregistro de nacimientos y defunciones y estimaciones ajustadas de nacimientos, mortalidad materna y perinatal por departamentos* tech. rep. (2006), 46. http://docplayer.es/40107048-Estudio-la-mortalidad-materna-y-perinatal-en-colombia-en-los-albores-del-siglo-xxi.html.
106. Schmertmann, C. P. & Gonzaga, M. R. Bayesian Estimation of Age-Specific Mortality and Life Expectancy for Small Areas With Defective Vital Records. *Demography* **55,** 1363–1388 (2018).
107. Golding, N. *et al.* Mapping under-5 and neonatal mortality in Africa, 2000-15: a baseline analysis for the Sustainable Development Goals. *The Lancet* **390,** 2171–2182. http://dx.doi.org/10.1016/S0140-6736(17)31758-0 (2017).
108. Wakefield, J. *et al.* Estimating under-five mortality in space and time in a developing world context. *Statistical Methods in Medical Research* **28,** 2614–2634 (2019).
109. National Institute of Statistics and Geography (INEGI) (Mexico). *Mexico Vital Registration 2009-2010* Aguascalientes, Mexico. https://en.www.inegi.org.mx/programas/mortalidad/.
110. National Institute of Statistics and Geography (INEGI) (Mexico). *Mexico Population and Housing Census 2010* Aguascalientes, Mexico, 2010. https://en.www.inegi.org.mx/programas/ccpv/2010/.
111. National Institute of Statistics and Geography (INEGI) (Mexico). *Geostatistical framework 2010 version 5.0 (Population and Housing Census 2010)* Aguascalientes, Mexico, 2010. http://en.www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825292812.
112. Enciso, G. F., Del Pilar Ochoa Torres, M. & Hernández, J. A. M. The subsystem of information on births. Case study of an indigenous region of Chiapas, Mexico. *Estudios Demográficos y Urbanos* **32,** 451–486 (2017).
113. Paulino, N. A., Vázquez, M. S. & Bolúmar, F. Indigenous language and inequitable maternal health care, Guatemala, Mexico, Peru and the Plurinational State of Bolivia. *Bulletin of the World Health Organization* **97,** 59–67. http://www.who.int/entity/bulletin/volumes/97/1/18-216184.pdf (Jan. 2019).
114. Gamlin, J. & Osrin, D. Preventable infant deaths, lone births and lack of registration in Mexican indigenous communities: health care services and the afterlife of colonialism. *Ethnicity and Health* **25,** 925–939. https://doi.org/10.1080/13557858.2018.1481496 (2020).
115. Dickman, P. W., Sloggett, A., Hills, M. & Hakulinen, T. Regression models for relative survival. *Statistics in Medicine* **23,** 51–64 (2004).

116. Luis, N. B. *El subregistro de nacimiento en los grupos vulnerables de México: análisis de las políticas públicas implementadas para abatirlo entre los años 2010 y 2014* PhD thesis (INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY, 2014), 1–121.
117. Simpson, D., Rue, H., Riebler, A., Martins, T. G. & Sørbye, S. H. Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science* **32,** 1–28. arXiv: 1403.4630 (2017).
118. Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H. & Bell, B. M. TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* **70.** arXiv: 1509.00660 (2016).
119. Thorson, J. T. & Kristensen, K. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research* **175,** 66–74 (2016).
120. R Core Team. *R: A Language and Environment for Statistical Computing* Vienna, Austria, 2018. https://www.r-project.org/.
121. Divino, F., Egidi, V. & Salvatore, M. A. Geographical mortality patterns in Italy: A Bayesian analysis. *Demographic Research* **20,** 435–466 (2009).
122. Gakidou, E., Cowling, K., Lozano, R. & Murray, C. J. Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: A systematic analysis. *The Lancet* **376,** 959–974. http://dx.doi.org/10.1016/S0140-6736(10)61257-3 (2010).
123. Chalasani, S. Understanding wealth-based inequalities in child health in India: A decomposition approach. *Social Science and Medicine* **75,** 2160–2169. http://dx.doi.org/10.1016/j.socscimed.2012.08.012 (2012).
124. Rutherford, M. E., Mulholland, K. & Hill, P. C. How access to health care relates to under-five mortality in sub-Saharan Africa: Systematic review. *Tropical Medicine and International Health* **15,** 508–519 (2010).
125. Dwyer-Lindgren, L. *et al.* Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature* **570,** 189–193. http://www.nature.com/articles/s41586-019-1200-9 (June 2019).
126. Dehnavieh, R. *et al.* The District Health Information System (DHIS2): A literature review and meta-synthesis of its strengths and operational challenges based on the experiences of 11 countries. *Health information management : journal of the Health Information Management Association of Australia* **48,** 62–75 (2019).
127. Onozaki, I. & Raviglione, M. Stopping tuberculosis in the 21st century: goals and strategies. *Respirology* **15,** 32–43 (2010).
128. World Health Organization. *The End TB Strategy* tech. rep. (World Health Organization, Geneva, 2015). https://www.who.int/teams/global-tuberculosis-programme/the-end-tb-strategy.
129. World Health Organization. *Global Tuberculosis Report 2020* tech. rep. (World Health Organization, Geneva, 2020), 1–208. https://www.who.int/publications/i/item/9789240013131.
130. World Health Organization. *Implementing the End TB Strategy* tech. rep. (World Health Organization, Geneva, 2015). https://www.who.int/teams/global-tuberculosis-programme/the-end-tb-strategy.
131. World Health Organization. *Framework and standards for country health information systems (2E)* tech. rep. (World Health Organization, Geneva, 2012), 1–63. http://www.who.int/healthmetrics/documents/hmn_framework200803.pdf.
132. World Health Organization. *Definitions and reporting framework for tuberculosis: 2013 revision* tech. rep. (World Health Organization, Geneva, 2020), 1–40. https://apps.who.int/iris/handle/10665/79199.
133. Kyu, H. H. *et al.* Global, regional, and national burden of tuberculosis, 1990-2016: results from the Global Burden of Diseases, Injuries, and Risk Factors 2016 Study. *The Lancet Infectious Diseases* **18,** 1329–1349. https://linkinghub.elsevier.com/retrieve/pii/S147330991830625X (Dec. 2018).

134. Uganda National Tuberculosis and Leprosy Programme. *National Tuberculosis and Leprosy Control Programme: revised national strategic plan 2015/16-2019/20* tech. rep. (Uganda Ministry of Health, Kampala, 2017), 1–52. file:///C:/Users/nyomb/Downloads/Revised-NTLP-Stragetic-Plan-2015-2020_25th-Jul2017-final.pdf.
135. Uganda National Tuberculosis and Leprosy Programme. *Uganda National TB and Leprosy Program: July 2019-June 2020 report* tech. rep. October 2020 (Uganda Ministry of Health, Kampala, 2020), 1–72.
136. Glaziou, P. & Floyd, K. *Latest developments in WHO estimates of TB disease burden* tech. rep. (WHO Global Task Force on TB Impact Measurement, Glion, 2018), 1–13.
137. Glaziou, P., Van Der Werf, M. J., Onozaki, I. & Dye, C. Tuberculosis prevalence surveys: Rationale and cost. *International Journal of Tuberculosis and Lung Disease* **12,** 1003–1008 (2008).
138. Karamagi, E. *et al.* Improving TB case notification in northern Uganda: evidence of a quality improvement-guided active case finding intervention. *BMC Health Services Research* **18,** 954. https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-018-3786-2 (Dec. 2018).
139. Shaweno, D. *et al.* Methods used in the spatial analysis of tuberculosis epidemiology: A systematic review. *BMC Medicine* **16,** 1–18 (2018).
140. Van Gurp, M. *et al.* Finding gaps in TB notifications: spatial analysis of geographical patterns of TB notifications, associations with TB program efforts and social determinants of TB risk in Bangladesh, Nepal and Pakistan. *BMC Infectious Diseases* **20,** 1–14 (2020).
141. Sudhakar, M. & Admassu, B. Determinants of help seeking and treatment seeking behavior of Tuberculosis patients - Gender perspective: A systematic review. *JBI Database of Systematic Reviews and Implementation Reports* **11,** 199–271 (2013).
142. Abebe, G. *et al.* Knowledge, health seeking behavior and perceived stigma towards tuberculosis among tuberculosis suspects in a rural community in Southwest Ethiopia. *PLoS ONE* **5,** 1–7 (2010).
143. Ahsan, G. *et al.* Gender difference in treatment seeking behaviors of tuberculosis cases in rural communities of Bangladesh. *Southeast Asian Journal of Tropical Medicine and Public Health* **35,** 126–135 (2004).
144. Kirirabwa, N. S. *et al.* A four-year trend in pulmonary bacteriologically confirmed tuberculosis case detection in Kampala-Uganda. *BMC Pulmonary Medicine* **19,** 1–8 (2019).
145. Uganda Ministry of Health. *The Uganda National Tuberculosis Prevalence Survey, 2014-2015: survey report* tech. rep. (Uganda Ministry of Health, Kampala, 2015), 1–162. https://www.health.go.ug/cause/the-uganda-national-tuberculosis-prevalence-survey-2014-2015-survey-report/.
146. Uganda National Tuberculosis and Leprosy Programme. *Manual for management and control of tuberculosis and leprosy (3E)* tech. rep. (Uganda Ministry of Health, Kampala, 2017), 1–162.
147. Uganda National Tuberculosis and Leprosy Programme. *National Tuberculosis and Leprosy Division: July 2017-June 2018 report* tech. rep. (Uganda Ministry of Health, Kampala, 2018), 1–49.
148. Uganda National Tuberculosis and Leprosy Programme. *National strategic plan for tuberculosis and leprosy control: 2020/21-2024/25* tech. rep. November 2017 (Uganda Ministry of Health, Kampala, 2020), 1–106.
149. Weiss, D. J. *et al.* Global maps of travel time to healthcare facilities. *Nature Medicine* **26,** 1835–1838 (2020).
150. Freeman, J. & Hutchison, G. B. Prevalence, incidence, and duration. *American Journal of Epidemiology* **112,** 707–723. https://academic.oup.com/aje/article/78756/PREVALENCE, (Nov. 1980).

151. World Health Organization. *Guidelines for treatment of drug-susceptible tuberculosis and patient care* 1–56. https://apps.who.int/iris/bitstream/handle/10665/255052/9789241550000-eng.pdf (World Health Organization, Geneva, 2017).
152. Tiemersma, E. W., van der Werf, M. J., Borgdorff, M. W., Williams, B. G. & Nagelkerke, N. J. Natural history of tuberculosis: Duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: A systematic review. *PLoS ONE* **6** (2011).
153. Payne, B. & Bellamy, R. Managing tuberculosis in people with HIV. *BMJ Clinical Evidence* (2007).
154. World Health Organization. *WHO consolidated guidelines on tuberculosis. Module 4: drug-resistant tuberculosis treatment* tech. rep. (World Health Organization, Geneva, 2020), 1–98. https://www.who.int/publications/i/item/9789240007048.
155. Falzon, D. *et al.* WHO guidelines for the programmatic management of drug-resistant tuberculosis: 2011 update. *European Respiratory Journal* **38,** 516–528 (2011).
156. Ledesma, J. R. *et al.* Global, regional, and national sex differences in the global burden of tuberculosis by HIV status, 1990-2019: results from the Global Burden of Disease Study 2019. *The Lancet Infectious Diseases* **3099,** 1–20 (2021).
157. Fullman, N. *et al.* Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. *The Lancet* **391,** 2236–2271. https://linkinghub.elsevier.com/retrieve/pii/S0140673618309942 (June 2018).
158. Glaziou, P., Dodd, P. J., Dean, A. S. & Floyd, K. *Methods used by WHO to estimate the global burden of TB disease: global tuberculosis report 2020* tech. rep. October (World Health Organization, Geneva, 2020). https://cdn.who.int/media/docs/default-source/hq-tuberculosis/global-tuberculosis-report-2020/tb2020_technical_appendix_20201014.pdf.
159. Buregyeya, E., Criel, B., Nuwaha, F. & Colebunders, R. Delays in diagnosis and treatment of pulmonary tuberculosis in Wakiso and Mukono districts, Uganda. *BMC Public Health* **14,** 1–10 (2014).
160. Glaziou, P., Dodd, P. J., Dean, A. S. & Floyd, K. *Methods used by WHO to estimate the global burden of TB disease: global tuberculosis report 2019* tech. rep. October (World Health Organization, Geneva, 2019). https://www.who.int/docs/default-source/documents/tuberculosis/technical-appendix-global-tb-report-2019methods-used-by-who-to-estimate-the-global-burden-of-tb-disease.pdf?sfvrsn=e9c1d7b0_2.
161. Glaziou, P., Dodd, P., Zignol, M., Sismanidis, C. & Floyd, K. *Methods used by WHO to estimate the global burden of TB disease: global tuberculosis report 2018* tech. rep. September (World Health Organization, Geneva, 2018), 1–36. https://www.who.int/tb/publications/global_report/gtbr2018_online_technical_appendix_global_disease_burden_estimation.pdf.
162. Togun, T., Kampmann, B., Stoker, N. G. & Lipman, M. Anticipating the impact of the COVID-19 pandemic on TB patients and TB control programmes. *Annals of Clinical Microbiology and Antimicrobials* **19,** 1–6. https://doi.org/10.1186/s12941-020-00363-1 (2020).
163. Shaweno, D., Trauer, J. M., Denholm, J. T. & McBryde, E. S. A novel Bayesian geospatial method for estimating tuberculosis incidence reveals many missed TB cases in Ethiopia. *BMC Infectious Diseases* **17,** 1–8 (2017).
164. Van Hest, R., Grant, A. & Abubakar, I. Quality assessment of capture-recapture studies in resource-limited countries. *Tropical Medicine and International Health* **16,** 1019–1041 (2011).
165. Lucas, T. C. D. *et al.* Model ensembles with different response variables for base and meta models: malaria disaggregation regression combining prevalence and incidence data. *BioRxiv,* 2–7. arXiv: 548719 (2019).
166. World Health Organization. *Global Tuberculosis Report 2019* 1–283 (World Health Organization, Geneva, 2019).

167. Uganda National Tuberculosis and Leprosy Programme. *National Tuberculosis and Leprosy Programme: annual report for 2017* tech. rep. (Uganda Ministry of Health, Kampala, 2017), 1–72.
168. Government of India Ministry of Health and Family Welfare. *National Health Policy 2017* tech. rep. (Government of India, New Delhi, 2017), 28. https://www.nhp.gov.in/nhpfiles/national_health_policy_2017.pdf.
169. Sundararaman, T. National Health Policy 2017: a cautious welcome. *Indian Journal of Medical Ethics* **2,** 69–71 (2017).
170. Government of India Ministry of Health and Family Welfare. *Situation Analyses: Backdrop to the National Health Policy 2017* tech. rep. (Government of India, New Delhi, 2017), 14.
171. Yadav, S. & Arokiasamy, P. Understanding epidemiological transition in India. *Global Health Action* **7** (2014).
172. Dandona, L. *et al.* Nations within a nation: variations in epidemiological transition across the states of India, 1990-2016 in the Global Burden of Disease Study. *The Lancet* **390,** 2437–2460 (2017).
173. Parliament of the Republic of India. *The Registration of Births and Deaths Act, 1969* New Delhi, 1969.
174. Subramanian, S. A Census and Statistics Act for India. *Calcutta Statistical Association Bulletin* **18,** 85–96. http://journals.sagepub.com/doi/10.1177/0008068319690204 (June 1969).
175. Office of the Registrar General & Census Commissioner. *Sample Registration System Statistical Report 2017* tech. rep. (Ministry of Home Affairs, Government of India, New Delhi, 2017), 337. http://www.censusindia.gov.in/vital_statistics/SRS_Report/9Chap%202%20-%202011.pdf.
176. Mahapatra, P. *An Overview of the Sample Registration System in India* in *Prince Mahidol Award Conference & Global Health Information Forum* (Institute of Health Systems, Bangkok, 2010), 1–13. http://www.ihs.org.in/IHS/PMahapatra-SRSInIndiaAnOverview.pdf.
177. Abouzahr, C. *et al.* The way forward. *Lancet Health Metrics Network, World Health Organization* **370,** 1791–99 (2007).
178. Mohanty, I. & Gebremedhin, T. A. Maternal autonomy and birth registration in India: Who gets counted? *PLoS ONE* **13,** 1–19 (2018).
179. Dandona, R. *et al.* Subnational mapping of under-5 and neonatal mortality trends in India: the Global Burden of Disease Study 2000-17. *The Lancet* **395,** 1640–1658 (2020).
180. Ahmad, O. B., Lopez, A. D. & Inoue, M. The decline in child mortality: a reappraisal. *Bulletin of the World Health Organization* **78,** 1175–91. http://www.ncbi.nlm.nih.gov/pubmed/11100613%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2560617 (2000).
181. Huang, H. C., Martinez, F., Mateu, J. & Montes, F. Model comparison and selection for stationary space-time models. *Computational Statistics and Data Analysis* **51,** 4577–4596 (2007).
182. Local Burden of Disease HIV Collaborators. Mapping subnational HIV mortality in six Latin American countries with incomplete vital registration systems. *BMC Medicine* **19,** 1–25. https://link.springer.com/article/10.1186/s12916-020-01876-4%20https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-020-01876-4 (Dec. 2021).
183. Marquez, N. *ar.matrix: Simulate Auto Regressive Data from Precision Matricies* 2018. https://cran.r-project.org/package=ar.matrix.
184. Ghosh, R. Child mortality in India: A complex situation. *World Journal of Pediatrics* **8,** 11–18 (2012).
185. Zomuanpuii, R. *et al.* Epidemiology of malaria and chloroquine resistance in Mizoram, northeastern India, a malaria-endemic region bordering Myanmar. *Malaria Journal* **19,** 1–11. https://doi.org/10.1186/s12936-020-03170-3 (2020).
186. Pletcher, S. D. Model fitting and hypothesis testing for age-specific mortality data. *Journal of Evolutionary Biology* **12,** 430–439 (1999).

187. Lucas, T. C. A translucent box: interpretable machine learning in ecology. *Ecological Monographs* **90,** 1–55 (2020).
188. Gupta, M., Rao, C., Lakshmi, P., Prinja, S. & Kumar, R. Estimating mortality using data from civil registration: a cross-sectional study in India. *Bulletin of the World Health Organization* **94,** 10–21 (2016).
189. Kotabagi, R., Chaturvedi, R. & Banerjee, A. Medical certification of cause of death. *Medical Journal Armed Forces India* **60,** 261–272 (2004).
190. Thacker, S. B. & Berkelman, R. L. Public Health Surveillance in the United States. *Epidemiologic Reviews* **10,** 164–190 (1988).
191. Raban, M. Z., Dandona, R. & Dandona, L. Essential health information available for India in the public domain on the internet. *BMC Public Health* **9,** 1–19 (2009).
192. Sebastiani, G., Massa, M. & Riboli, E. Covid-19 epidemic in Italy: evolution, projections and impact of government measures. *European Journal of Epidemiology* **35,** 341–345. https://doi.org/10.1007/s10654-020-00631-6 (2020).
193. Alicandro, G., Remuzzi, G. & La Vecchia, C. Italy's first wave of the COVID-19 pandemic has ended: no excess mortality in May, 2020. *The Lancet* **396,** e27–e28. http://dx.doi.org/10.1016/S0140-6736(20)31865-1 (2020).
194. Institute for Health Metrics and Evaluation (IHME). *COVID-19 Mortality, Infection, Testing, Hospital Resource Use, and Social Distancing Projections* Seattle, WA, 2020. http://www.healthdata.org/covid/data-downloads.
195. La Maestra, S., Abbondandolo, A. & De Flora, S. Epidemiological trends of COVID-19 epidemic in Italy over March 2020: From 1,000 to 100,000 cases. *Journal of Medical Virology* **92,** 1956–1961 (2020).
196. Albitar, O., Ballouze, R., Ooi, J. P. & Sheikh Ghadzi, S. M. Risk factors for mortality among COVID-19 patients. *Diabetes Research and Clinical Practice* **166,** 108293. https://doi.org/10.1016/j.diabres.2020.108293 (2020).
197. Vestergaard, L. S. *et al.* Excess all-cause mortality during the COVID-19 pandemic in Europe - preliminary pooled estimates from the EuroMOMO network, March to April 2020. *Eurosurveillance* **25.** http://dx.doi.org/10.2807/1560-7917.ES.2020.25.26.2001214 (2020).
198. Pasquariello, P. & Stranges, S. Excess mortality from COVID-19: a commentary on the Italian experience. *International Journal of Public Health* **65,** 529–531. https://doi.org/10.1007/s00038-020-01399-y (2020).
199. Mannucci, E., Nreu, B. & Monami, M. Factors associated with increased all-cause mortality during the COVID-19 pandemic in Italy. *International Journal of Infectious Diseases* **98,** 121–124. https://doi.org/10.1016/j.ijid.2020.06.077 (2020).
200. Serfling, R. E. Methods for Current Statistical Analysis of Excess Pneumonia-Influenza Deaths. *Public Health Reports (1896-1970)* **78,** 494 (1963).
201. Noufaily, A. *et al.* An improved algorithm for outbreak detection in multiple surveillance systems. *Statistics in Medicine* **32,** 1206–1222 (2013).
202. Michelozzi, P. *et al.* Temporal dynamics in total excess mortality and COVID-19 deaths in Italian cities. *BMC Public Health* **20,** 1238. https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-020-09335-8 (Dec. 2020).
203. Lambert, P. C., Smith, L. K., Jones, D. R. & Botha, J. L. Additive and multiplicative covariate regression models for relative survival incorporating fractional polynomials for time-dependent effects. *Statistics in Medicine* **24,** 3871–3885 (2005).
204. Wu, J., McCann, A., Katz, J., Peltier, E. & Singh, K. D. *Tracking the True Toll of the Coronavirus Outbreak* 2020. https://nyti.ms/2WbTKbS (2020).
205. The Economist. *COVID-19 Data: Tracking COVID-19 excess deaths across countries* 2020. https://econ.st/3oNlNL0 (2020).
206. FT Visual & Data Journalism team. *FT Coronavirus Tracker* 2020. https://on.ft.com/37ie1Dj (2020).

207. U.S. National Center for Health Statistics. *Excess Deaths Associated with COVID-19* 2021. https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm (2021).
208. Italian National Institute of Statistics (Istat). *All-cause mortality by Italian municipality, January 2015 through August 2020 (release: 22 October 2020).* Rome, 2020. https://www.istat.it/it/archivio/240401.
209. Italian National Institute of Statistics (Istat). *Resident municipal population by age, sex and marital status* Rome, 2020. https://dati.istat.it.
210. Dowd, J. B. *et al.* Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences of the United States of America* **117,** 9696–9698 (2020).
211. Cox, D. R. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* **34,** 187–220 (1972).
212. Ederer, F., Axtell, L. M. & Cutler, S. J. The Relative Survival Rate: A Statistical Methodology. *National Cancer Institute Monographs* **6,** 101–121. https://pubmed.ncbi.nlm.nih.gov/13889176/ (1961).
213. Thorson, J. T. & Barnett, L. A. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science* **74,** 1311–1321 (2017).
214. Woolf, S. H., Chapman, D. A., Sabo, R. T., Weinberger, D. M. & Hill, L. Excess Deaths From COVID-19 and Other Causes, March-April 2020. *JAMA* **324,** 510. https://jamanetwork.com/journals/jama/fullarticle/2768086 (Aug. 2020).
215. Yu, W., Vaneckova, P., Mengersen, K., Pan, X. & Tong, S. Is the association between temperature and mortality modified by age, gender and socio-economic status? *Science of the Total Environment* **408,** 3513–3518. http://dx.doi.org/10.1016/j.scitotenv.2010.04.058 (2010).
216. Stafoggia, M. *et al.* Factors affecting in-hospital heat-related mortality: A multi-city case-crossover analysis. *Journal of Epidemiology and Community Health* **62,** 209–215 (2008).
217. Bayer, R. & Fairchild, A. L. Surveillance and privacy. *Science* **290,** 1898–1899. https://www.sciencemag.org/lookup/doi/10.1126/science.290.5498.1898 (Dec. 2000).
218. León Cabrera, J. M. & Kurmanaev, A. *Ecuador's Death Toll During Outbreak Is Among the Worst in the World* New York, Apr. 2020. https://nyti.ms/2WbTKbS.
219. Duryea, S., Olgiati, A. & Stone, L. *The Under-Registration of Births in Latin America* 2006.
220. Boing, A. F., Boing, A. C., Cordes, J., Kim, R. & Subramanian, S. V. Quantifying and explaining variation in life expectancy at census tract, county, and state levels in the United States. *Proceedings of the National Academy of Sciences of the United States of America* **117,** 17688–17694 (2020).
221. Büyüm, A. M., Kenney, C., Koris, A., Mkumba, L. & Raveendran, Y. Decolonising global health: If not now, when? *BMJ Global Health* **5,** 1–4 (2020).
222. Graetz, N. *et al.* Mapping disparities in education across low- and middle-income countries. *Nature* **577,** 235–238 (2020).
223. Kinyoki, D. K. *et al.* Mapping child growth failure across low- and middle-income countries. *Nature* **577,** 231–234 (2020).