Discussion

Nathaniel Henry

In this thesis, I develop a class of methods to make health surveillance data interpretable and relevant to policy at the local level. To do so, I extend traditional spatial statistical modeling approaches to synthesize two or more health data sources, compensating for the limitations of each individual 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. I demonstrate that this framework can also be applied to understand health surveillance gaps in high-income countries, which previous research has often treated as complete.

This discussion summarizes 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.

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. Chapter 2 identifies the core issue of latent, spatially-varying bias as a barrier to utilization 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[1–5], as well as the limitations of these approaches for spatial analysis.[6–8] 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[8,9] by estimating the ratio

between survey-based and routinely-collected data sources, subject to priors about the rela-
tionship between these sources. I apply this model to estimate the neonatal mortality rate
(NMR) by municipality in Mexico. Context-specific priors for CRVS bias were applied based
on previous research that demonstrated birth and death under-reporting in indigenous and
socially-marginalized communities across Mexico.[10–13]

While strong priors are appropriate to a data context where health surveillance is nearly
complete nationwide, the same class of model can applied in an infectious disease con-
text where case notification completeness is known to be largely incomplete and variable
across the country. In the context of global tuberculosis (TB) control, disaggregated esti-
mates of TB prevalence and incidence are highly desirable for national TB control programs
(NTPs) as a tool for targeting interventions and identifying treatment gaps.[14] In Chapter
3, I explore the TB data context in Uganda, which is representative of the situation in
many high-burden countries: while TB case notifications from routine diagnosis are col-
lected 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 cas-
cade, making unadjusted notifications an inappropriate proxy for subnational variation in
TB incidence.[15,16] A national TB prevalence survey conducted in 2014-2015 does not suf-
fer from the same underreporting bias, but this surveys was not powered for subnational
estimation.[17] 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.[UgandaNationalTuberculosisandLeprosyProgramme2020a]

While reporting completeness is a crucial measure of health data quality, routine surveil-
lance 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 Sys-
tem proposes that CRVS (and other health surveillance data) should also meet stringent
standards for timeliness, accuracy in reporting outcomes, and confidentiality.[18] In addition
to their utility for policy, CRVS birth and death certification enable legal rights for covered
individuals.[19,20] 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 disaggre-

gated analysis of neonatal, infant, and child mortality using a novel space-time-age modeling framework.[21,22] This analysis demonstrates the value of spatial modeling 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.[23] Previous analyses have underlined the variable completeness of the SRS at the state level.[24,25] In this chapter, I demonstrates 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.[26,27] 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.[28]

Previous geospatial and small-area research has drawn a distinction between incomplete and complete data systems. Research in low- and middle-income countries has often assumed incompleteness in health surveillance systems, requiring epidemiologists to develop corrections for reporting or otherwise limit their analyses.[16,29,30] Conversely, in high-income countries, spatial analyses often treat health surveillance as complete, needing correction only for stochastic fluctuations associated with small sample sizes.[31,32] 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 cause of death reporting for COVID-19 missed 30% of the true disease burden nationwide, and 42% of true disease 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. A wider array of tools is needed, including the models proposed in this thesis, to correct for underlying biases in health surveillance data at the local level.

2 Methodological discussion

This section reviews the statistical methods used across the chapters, highlighting strengths and limitations of this class of models.

2.1 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.[33–35] Measurement facilitates management: tracking health status across social and geographical groupings can identify inequity, prompting responsive health policy and reform.[36,37] 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,[9,38] they can be extended to measure differences across other dimensions such as age group, as demonstrated in Chapters 4 and 5, without any adjustments to correct for stochastic noise.

Country-specific models of health can be parameterized 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,[10,11] 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,[39] and was expected to outweigh bias in modeled 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 etiology 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 low- and middle-income contexts, small-area modeling offers can supplement existing approaches that have been used to estimate health surveillance data completeness. Local audits have been developed to estimate the completeness of CRVS data and infectious disease notifications;[4,10,40–42] While these studies ofer 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 is preferred over the previous capture-recapture method, which makes unrealistically strong assumptions about independence between data sources.[6,7]

Finally, the Bayesian hierarchical modeling approach deployed here propagates uncertainty from input data and parameter fits in a principled way, allowing for nuanced analysis about the confidence of results.[43] 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 visualized 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 realizations.[44]

2.2   Methodological limitations

Despite their flexibility, this class of semiparametric 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 formalize the assumption that unobserved risk factors and covariates are likely to vary according to an underlying spatial structure: in other words, two neighboring 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,[38] 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:[9] if strong priors are inaccurate, they can bias the model in the direction of existing expert opinion. 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.

Separate from model implementation, 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. In fact, uncertainty is arguably the most important aspect of small area models. New communication strategies are needed to intuitively demonstrate variability across many possible realizations of these spatial models.[44]

2.3 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, modeling, 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 maximize 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 visualizations can also facilitate the uptake of surveillance-based disease maps as policy tools: importing modeled 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 modeled 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 survey's cluster sampling design can influence knowledge about the disease distribution for years to come; future survey designs could consider relative uncertainty in modeled preva-

lence in order to place survey clusters in a configuration that maximizes 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 modeling 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 formalizing this geographical knowledge into bespoke spatial statistical models, public health researchers can reveal the scope of health inequalities at the local level.

3   Conclusions

This thesis aims to redefine the role of modeling in countries with 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 their conception, policymakers interpret models as a superior alternative to imperfect health surveillance data, reducing the momentum to improve CRVS quality.[19,45] 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, 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 emphasized that CRVS systems are more than a data source: they are a process for guaranteeing legal and civil rights to a country's citizens.[46]

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 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 subpopulations. This thesis strengthens the link between developing national health surveillance systems and improved health outcomes, leading to better health policies and a strong basis for human rights worldwide.

4 References

1. 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).

2. 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).

3. 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 (1996).

4. Frias, P. G. de, Szwarcwald, C. L., Souza, P. R. B. de, 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).

5. Murray, C. J. L., 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).

6. Tilling, K. Capture-recapture methods - Useful or misleading? International Journal of Epidemiology 30, 12–14 (2001).

7. Cormack, R. M. Problems with using capture-recapture in epidemiology: An example of a measles epidemic. Journal of Clinical Epidemiology 52, 909–914 (1999).

8. 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).

9. 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).

10. 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).

11. 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).

8

12.   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).

13.   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 (2019).

14.   Glaziou, P. & Floyd, K. Latest developments in WHO estimates of TB disease burden. 1–13 (2018).

15.   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).

16.   Shaweno, D. et al. Methods used in the spatial analysis of tuberculosis epidemiology: A systematic review. BMC Medicine 16, 1–18 (2018).

17.   Uganda Ministry of Health. The Uganda National Tuberculosis Prevalence Survey, 2014-2015: survey report. 1–162 https://www.health.go.ug/cause/the-uganda-national-tuberculosis-prevalence-survey-2014-2015-survey-report/ (2015).

18.   United Nations Statistics Division. Principles and Recommendations for a Vital Statistics System Revision 3. (2014).

19.   Setel, P. W. et al. A scandal of invisibility: making everyone count by counting everyone. Lancet 370, 1569–1577 (2007).

20.   Duryea, S., Olgiati, A. & Stone, L. The Under-Registration of Births in Latin America. 25 (2006) doi:10.2139/ssrn.1820031.

21.   Dandona, R., Pandey, A. & Dandona, L. A review of national health surveys in India. Bulletin of the World Health Organization 94, 286–296A (2016).

22.   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).

9

23. Office of the Registrar General & Census Commissioner. Sample Registration System Statistical Report 2017. 337 http://www.censusindia.gov.in/vital_statistics/SRS_Report/9Chap%202%20-%202011.pdf (2017).

24. Bhat, P. N. M. Completeness of India's sample registration system: An assessment using the general growth balance method. Population Studies 56, 119–134 (2002).

25. Mahapatra, P. An Overview of the Sample Registration System in India. in Prince mahidol award conference & global health information forum 1–13 (Institute of Health Systems, 2010).

26. Parliament of the Republic of India. The Registration of Births and Deaths Act, 1969. vols Act 18 (1969).

27. Mohanty, I. & Gebremedhin, T. A. Maternal autonomy and birth registration in India: Who gets counted? PLoS ONE 13, 1–19 (2018).

28. Abouzahr, C. et al. The way forward. Lancet Health Metrics Network, World Health Organization 370, 1791–99 (2007).

29. Adair, T. & Lopez, A. D. Estimating the completeness of death registration: An empirical method. PLoS ONE 13, 1–19 (2018).

30. Zeng, X. et al. Measuring the completeness of death registration in 2844 Chinese counties in 2018. BMC medicine 18, 176 (2020).

31. Papoila, A. L. et al. Stomach cancer incidence in Southern Portugal 1998-2006: A spatio-temporal analysis. Biometrical Journal 56, 403–415 (2014).

32. 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).

33. Government of India Ministry of Health and Family Welfare. National Health Policy 2017. 28 https://www.nhp.gov.in/nhpfiles/national_health_policy_2017.pdf (2017).

34. Uganda National Tuberculosis and Leprosy Programme. National strategic plan for tuberculosis and leprosy control: 2020/21-2024/25. 1–106 (2020).

35. 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).

36. Roberts, M. J., Hsiao, W., Berman, P. & Reich, M. R. Getting health reform right: a guide to improving performance and equity. (2008).

37. Frenk, J. Bridging the divide: global lessons from evidence-based health policy in Mexico. Lancet 368, 954–961 (2006).

38. Diggle, P. J. & Giorgi, E. Model-based geostatistics for prevalence mapping in low-resource settings. Journal of the American Statistical Association 111, 1096–1120 (2016).

39. Uganda National Tuberculosis and Leprosy Programme. Uganda National TB and Leprosy Program: July 2019-June 2020 report. 1–72 (2020).

40. Szwarcwald, C. L., Frias, P. G. de, Júnior, P. R. B. deSouza, 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).

41. Almeida, W. D. S. de 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).

42. 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. 46 http://docplayer.es/40107048-Estudio-la-mortalidad-materna-y-perinatal-en-colombia-en-los-albores-del-siglo-xxi.html (2006).

43. McElreath, R. Statistical Rethinking. (Taylor & Francis, 2016). doi:10.1080/09332480.2017.1302722.

44. 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).

45.   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).

46.   AbouZahr, C. et al. Towards universal civil registration and vital statistics systems: The time is now. The Lancet 386, 1407–1418 (2015).

325