
Mapping the relationship between tuberculosis burden and case notifications in Uganda

Nathaniel Henry

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
5 Therapy (DOTS, 1994-2005) and Stop TB (2006-2014) plans as the coordinating principles for addressing the tuberculosis epidemic worldwide.¹ 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.² These goals are
10 consistent with Sustainable Development Goal 3.3, which commits United Nations member countries to end the TB epidemic by 2030.³ Timely and accurate TB case detection serves as a critical component for two pillars of the End TB strategy: early detection enables for adequate patient-centered care, while mandatory and universal registration of TB disease allows countries to properly allocate resources and track progress towards goals.⁴

15 To achieve the End TB Strategy, the World Health Organization has developed tools for systematically building health system capacity⁵⁻⁷ and has tracked progress and shared best practices through the release of annual Global Tuberculosis Reports.³ Both these reports and external analyses have concluded that while many countries have increased their capacity to identify and treat tuberculosis, particularly drug-resistant TB and TB among people living
20 with HIV, the majority of high-burden countries are not on track to meet the End TB 2030 targets.^{3,8} The gap between 2030 targets and observed performance can be explained by a combination of social determinants of health and health system factors, including remaining incompleteness in TB case detection in high-burden settings.³

The recent history of TB control in Uganda, one of the 30 high-burden countries recognized
25 by the end TB program, 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’s 2015/16-2019/20 strategic plan mirrors the core objectives of the global End TB plan, calling for a 5% reduction in TB incidence as well as a 30% increase in case notifications over a four-year time period. 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 2019.⁹ Two of the principle challenges to universal TB detection and treatment identified in the 2020 report were (1) stubbornly low case detection rates in the far eastern and southwestern regions of the country; and (2) the threat of service disruptions related to COVID-19, which could reduce health system capacity and adherence to TB treatment regimens.⁹ New approaches are needed to ensure universal access to TB treatment and, ultimately, to eliminate TB nationwide.

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.¹⁰ TB prevalence and incidence 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.¹¹

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 lab tests on the sputum of potential cases to confirm the presence of TB.¹² 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.^{3,12}

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

60 as the DHIS2, reporting and aggregation of notified TB cases may happen continuously.^{6,13} Aggregate case notifications are a potentially appealing source of subnational data due to their broad geographic coverage and wide availability. However, case notifications may be incomplete due to lack of health system access, misdiagnosis, and under-reporting of TB cases from private facilities.¹²

65 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.¹⁴ Of the 168 identified studies, 161 (96%) used TB case notifications as the data underlying the spa-
70 tial analysis, although no studies 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 Gorp 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.¹⁵ The
75 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.

The limitations identified in case notifications present a a major challenge to assessing spatial variability in TB burden across low-resource settings. Given the small sample sizes and low
80 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
85 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.

2 Methods

In this chapter, I develop a geostatistical modeling approach that incorporates both TB prevalence survey data and a time series of case notifications to estimate TB prevalence and
90 notification incompleteness by district across Uganda. I first develop crosswalks between the differing age groups, case definitions, and indicators reported across annual notifications and the national TB prevalence survey. I then present a geostatistical 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
 95 modified by spatially-varying incompleteness.

2.1 Data preparation

The joint geostatistical modeling 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. I downloaded cluster-level results, which included place names
 100 for each cluster, that were published alongside the 2014-15 Uganda National TB Prevalence Survey.¹⁶ Using spatial information from Google Maps and OpenStreetMap, I associated each cluster with the GPS location of the sampled township. 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
 105 sputum samples using both a smear test and a culture test; I included individuals who tested positive for either test to approximate the case definition for a confirmed pulmonary TB case in a clinical setting.¹⁷ Of the 128 districts in Uganda, 50 were associated with at least one cluster from the 2014-15 prevalence survey.

Total case notifications were reported by district across Uganda from 2015 through 2019 in
 110 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
 115 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 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 region to 19% in the Moroto region; meanwhile, the proportion of cases diagnosed in children under age 15
 120 varied from 6% in the Lira region to 19% in the Moroto region.¹⁸ 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
 125 population denominators.¹⁹

In populations where disease burden is relatively stable, the prevalence and incidence of a disease are related by the average duration of the disease: $Prevalence = Incidence *$

*Duration.*²⁰ When treated, the duration of TB is widely accepted to be six months,³ while the average duration of untreated pulmonary TB is estimated to be approximately three years.²¹ In a study of two districts in Uganda, the median delay between the onset of active tuberculosis and the beginning of treatment was estimated at four weeks.²² An estimate for the average duration of disease would combine the experiences of untreated individuals and treated individuals after accounting for treatment delay, and would fall somewhere between the two extremes of six months and three years. In their most recent published estimates that include estimates of both TB prevalence and incidence, both the WHO and the Institute for Health Metrics and Evaluation (IHME) estimate that the ratio between TB prevalence and incidence in Uganda is approximately 1, indicating an estimated average duration of approximately one year.^{3,8} I use this estimate to crosswalk between measures of TB incidence and prevalence in the joint geostatistical model.

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 an estimate of a pastoral lifestyle.²³ Following Van Gorp and colleagues,¹⁵ I included covariates for the reporting completeness surface that are associated with health system access: these included average travel time to health facilities, nighttime light intensity, and urban land cover.^{24,25} All gridded covariates were aggregated to the district level using population weights derived from the WorldPop project.¹⁹

2.2 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). Prevalences varies according to an intercept, α , fixed effects $\overline{\beta^p}$ 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 parameterized using the interpretable BYM2 spatial model formulation described by Riebler and colleagues.²⁶ 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:

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

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 parameterized 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:

$$Y_i^{prev} = \text{Poisson}(N_i^{prev}, p_{s(i)})$$

If only data from the TB prevalence survey was incorporated into this model, it would 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:

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

In this formulation, the observed number of case notifications are estimated to be centered 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, $1/D$, which has been fixed at one. 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^{27,28}. 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. The model was fit in R v.4.0.3 using the package Template Model Builder v.1.7.18^{27,29}.

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 varied from less than 1 in 4 cases reported to near-complete reporting in 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.

3.1 Spatial variation in tuberculosis prevalence across Uganda

Across the 128 districts of Uganda, estimated pulmonary TB prevalence in adults ranged from 140 per 100,000 (95% uncertainty intervals: 134 to 147) in the district of Wakiso bordering Kampala, to a maximum of 1,468 per 100,000 (1,108 to 1,977) in the northeastern district of Napak, which borders the Bokora Wildlife Reserve. As shown in Panel B of Figure 1 Prevalence was generally lowest in the southern districts bordering Lake Victoria and highest in the northern and northwestern regions of the country, mirroring programmatic experience within the NTLP.²³ The densely-populated districts of Kampala and Jinja experienced middling tuberculosis burden relative to the rest of the country, with estimated TB prevalence of 557 (547 to 568) and 537 (423 to 656) per 100,000, respectively.

Panel A of Figure 1 shows the raw prevalence estimates derived by aggregating the 2014-2015 national TB prevalence survey by district, while Panel C shows the relationship between these raw survey-based estimates (on the x axis) and prevalence estimated generated from the joint model (on the y axis). In general, modeled estimates tend to follow the same overall trends as the raw survey data, with higher prevalence observed in the prevalence survey corresponding to higher modeled estimates in a given district. However, the model has substantially smoothed these raw estimates: the highest raw estimate of nearly 2,500 per 100,000 has been pulled down, while many clusters with zero observed TB-positive

individuals have higher model-estimated prevalence. These modifications 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.

220 3.2 Completeness of case notifications increases over time

Figure 2, below, shows model estimates of case notification completeness by district for two years, 2017 and 2019. 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.⁹ In 2017, estimated case notification completeness varied
 225 from 21% (17% to 25%) in the rural Otuke district to 99% (92% to 100%) in Kampala. In 2019, case notification completeness ranged from an estimated 26% (21% to 31%) in the southwestern Kiruhira district to, once again, 99% (93% to 100%) reporting completeness in Kampala. Population was associated with greater reporting completeness, as the urban districts of Kampala, Wakiso, Mukono, Gulu, and Jinja all had an estimated reporting
 230 completeness of 80% or greater in 2019.

The large majority of districts, 107 out of 128, experienced an increase in reporting completeness over the duration of the study, with an average increase of eight percentage points. Although the joint model estimated that the remaining 21 districts experienced a decline in case notifications, this decline was only significant in one district, Kiruhura, which dropped
 235 from 48% completeness (39% to 56%) in 2017 to 26% (21% to 31%) in 2019, the lowest reporting completeness nationally.

3.3 Estimating the TB prevalence to notification ratio (PNR) by district

Based on the fitted model parameters, TB case notification rates were negatively associated with travel time to health facilities, and positively associated with nighttime lights, urban-
 240 icity, and time. In addition, completeness was empirically associated with the number of notifications per prevalent case observed in the raw data. The x axis of Figure 3, below, shows the ratio between the estimated TB prevalence rate and the TB notification rate in the raw data, including only districts where at least one prevalent case was observed in the 2014-15 prevalence survey. The y-axis shows the estimated completeness generated from the
 245 joint model in 2017 (left) and 2019 (right). Although the sample sizes from the prevalence survey are small and must be interpreted with caution, a weakly positive relationship is observed between “empirical reporting completeness” and estimated completeness from the joint model, suggesting that the model is able to indirectly estimate completeness through the relationship between prevalence survey and notification data.

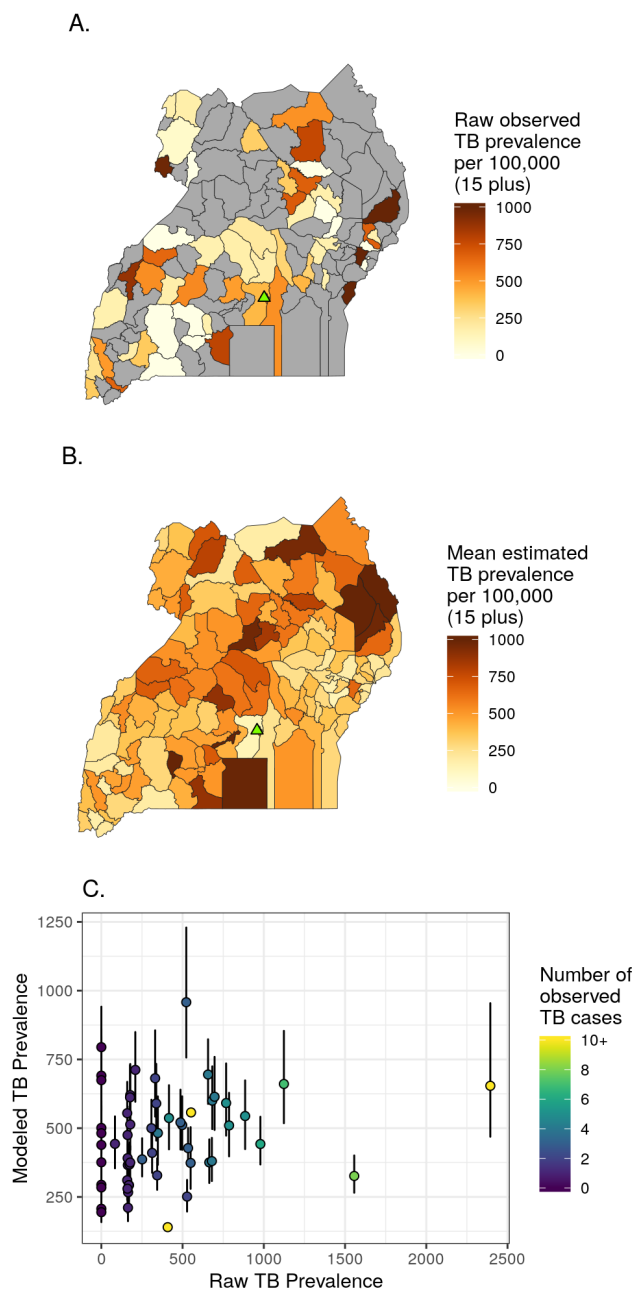


Figure 1: Model results compared to geopositioned prevalence survey data. Panel A: Data from the 2014-2015 Uganda National Tuberculosis Prevalence Survey, geo-positioned and aggregated to the district level. Out of 128 districts, 78 contained no underlying prevalence survey clusters. Panel B: TB prevalence estimated from the joint data model. Panel C: scatter plot between raw TB prevalence survey data and estimated TB prevalence from the joint model, with uncertainty. The model generally smooths extreme data observations across districts.

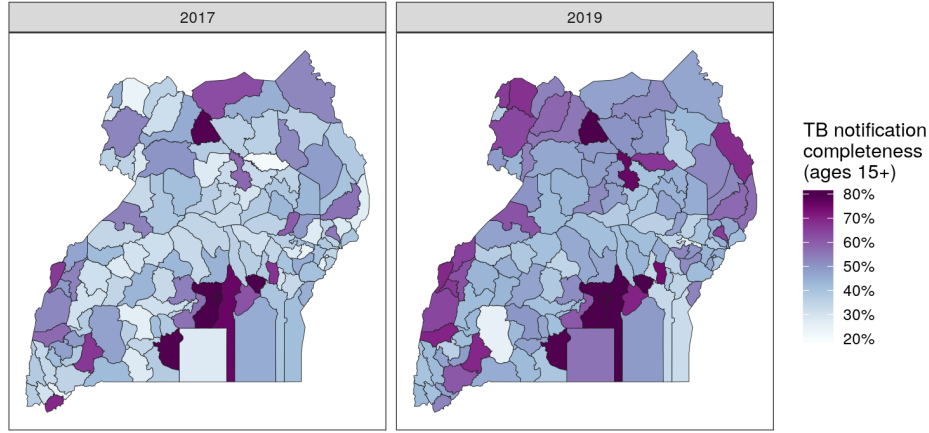


Figure 2: Mean estimated completeness for district-level TB case notifications in 2017 and 2019, as predicted by the joint data model. Completeness is estimated for pulmonary TB cases in people aged 15 and above.

3.4 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 4 below shows the mean estimated TB prevalence based on this model. Compared to Panel A of Figure 1, the prevalence-only model captures and smooths extreme values within districts observed in the prevalence survey, but reverts to values close to 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 13 to 869 deaths per 100,000, with an average spread of 209. 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.

Panels B and C in Figure 4 visualize the effect of indirectly incorporating case notifications into the modeled 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 northeast and decreasing estimated prevalence in the coastal south. However, given the wide uncertainty associated with the prevalence-only estimates, just four districts of 128 experience significant changes between the two model formulations.

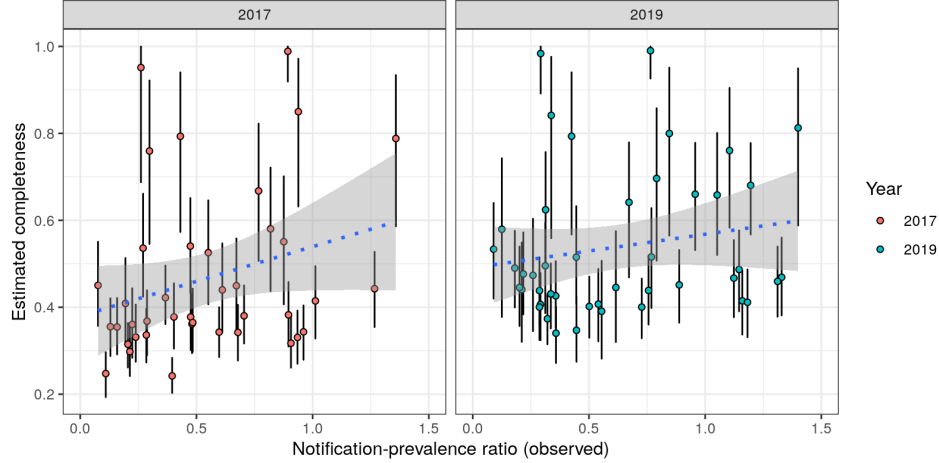


Figure 3: Scatter plot showing the relationship between prevalence data, case notification reporting, and estimates of notification completeness in 2017 and 2019. The x-axis shows the ratio between the number of reported TB cases and raw prevalence from the Uganda National TB Prevalence Survey as an empirical measure of notification completeness. The y-axis displays completeness estimates predicted by the joint model, with 95 percent uncertainty bounds. In general, both empirical and estimated completeness generally rose across districts from 2017 to 2019. Trend lines in both years indicate that increased empirical completeness is weakly positively associated with increased estimated completeness.

4 Discussion

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.¹⁴ In this chapter, I demonstrate how TB case notifications can be used to indirectly inform estimates of TB burden by indirectly fitting a second spatial surface for case notification completeness. While this completeness surface incorporated data from spatial covariates related to health systems access,¹⁵ it also corresponds to the notification-to-prevalence ratio which may be the best proxy for notification completeness in the unadjusted data.

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. Two southern districts, Wakiso and Butaleja, have 95% uncertainty intervals that fall below the prevalence threshold of 200 per 100,000; on the other hand, the districts of Moroto and Napak have 95% UIs that fall entirely above the high prevalence threshold of 1,000 per 100,000. These high-precision

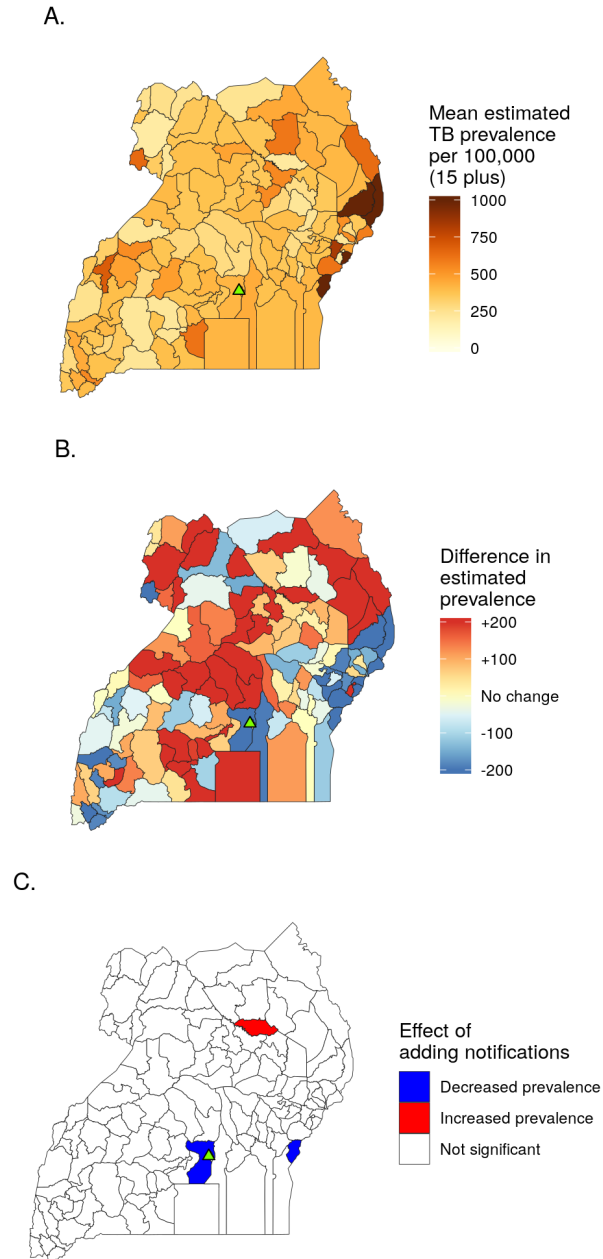


Figure 4: Effect of adding notifications to the joint estimation model. Panel A: Mean estimated TB prevalence based on an equivalent spatial model incorporating only prevalence data from the Uganda National TB Prevalence Survey. Panel B: Difference in mean estimated TB prevalence between the joint data model and a model that only incorporated prevalence data. Panel C: Effect of adding case notifications, taking uncertainty into account. Note that because of the wide uncertainty associated with the prevalence-only model, only four districts had non-overlapping UIs between the two model specifications; all four of these districts were sampled in the national prevalence survey.

identifications have programmatic value, and the differentiation between districts will allow
 285 the NTLP to use these results to allocate resources with greater confidence.

The space-time surface of notification completeness generated by the joint model offers a new
 tool for targeting improvements in TB case-finding. The model identified 32 districts where
 case reporting was less than half complete with 95% certainty; four of these, Kiruhura,
 Butebo, Namayingo, and Mayuge, were predicted with 95% certainty to have estimated
 290 case finding of less than 40%. Measures of notification completeness could typically only
 be estimated through expensive audits in a small number of districts; while limited by the
 model assumptions, this joint modeling approach offers a supplementary method to tracking
 notification completeness that can be estimated nationwide each year. In the near future,
 this model could be applied to 2020 notifications to explore spatially-varying barriers to
 295 accessing TB care that may have arisen in the wake of the COVID-19 pandemic.³⁰

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 2019, 13 countries in Africa and six countries in Asia had completed
 300 just one TB prevalence survey since 2007.³ Many of these countries rely on notifications
 data to track annual change in TB prevalence: by combining these 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
 305 reporting completeness, and this information could be used to develop highly predictive
 spatial covariates for a latent completeness surface.^{11,15,31}

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
 310 be the gold standard for surveillance, but small sample sizes and gaps in temporal coverage
 limit their use for routine surveillance. Previous geospatial studies have combined these
 surveys with routine surveillance data, using an estimated crosswalk between data types;^{32,33}
 these models could be made more flexible using by adapting the joint model approach
 described in this chapter.

4.2 Limitations

While this model estimates 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 TB incidence in Uganda has fallen by less than 5% since 2015, as estimated by the NTLP, WHO, and IHME,^{3,8,9} the joint model assumes that subnational variation in TB prevalence remains constant over the time period 2015-2019. Conversely, case notification reporting completeness was modeled as a space-time surface to capture documented increases in case notifications over the same time period.^{9,34} 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.

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. In the next chapter, the same principle is applied to estimate the quality of cause-specific mortality data in Italy, a country where mortality data is typically considered unbiased and complete.

5 References

1. Onozaki, I. & Raviglione, M. Stopping tuberculosis in the 21st century: goals and strategies. *Respirology* 15, 32–43 (2010).
2. World Health Organization. The End TB Strategy. <https://www.who.int/teams/global-tuberculosis-programme/the-end-tb-strategy> (2015).
3. World Health Organization. Global Tuberculosis Report 2019. 1–283 (World Health Organization, 2019).

- 345 4. World Health Organization. Implementing the End TB Strategy. <https://www.who.int/teams/global-tuberculosis-programme/the-end-tb-strategy> (2015).
5. World Health Organization. Framework and standards for country health information systems (2E). 1–63 http://www.who.int/healthmetrics/documents/hmn_framework200803.pdf (2012).
6. World Health Organization. Definitions and reporting framework for tuberculosis: 2013 revision. 1–40 <https://apps.who.int/iris/handle/10665/79199> (2020).
- 350 7. World Health Organization. Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies. vol. 35 1–92 <http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.35.021103.105711> (2010).
8. 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 (2018).
- 355 9. Uganda National Tuberculosis and Leprosy Programme. Uganda National TB and Leprosy Program: July 2019-June 2020 report. 1–72 (2020).
10. Glaziou, P. & Floyd, K. Latest developments in WHO estimates of TB disease burden. 1–13 (2018).
11. 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).
- 360 12. 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).
13. 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).
- 365 14. Shaweno, D. et al. Methods used in the spatial analysis of tuberculosis epidemiology: A systematic review. *BMC Medicine* 16, 1–18 (2018).

15. Van Gorp, 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).
16. 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).
370
17. Uganda National Tuberculosis and Leprosy Programme. Manual for management and control of tuberculosis and leprosy (3E). 1–162 (2017).
18. Uganda National Tuberculosis and Leprosy Programme. National Tuberculosis and Leprosy Division: July 2017-June 2018 report. 1–49 (2018).
- 375 19. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* 4, 2–5 (2017).
20. Freeman, J. & Hutchinson, G. B. Prevalence, incidence, and duration. *American Journal of Epidemiology* 112, 707–723 (1980).
21. Tiemersma, E. W., Werf, M. J. van der, Borgdorff, M. W., Williams, B. G. & Nagelkerke, N. J. D. Natural history of tuberculosis: Duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: A systematic review. *PLoS ONE* 6, (2011).
380
22. 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).
23. Uganda National Tuberculosis and Leprosy Programme. National strategic plan for tuberculosis and leprosy control: 2020/21-2024/25. 1–106 (2020).
- 385 24. Weiss, D. J. et al. Global maps of travel time to healthcare facilities. *Nature Medicine* 26, 1835–1838 (2020).
25. 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).

26. Riebler, A. et al. An intuitive Bayesian spatial model for disease mapping that
 390 accounts for scaling. *Statistical Methods in Medical Research* 25, 1145–1165 (2016).
27. Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H. & Bell, B. M. TMB: Auto-
 matic Differentiation and Laplace Approximation. *Journal of Statistical Software* 70,
 (2016).
28. 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).
- 395 29. R Core Team. R: A Language and Environment for Statistical Computing. (2018).
30. 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 (2020).
31. Shaweno, D., Trauer, J. M., Denholm, J. T. & McBryde, E. S. A novel Bayesian
 geospatial method for estimating tuberculosis incidence reveals many missed TB
 400 cases in Ethiopia. *BMC Infectious Diseases* 17, 1–8 (2017).
32. Dwyer-Lindgren, L. et al. Mapping HIV prevalence in sub-Saharan Africa between
 2000 and 2017. *Nature* 570, 189–193 (2019).
33. 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 (2019).
- 405 34. Uganda National Tuberculosis and Leprosy Programme. National Tuberculosis and
 Leprosy Programme: annual report for 2017. 1–72 (2017).