

RESEARCH ARTICLE

Open Access

Children as sentinels of tuberculosis transmission: disease mapping of programmatic data



Kenneth S. Gunasekera¹, Jon Zelner², Mercedes C. Becerra³, Carmen Contreras⁴, Molly F. Franke³, Leonid Lecca^{3,4}, Megan B. Murray³, Joshua L. Warren^{5†} and Ted Cohen^{1*†}

Abstract

Background: Identifying hotspots of tuberculosis transmission can inform spatially targeted active case-finding interventions. While national tuberculosis programs maintain notification registers which represent a potential source of data to investigate transmission patterns, high local tuberculosis incidence may not provide a reliable signal for transmission because the population distribution of covariates affecting susceptibility and disease progression may confound the relationship between tuberculosis incidence and transmission. Child cases of tuberculosis and other endemic infectious disease have been observed to provide a signal of their transmission intensity. We assessed whether local overrepresentation of child cases in tuberculosis notification data corresponds to areas where recent transmission events are concentrated.

Methods: We visualized spatial clustering of children < 5 years old notified to Peru's National Tuberculosis Program from two districts of Lima, Peru, from 2005 to 2007 using a log-Gaussian Cox process to model the intensity of the point-referenced child cases. To identify where clustering of child cases was more extreme than expected by chance alone, we mapped all cases from the notification data onto a grid and used a hierarchical Bayesian spatial model to identify grid cells where the proportion of cases among children < 5 years old is greater than expected. Modeling the proportion of child cases allowed us to use the spatial distribution of adult cases to control for unobserved factors that may explain the spatial variability in the distribution of child cases. We compare where young children are overrepresented in case notification data to areas identified as transmission hotspots using molecular epidemiological methods during a prospective study of tuberculosis transmission conducted from 2009 to 2012 in the same setting.

Results: Areas in which childhood tuberculosis cases are overrepresented align with areas of spatial concentration of transmission revealed by molecular epidemiologic methods.

(Continued on next page)

* Correspondence: theodore.cohen@yale.edu

†Joshua L. Warren and Ted Cohen are the co-senior authors.

¹Department of Epidemiology of Microbial Diseases, Yale School of Public Health, 60 College Street, New Haven, CT 06520, USA

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Conclusions: Age-disaggregated notification data can be used to identify hotspots of tuberculosis transmission and suggest local force of infection, providing an easily accessible source of data to target active case-finding intervention.

Keywords: Spatial, Hotspot, Tuberculosis, Modeling, Age-structure, Pediatric, Population surveillance/*methods, Tuberculosis/epidemiology/*prevention & control/transmission

Background

The End TB Strategy's ambitious goals to reduce tuberculosis incidence require new interventions to interrupt transmission [1]. This has led to a renewed interest in active case-finding strategies, in which risk groups are screened to identify infectious individuals before they present to care [2, 3]. Because untargeted community-based active case-finding has not consistently demonstrated population-level benefits [4–7], there has been interest in new practical approaches to focus case-finding to population groups among whom risk is concentrated. One such approach is to target active case-finding to hotspots, areas in which transmission is most intense [8]. While evidence supporting the impact of targeting screening in hotspots is currently limited [9], mathematical modeling suggests that such targeting can produce substantial population-wide reductions in transmission [10, 11].

Conclusive evidence of hotspot transmission typically relies on access to detailed spatial and pathogen genetic data [12–14]. While spatial information is often available in public health reporting systems (e.g., home location), in high-transmission/lower-income settings, resources for genetic sequencing of pathogens are typically only available in research studies. Thus, methods to robustly identify hotspots from routine reporting data would be valuable [15]. However, given that high local rates of tuberculosis notifications may reflect spatially aggregated risk for progression of infection, migration of individuals infected with tuberculosis into the area [16], or spatial heterogeneity in diagnostic capacity [17], finding new ways to probe routine surveillance data to find evidence of local transmission is a priority.

Spatial differences in the age distribution of tuberculosis cases in a single city may provide a signal for local transmission intensity [18]. In locations where disease transmission is more intense, cases are systematically younger than in locations where disease transmission is less intense [19]. We aimed to test this previously posited, but to our knowledge yet untested, idea that areas where children are overrepresented in tuberculosis case notification data are areas where recent transmission events are concentrated. We tested this hypothesis using case notification data from Lima, Peru, where we were able to compare our inference to a prospective

molecular epidemiology study conducted in the same setting several years later [20, 21]. This comparison provided an opportunity to examine whether routinely collected tuberculosis notification data can be used to identify transmission hotspots.

Methods

Study setting and population

We examined data from all tuberculosis cases notified to Peru's National Tuberculosis Program from two of Lima's four health districts, Lima Ciudad and contiguous catchment areas of Lima Este, between January 1, 2005 and December 31, 2007. Patient demographic and clinical information was available within the notification data as well as household address, which was identified on high-resolution maps created using Google Earth. Additional details of the study design and mapping procedures have been described previously [22, 23].

Our interest was in identifying areas in which young children were overrepresented in these routinely collected notification data from 2005 to 2007 and whether they correlated with areas identified as transmission hotspots during a prospective study of tuberculosis transmission conducted from 2009 to 2012 [21]. The latter study included molecular epidemiological characterization of culture-positive cases of drug-susceptible and drug-resistant tuberculosis from adults older than 15 years using 24-loci mycobacterial interspersed repetitive units-variable-number tandem repeats (MIRU-VNTR). Spatial aggregation of *Mycobacterium tuberculosis* (*M.tb*) strains identified by MIRU-VNTR genotype was presumed to indicate transmission.

Data visualization and modeling

We visualized spatial clustering of child cases < 5 years old in the notification data using a log-Gaussian Cox process (LGCP) to model the intensity function driving the point process describing the distribution of child cases. We used the *lgcp* package and defined the Gaussian process with an exponential covariance function and weakly informative priors on all model parameters (details provided in Additional file 1: Supplementary Information) [24]. All data visualization and analysis were performed using R 4.0.1.

Next, we aimed to determine if the clustering of child cases observed in the exploratory maps was more extreme than would be expected by chance alone. Point-level census and covariate data that may explain spatial variability in the distribution of child cases through effect on overall risk were not available for this analysis. Due to the large number of unique spatial locations observed in the data (10,198) and the well-known difficulties associated with using a Gaussian process to analyze point-referenced spatial data when the sample size is large [25], we opted for a method that approximates the point-referenced model while offering computational improvements [26]. Specifically, we overlaid a grid on the convex hull of the case notification data and modeled the proportion of reported tuberculosis cases that occurred among children in each grid cell using a hierarchical Bayesian spatial modeling framework. We chose the grid cell sizes to be small in order to ensure that the risk within each grid cell was homogeneous. We considered multiple sizes in subsequent sensitivity analyses. As the size of the grid cells gets smaller, our approximation to the point-referenced geostatistical model improves.

By modeling the proportion of the tuberculosis cases that were children (as opposed to simply modeling the number of child cases), we used the distribution of adult cases to control for unobserved factors that may explain the spatial variability in the distribution of child cases. Under this modeling framework, we expect that the local proportion of child cases will be higher than the expected proportion of child cases over the entire study area in areas where there is local transmission. The hierarchical model structure allows us to identify where this occurs and allows us to describe the certainty with which the proportion is higher.

To do this, we use a logistic regression framework to model the grid cell-specific proportions such that:

$$Y_i | \theta_i \sim \text{Binomial}(n_i, \theta_i), i = 1, \dots, m$$

$$\ln\left(\frac{\theta_i}{1 - \theta_i}\right) = \mu + \phi_i$$

where Y_i is the number of child cases observed in grid cell i , n_i is the total number of child and adult cases in the grid cell, m is the total number of grid cells, and θ_i represents the proportion of the total cases in the grid cell that are due to children. We define child cases as those < 5 years old and adult cases as those > 15 years old to clearly separate recent infection among young children from more distant infection among adults (expecting that cases among older children and young adults between ages 5 and 15 represent a mix of recent infection and infection that happened earlier in their lives). We model these proportions on the logit scale as a function of an overall mean, μ (fixed effect), and a grid

cell-specific deviation from that mean, ϕ_i (random effect).

We anticipate that the proportion of child cases in grid cells that are close together may be similar. To account for this potential spatial correlation and to obtain spatially smoothed risk estimates, we estimated the ϕ_i parameters using a conditional autoregressive (CAR) model such that:

$$\phi_i | \phi_{-i}, \tau^2, \rho \sim N\left(\frac{\rho \sum_{j=1}^n w_{ij} \phi_j}{\rho \sum_{j=1}^n w_{ij} + 1 - \rho}, \frac{\tau^2}{\rho \sum_{j=1}^n w_{ij} + 1 - \rho}\right)$$

where ϕ_{-i} is the vector of parameters excluding ϕ_i , w_{ij} is equal to one if grid cells i and j share a common border or point and is equal to zero otherwise, τ^2 describes the variability in the ϕ_i parameters, and $\rho \in (0, 1)$ describes their strength of spatial correlation. As a result, this model is flexible enough to accommodate a wide range of spatial patterns as well as the possibility that there is no spatial variability in the proportion of child cases (i.e., τ^2 near zero indicates that all ϕ_i are near zero). Additionally, examining the posterior distributions of ϕ_i allows us to determine if the grid cell proportion differs substantially from the overall mean.

We selected weakly informative prior distributions for all model parameters and used the *CAR.Leroux* function in the *CARBayes* package to obtain posterior samples for all parameters [27]. Details are provided in Additional file 1: Supplementary Information [28]. Using the posterior samples from each ϕ_i , we estimate the posterior probability that ϕ_i is larger than zero, which would suggest recent transmission based on our hypothesis.

Results

Analysis of notification data

Of the total 11,711 notified tuberculosis cases over the study period, there were 332 children < 5 years old, and 10,352 adults > 15 years old. The LGCP modeled intensity of the cases among children < 5 years old is given in Fig. 1.

We fit the hierarchical Bayesian spatial model to the case notification data collected from 2005 to 2007 aggregated into a 200 m × 200 m grid within the convex hull of the data. The model suggested six grid cells in which > 95% of the posterior distribution of the random effect terms were above zero and an additional eight grid cells in which > 90% of the posterior distribution was above zero (Fig. 2). Examination of the posterior estimate of the spatial correlation parameter, ρ , suggested that the excess variability observed in the data was spatially structured (posterior mean 0.75, 95% credible interval 0.24–0.98). Posterior summaries of the remaining parameters are provided in Additional file 1: Table S1.

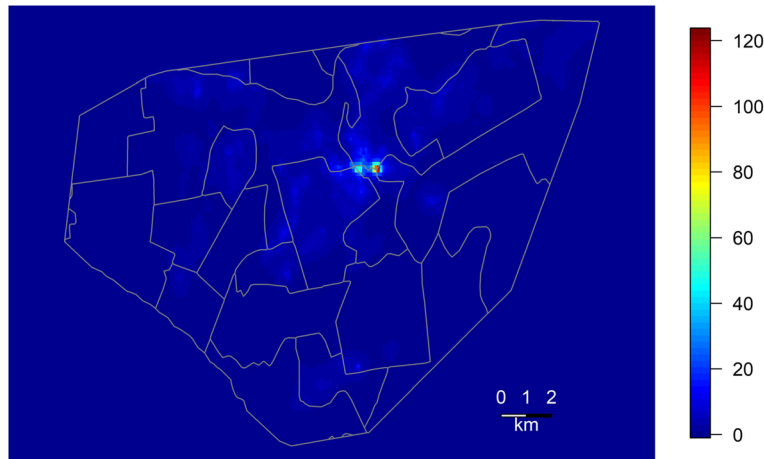


Fig. 1 Disease mapping of young children in Peru's National Tuberculosis Program data. Log-Gaussian Cox process modeled intensity of the cases of tuberculosis among children < 5 years old notified to the Peru's National Tuberculosis Program within two of Lima's four health districts, Lima Ciudad and contiguous catchment areas of Lima Este, between January 1, 2005 and December 31, 2007

Comparison to prospective molecular epidemiological study

Figure 3a, reproduced with permission from Zelner et al., shows areas in which there was statistically significant spatial aggregation of specific *M.tb* MIRU-VNTR genotypes, consistent with localized transmission of these strain types [21]. In Fig. 3b, we overlay the grid from Fig. 2 to demonstrate the proximity between areas where children < 5 years old are overrepresented in case notification data and areas where specific strains are concentrated. In Additional file 1: Figs. S1-S2, we show that these findings are insensitive to assumed grid cell size

and age cutoffs for the definitions of young child and adult cases. Figure 4a, also reproduced with permission from Zelner et al., shows the spatial variation in annual per capita incidence of tuberculosis by healthcare catchment area [21]. We similarly overlay the grid from Fig. 2 to create Fig. 4b to demonstrate the proximity between areas where child cases are overrepresented and high local incidence.

Discussion

In this paper, we evaluated whether routinely collected, age-disaggregated notification data can be used to

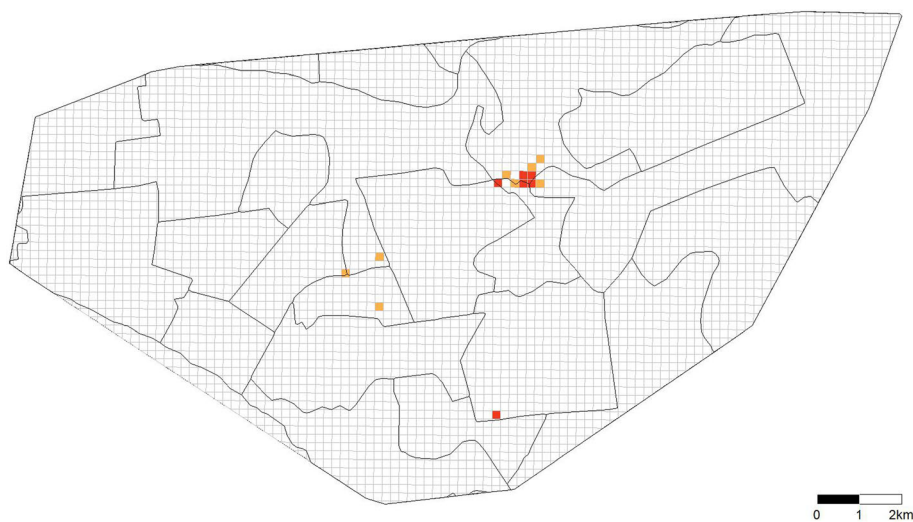
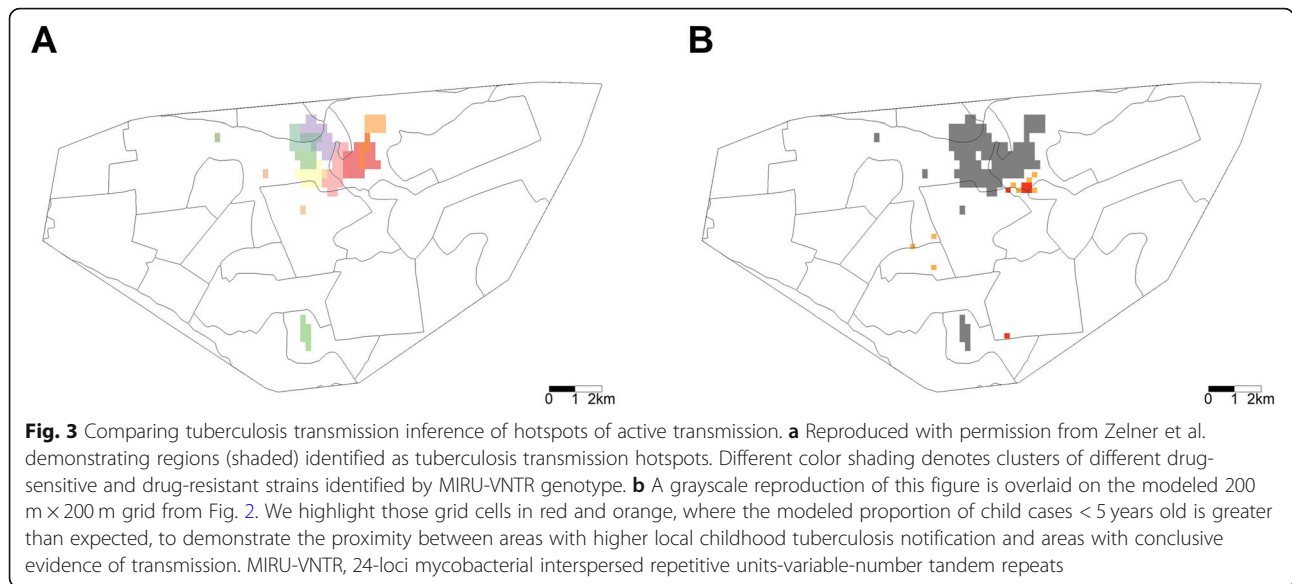


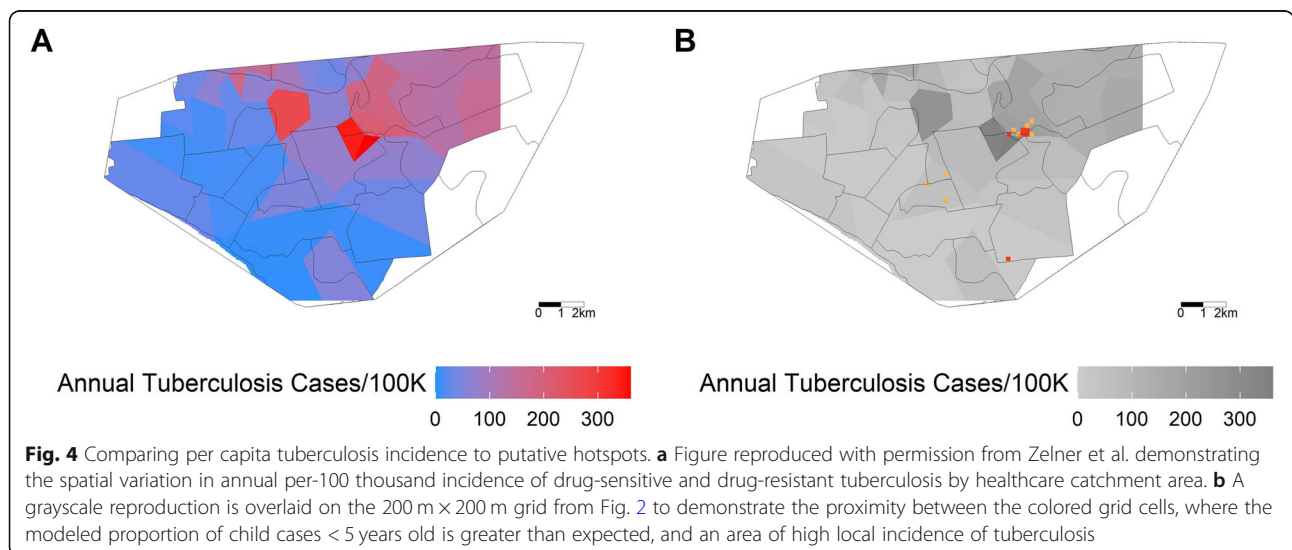
Fig. 2 Identifying areas with local overrepresentation of young children in tuberculosis notification data. Hierarchical Bayesian spatial model fit to the child cases < 5 years old and adult cases > 15 years old in the notification data aggregated into 200 m x 200 m grid cells overlaid on the convex hull of the data. The model suggested six grid cells (red) in which > 95% of the posterior distribution of the random effect terms were above zero, and an additional eight grid cells (orange) in which > 90% of the posterior distribution was above zero. The proportion of child cases in these grid cells is greater than expected over the study region, suggesting recent tuberculosis transmission based on our hypothesis



identify hotspots of spatially concentrated tuberculosis transmission. Our analysis, based on routine data collected from 2005 to 2007, pinpointed a region where child cases of tuberculosis were overrepresented relative to the number of adult cases in the area. This region was previously identified as an area of high transmission using molecular genetic data from a prospective study conducted from 2009 to 2012 [21]. This concordance of transmission inference obtained using different methods and datasets supports the use of routinely collected age-disaggregated notification data to identify areas of local transmission intensity.

Child cases have been suggested as a useful signal of transmission intensity for tuberculosis as well as other

infectious disease [29]. For example, a number of studies used the age prevalence of tuberculin-skin test positivity to measure risks of infection from household and community exposure [30, 31]. Previous studies have suggested that areas with high childhood tuberculosis case notification rates may correspond to areas of active transmission [32–34]; however, only one included covariates to account for potential non-transmission explanations of the spatial distribution of child cases [34]. Thus, our analysis is the first to provide molecular and epidemiological evidence to corroborate inferences of local tuberculosis transmission with attempts to control for unobserved, spatially heterogeneous, non-transmission factors that may explain the distribution of child



cases (such as risk factors for progression of infection, migration of infected individuals into the area, and/or local diagnostic capacity).

Considering that both the routine notification data and the prospective molecular epidemiology study included tuberculosis cases separated by as many as 6 years, we also note that the identified hotspot appears to have been persistent over several years. This suggests that tuberculosis transmission hotspots identified from notification data may be observable for long enough periods of time to guide targeted interventions, such as spatially focused active case-finding.

It is important to note several simplifying assumptions in our analysis. Given the absence of detailed information on the distribution of covariates in the source population, we incorporated all spatial heterogeneity in the distribution of child cases into the random effect term of the model. As a result, our model necessarily attributes all spatial variability in the modeled proportions to possible recent transmission. If there are other non-transmission-related factors that impact the proportion of total cases that occurred in children, this could lead to a grid cell being incorrectly labeled as a transmission “hotspot.” However, given the consistency of our results with the previous findings that more directly measure transmission, this may not be a major issue in this work. Our hierarchical Bayesian spatial modeling approach (as well as the LGCP intensity modeling approach) is flexible enough to incorporate local covariate data as regression components. Future study should include such information when available.

Though we provide compelling evidence, we must be cautious interpreting that age-disaggregated data will always provide a reliable signal of transmission. Molecular evidence of transmission against which we compare transmission inference was only available for those > 15 years old. Thus, we are unable to biologically link childhood cases to the identified clusters of transmission. Furthermore, accurately diagnosing tuberculosis among children is difficult. While it is clear that missing child cases in notification data likely underestimate transmission, it is unclear how false positives may affect signal detection. In addition, though we demonstrate that the putative hotspot persists over time, it is not possible to assess how mobility over the time period through which all data from these two studies was collected may affect hotspot detection. It is important to note that our findings do not imply an either-or choice between genetic and age-incidence data: future analyses exploring the impact of combining granular molecular genetic data with age-incidence data in a single model could improve the predictive capacity of such models.

This methodology may be adapted to settings in which high-resolution residence data is not readily

available. For example, in settings where residential geocoding is not feasible, it may be reasonable to model the proportion of child cases in the smallest recorded unit to which the household belongs (such as modeling the proportion in the neighborhood, community, and/or administrative unit).

Conclusions

In summary, we show that age-disaggregated tuberculosis notification data may be used to investigate potential hotspots of tuberculosis transmission. This suggests that the use of models leveraging widely available data should be explored as tools for targeting case-finding and treatment efforts in high-transmission locations in the hope of maximizing the direct and indirect protective benefits of active screening approaches.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12916-020-01702-x>.

Additional file 1 : Supplementary Information, Table S1, Figure S1-S2. Supplementary Information – Log-Gaussian Cox process details and hierarchical Bayesian spatial model details. **Table S1** – Hierarchical Bayesian spatial model posterior parameter estimates. **Fig S1** – Sensitivity analysis to child and adult age cut-offs. **Fig S2** – Sensitivity analysis to grid size.

Abbreviations

TB: Tuberculosis; MIRU-VNTR: 24-loci mycobacterial interspersed repetitive units-variable-number tandem repeats; *M.tb*: *Mycobacterium tuberculosis*; LGCP: Log-Gaussian Cox process; CAR: Conditional autoregressive

Acknowledgements

We thank the study nurses and Socios En Salud for assistance with data collection and mapping, and the Peruvian Ministry of Health centers for their support for the studies.

Authors' contributions

KSG, JLW, and TC designed the study, analyzed the data, and wrote the first draft of the manuscript. JZL, MCB, CC, MFF, LL, and MBM critically revised the manuscript. MCB and MM coordinated the molecular epidemiological study which verified areas of transmission. All authors read and approved the final manuscript.

Funding

This work was supported by the National Institutes of Health Medical Scientist Training Program Training Grant [T32GM007205 to KSG], the Fogarty International Center Global Health Equity Scholars Program [D43TW010540 to KSG], and the National Institutes of Allergy and Infectious Diseases [U01AI057786 and U19AI076217 supported MCB, CC, MFF, LL, MBM, and TC]. The funding organizations had no role in the design, collection, analysis, and interpretation of data, or in the writing of the manuscript.

Availability of data and materials

Additional data are available on reasonable request to MCB and MM. All requests for data access will need to specify the planned use of data and will require approval from MCB and MM before release.

Ethics approval and consent to participate

The study protocol for the prospective molecular epidemiology investigation was approved by the Harvard University Institutional Review Board (Ref. No. 19332). The study protocol for data collection and spatial analyses of cases notified to the Peru National TB Program from 2005 to 2007 was approved

by the Research Ethics Committee of the National Institute of Health of Peru (Ref. No. 085-2007). Consent was not required as this is a secondary data analysis of previously published, de-identified data.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no conflicts of interest.

Author details

¹Department of Epidemiology of Microbial Diseases, Yale School of Public Health, 60 College Street, New Haven, CT 06520, USA. ²Department of Epidemiology, University of Michigan School of Public Health, 267 SPH Tower, 1415 Washington Heights, Ann Arbor, MI 48109, USA. ³Department of Global Health and Social Medicine, Harvard Medical School, 641 Huntington Avenue, Boston, MA 02115, USA. ⁴Socios En Salud, Lima, Peru. ⁵Department of Biostatistics, Yale School of Public Health, 60 College Street, New Haven, CT 06520, USA.

Received: 1 April 2020 Accepted: 9 July 2020

Published online: 02 September 2020

References

- World Health Organization. The End TB Strategy. Geneva; 2015.
- Houben R, Menzies NA, Sumner T, Huynh GH, Arinaminpathy N, Goldhaber-Fiebert JD, et al. Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models. *Lancet Glob Health*. 2016;4(11):e806–e15.
- Dheda K, Barry CE 3rd, Maartens G. *Tuberc*. *Lancet*. 2016;387(10024):1211–26.
- Kranzer K, Afnan-Holmes H, Tomlin K, Golub JE, Shapiro AE, Schaap A, et al. The benefits to communities and individuals of screening for active tuberculosis disease: a systematic review. *Int J Tuberc Lung Dis*. 2013;17(4):432–46.
- Corbett EL, Bandason T, Duong T, Dauya E, Makamure B, Churchyard GJ, et al. Comparison of two active case-finding strategies for community-based diagnosis of symptomatic smear-positive tuberculosis and control of infectious tuberculosis in Harare, Zimbabwe (DETECTB): a cluster-randomised trial. *Lancet*. 2010;376(9748):1244–53.
- Marks GB, Nguyen NV, Nguyen PTB, Nguyen T-A, Nguyen HB, Tran KH, et al. Community-wide screening for tuberculosis in a high-prevalence setting. *N Engl J Med*. 2019;381(14):1347–57.
- Calligaro GL, Zijenah LS, Peter JG, Theron G, Buser V, McNERNEY R, et al. Effect of new tuberculosis diagnostic technologies on community-based intensified case finding: a multicentre randomised controlled trial. *Lancet Infect Dis*. 2017;17(4):441–50.
- World Health Organization. Systematic screening for active tuberculosis: principles and recommendations. Geneva; 2013.
- Cudahy PGT, Andrews JR, Bilinski A, Dowdy DW, Mathema B, Menzies NA, et al. Spatially targeted screening to reduce tuberculosis transmission in high-incidence settings. *Lancet Infect Dis*. 2019;19(3):e89–95.
- Dowdy DW, Golub JE, Chaisson RE, Saraceni V. Heterogeneity in tuberculosis transmission and the role of geographic hotspots in propagating epidemics. *Proc Natl Acad Sci U S A*. 2012;109(24):9557.
- Hickson RL, Mercer GN, Lokuge KM. A metapopulation model of tuberculosis transmission with a case study from high to low burden areas. *PLoS One*. 2012;7(4):e34411.
- Meehan CJ, Moris P, Kohl TA, Pečerska J, Akter S, Merker M, et al. The relationship between transmission time and clustering methods in *Mycobacterium tuberculosis* epidemiology. *EBioMedicine*. 2018;37:410–6.
- Ribeiro FK, Pan W, Bertolde A, Vinhas SA, Peres RL, Riley L, et al. Genotypic and spatial analysis of *Mycobacterium tuberculosis* transmission in a high-incidence urban setting. *Clin Infect Dis*. 2015;61(5):758–66.
- Middelkoop K, Mathema B, Myer L, Shashkina E, Whitelaw A, Kaplan G, et al. Transmission of tuberculosis in a South African community with a high prevalence of HIV infection. *J Infect Dis*. 2015;211(1):53–61.
- Theron G, Jenkins HE, Cobelens F, Abubakar I, Khan AJ, Cohen T, et al. Data for action: collection and use of local data to end tuberculosis. *Lancet*. 2015;386(10010):2324–33.
- Mathema B, Andrews JR, Cohen T, Borgdorff MW, Behr M, Glynn JR, et al. Drivers of tuberculosis transmission. *J Infect Dis*. 2017;216:644–53.
- MacPherson P, Khundi M, Nliwasa M, Choko AT, Phiri VK, Webb EL, et al. Disparities in access to diagnosis and care in Blantyre, Malawi, identified through enhanced tuberculosis surveillance and spatial analysis. *BMC Med*. 2019;17(1):21.
- Rieder HL, Chadha VK, Nagelkerke NJD, van Leth F, van der Werf MJ. Guidelines for conducting tuberculin skin test surveys in high-prevalence countries [second edition]. *Int J Tuberc Lung Dis*. 2011;15(1):S1–S25.
- Anderson RM, May RM. *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press; 1991.
- Becerra MC, Huang C-C, Lecca L, Bayona J, Contreras C, Calderon R, et al. Transmissibility and potential for disease progression of drug resistant *Mycobacterium tuberculosis*: prospective cohort study. *BMJ*. 2019;367:I5894.
- Zelner JL, Murray MB, Becerra MC, Galea J, Lecca L, Calderon R, et al. Identifying hotspots of multidrug-resistant tuberculosis transmission using spatial and molecular genetic data. *J Infect Dis*. 2016;213(2):287–94.
- Lin H, Shin S, Blaya JA, Zhang Z, Cegielski P, Contreras C, et al. Assessing spatiotemporal patterns of multidrug-resistant and drug-sensitive tuberculosis in a South American setting. *Epidemiol Infect*. 2011;139(11):1784–93.
- Manjourides J, Lin H-H, Shin S, Jeffery C, Contreras C, Cruz JS, et al. Identifying multidrug resistant tuberculosis transmission hotspots using routinely collected data. *Tuberculosis*. 2012;92(3):273–9.
- Taylor BM, Davies TM, Rowlingson BS, Diggle PJ. Bayesian Inference and Data Augmentation Schemes for Spatial, Spatiotemporal and Multivariate Log-Gaussian Cox Processes in R. *J Stat Softw*. 2015;63(7):48.
- Heaton MJ, Datta A, Finley AO, Furrer R, Guinness J, Ghananiyogi R, et al. A case study competition among methods for analyzing large spatial data. *J Agric Biol Environ Stat*. 2019;24(3):398–425.
- Banerjee SCB, Gelfand A. *Modeling and analysis for point patterns. Hierarchical modeling and analysis for spatial data*. 2nd ed. Boca Raton: Chapman and Hall/CRC; 2014. p. 199–255.
- Lee D. CARBayes: an R package for Bayesian spatial modeling with conditional autoregressive priors. *J Stat Softw*. 2013;55(13):24.
- Geweke JF. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo JM, Berger JO, Dawid AP, AFM S, editors. *Bayesian statistics*. 4th ed. Oxford: Clarendon Press; 1991.
- Rodriguez-Barraquer I, Salje H, Cummings DA. Opportunities for improved surveillance and control of dengue from age-specific case data. *eLife*. 2019;8:e45474.
- Zelner JL, Murray MB, Becerra MC, Galea J, Lecca L, Calderon R, et al. Age-specific risks of tuberculosis infection from household and community exposures and opportunities for interventions in a high-burden setting. *Am J Epidemiol*. 2014;180(8):853–61.
- Middelkoop K, Bekker L-G, Morrow C, Zwane E, Wood R. Childhood tuberculosis infection and disease: a spatial and temporal transmission analysis in a South African township. *S Afr Med J*. 2009;99(10):738–43.
- Sales CM, Figueiredo TA, Zandonade E, Maciel EL. Spatial analysis on childhood tuberculosis in the state of Espírito Santo, Brazil, 2000 to 2007. *Rev Soc Bras Med Trop*. 2010;43(4):435–9.
- Venâncio T, Tuan T, Nascimento L. Incidence of tuberculosis in children in the state of São Paulo, Brazil, under spatial approach. *Ciênc Saúde Coletiva*. 2015;20(5):1541–7.
- Alene KA, Viney K, McBryde ES, Clements ACA. Spatiotemporal transmission and socio-climatic factors related to paediatric tuberculosis in north-western Ethiopia. *Geospat Health*. 2017;12(2):575.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.