PERSPECTIVE

Subnational burden estimates to find missing people with tuberculosis: wrong but useful?

Sandra Alba[1*], Christina Mergenthaler[1], Mirjam I. Bakker[1] and Ente Rood[1]

Abstract

Efforts to combat tuberculosis (TB) require reliable national and subnational data for planning, monitoring and evaluation. Yet, reliable subnational estimates of TB burden are hard to come by—especially at the lower levels of disaggregation such as district, community, or ward level. Several approaches have been proposed to generate subnational estimates of TB burden. However, ascertaining the accuracy of modelled estimates and ensuring their use for TB program planning remains a challenge, thereby raising questions about their usefulness. In this perspective article, we review several subnational TB models to gain insights into their accuracy, purpose and use as a starting point to reflect on their usefulness in finding the missing people with TB. We argue that despite concerns about their accuracy, subnational TB models can help pinpoint areas that deserve more programmatic attention (spatial targeting) and better understand the effectiveness of interventions (programmatic learning). Furthermore, increasing the use of these models can help improve both their accuracy and usefulness in the long run—if estimates are systematically compared against programmatic data and models are improved to better capture reality on the ground. As such, we conclude that subnational TB models represent an essential evidence-based learning tool to guide the search for the missing people with TB.

Keywords Modelling, Subnational TB, Accuracy, Use

Background

Essentially all models are wrong, but some are useful— this famous aphorism attributed to the British statistician George Box [1] acknowledges that models cannot capture the complexities of reality but could nevertheless be useful if users take model limitations into account. As modellers and technical advisors to tuberculosis (TB) control programs in several high-burden countries, this aphorism also resonates for models to estimate subnational TB burden. In this perspective article, we review several subnational TB models to gain insights into their accuracy, purpose and use as a starting point to reflect on their usefulness in finding the missing people with TB.

Efforts to combat TB require reliable national and subnational data for planning, monitoring and evaluation. TB remains a major global public health challenge, with millions of people affected each year [2]. As an infectious disease, it tends to cluster geographically as a result of person-to-person transmission. In addition, risk factors for TB are also clustered since social and environmental factors are more prevalent in certain geographical areas (poverty, overcrowded living conditions, poor access to healthcare, etc.). Broadly speaking, subnational TB data can inform TB control strategies by providing insights to understand the local epidemiology of the disease, develop locally specific interventions, and prioritise resource allocation. More specifically, by pinpointing geographic areas

*Correspondence:
Sandra Alba
s.alba@kit.nl
[1] KIT Royal Tropical Institute, Amsterdam, Netherlands

with higher TB burden, subnational TB models can play a crucial role in the search for the millions of 'missing people with TB' (individuals affected with TB that are either undiagnosed or diagnosed but not reported in official TB statistics [3]. Yet, reliable subnational estimates of TB burden are hard to come by – especially at the lower levels of disaggregation such as district, community, or ward level. Currently, TB notifications reported through health facilities or district-level TB registers are the main source of subnational TB data, but subnational TB case notification rates are more often reflective of TB programme efforts and access to healthcare than TB burden [4, 5] especially in settings where large fractions of cases are unaccounted for by routine systems. As a result, high notifications indicate an unknown mix of high burden and high performance and vice versa for low relative notifications. National population-based prevalence surveys provide a direct measurement of the burden of disease, but they are typically not powered to provide subnational granularity.

Given the lack of empirical data, several approaches have been proposed to generate subnational estimates of TB burden. A literature scan suggests that scientific interest in subnational TB modelling started to grow from 2010 onwards. Yet efforts to understand spatial heterogeneity in disease occurrence and underlying risk factors were already ongoing in prior years with the use of geographic information systems and spatial statistics to analyse small-scale TB data [6]. Conceptually two types of geo-spatial approaches to estimating TB burden can be distinguished: (1) classify or rank subnational regions by TB burden and; (2) provide point estimates of TB burden for subnational regions. The classification or ranking of approaches does not necessarily require statistical or mathematical modelling and can be done with multi-criteria decision-making algorithms (such as for example the MATCH approach [7]). Conversely, providing point estimates of TB burden for subnational regions can be

computationally intensive and requires methodological rigour and ample data. Over the last decade interest in this type subnational TB modelling increased, coinciding with a growing application of statistical and mathematical models to evaluate health programmes [8] advancements in computational methods and TB data availability [9] and a recognition that new approaches were necessary to enable tuberculosis elimination [10–12]. All three TB burden indicators are of interest for subnational modelling: incidence, mortality and prevalence. While modelling TB mortality and incidence is relevant for global reporting [13, 14], prevalence is appealing to modellers as TB surveys provide ample empirical data for modelling efforts [15, 16].

Building on previous work by Garnett et al. [8] we propose that models to estimate subnational TB burden can be categorised as either mathematical or statistical (Table 1), each corresponding to distinct types of reasoning, varying degrees to which empirical data is needed and different ways of factoring in the role of chance. Here we consider Bayesian approaches as statistical and inductive although Bayesian probability theory can also be considered a mathematical framework and Bayesian approaches include some elements of deductive reasoning when they build on prior knowledge. We also classified machine learning models as statistical because they often involve estimating parameters (in fact many machine learning algorithms are based on statistical regression models). But one could also argue that they constitute a separate computational modelling paradigm that goes beyond traditional statistics by leveraging techniques to handle large datasets and optimise complex models.

Despite the wealth of approaches available, ascertaining the accuracy of modelled estimates and ensuring their use for TB program planning remains a challenge, thereby raising questions about their usefulness. From a technical perspective, the accuracy of model predictions

Table 1 Overview of modelling approaches for subnational TB

Modelling paradigm	Type of reasoning	Use of empirical data	Role of chance	Examples of models
Mathematical: Produces definite TB outcomes based on pre-specified inputs and relationships with other TB or non-TB data	Deductive: relies on known relationships and specific inputs to yield TB outcomes	No empirical data is necessary about the TB indicator to be modelled as long as relationships with other TB or non-TB data can be formulated	Variation can be introduced in the model simulations to account for inherently random effects and reflect variability in model inputs	Deterministic and stochastic compartmental models Individual-based deterministic and stochastic models
Statistical: Uses existing empirical TB data to estimate probabilities and make predictions for unobserved cases	Inductive: Draws general conclusions about a TB outcome from specific observations and associations	Empirical data for the TB indicator to be modelled is necessary for certain areas for model building (the same areas for which risk factors are also available)	Variation inherent to model building as associations between the TB outcome and risk factors are estimated based on probability theory	Frequentist regression models Bayesian approaches Machine learning

is very difficult to ascertain in the absence of an empirical ground truth to use for validation. The 'Benchmarking, reporting, and review' approach proposes several benchmarks to circumvent empirical evidence when assessing the quality of a TB model [18], but its applicability to subnational models is limited when the benchmarking data is not available subnationally. In addition, models are often technically complex which can limit their use by TB program planners. Indeed, planners may not have the expertise to understand subnational models, negatively affecting their trust in modelled estimates and willingness to engage in discussions about their use. Recommendations to increase the quality of model outputs for country-level TB policy-making recognise the need to involve local stakeholders in the modelling process [18, 19], but in our experience this is not always enough to ensure models become part and parcel of in-country decision-making processes.

Ensuring the accuracy and use of models for TB subnational estimation is key to fulfil their potential to find and treat missing people with TB and will determine the model's impact on TB control. While accurate model predictions are needed to ensure programmatic decisions are based on valid estimates, in practice a model needs to be trusted and used to have an impact. In this paper, we provide our perspective on current efforts by modellers to ensure the accuracy and use of their modelling approach. We present some selected examples of statistical and mathematical models that generated TB subnational estimates of TB prevalence, incidence and mortality. Thereafter we reflect on (1) efforts made by modellers to validate the accuracy of estimates; (2) the models' purpose and reported use for TB planning. Finally, we build on this information to reflect on the usefulness of subnational estimates to find the missing people with TB.

Existing models: accuracy, purpose and use

In our experience, statistical models of TB burden are very common for TB prevalence and there are also some—albeit fewer—examples of models for TB incidence and mortality. One commonly used statistical approach to predicting subnational TB prevalence leverages TB prevalence survey data with Bayesian regression modelling as documented for instance in Bangladesh [23] and Ethiopia [24] or the TB Hackathon in Pakistan [25] just to name a few. Examples of Bayesian models estimating incidence include a geospatial Bayesian model to link case notifications to unobserved TB incidence in Ethiopia, allowing for differences in case detection identified through the presence of health facilities at a local level [20]. In another application, modellers estimated municipal-level incidence and fraction of individuals treated

based on available notification and mortality data in Brazil [21] (although the model is referred to as mathematical by the authors we refer to it as statistical according to the definitions in Table 1). To the best of our knowledge, there is only one documented application to date of a statistical model used to estimate subnational TB mortality: a Bayesian regression model that used vital registration data from the national mortality information system and TB case notifications statistically to predict TB mortality at municipal level in Brazil [22].

Conversely, we find mathematical models are mostly used for TB incidence and mortality but we are not aware of any for TB prevalence. Compartmental models, a commonly used approach for the mathematical modelling of infectious diseases, have been used to estimate TB incidence at the health zone level in the South-Kivu Province of the Democratic Republic of Congo based on population density and TB notification data [27]. In many ways, the SubSET multiplicative model used to estimate district-level incidence in Indonesia [28] can also be considered as a mathematical model according to our definitions in Table 1. This method uses WHO-estimated TB incidence for the country and known ecological predictors of TB (e.g. population size, urbanization, socioeconomic indicators) to deduce district-level values of incidence. The mathematical TIME model stands out as the only model producing two provincial-level outcome indicators (mortality and incidence) in South Africa [30]. The TIME model is an age-structured, dynamic, compartmental transmission model of TB with a user-friendly interface that automatically incorporates country data on TB notifications and demographic projections and can be applied at subnational level.

Reflecting on these selected statistical and mathematical models, an association between types of modelling approaches and TB outcome variables can be discerned. The widespread availability of empirical data at the subnational level for model building (e.g. cluster locations of the TB prevalence surveys) likely explains the ample use of statistical models to estimate TB prevalence. Conversely, mathematical models are appealing to model outcomes—such as incidence—that cannot be measured directly [29] and for which therefore require building on assumptions about TB transmission dynamics and disease progression. This explanation resonates with the fact that statistical models for TB incidence tend to include complex mathematical components, e.g. to allow for local differences in case detection rates and to estimate TB incidence from notifications [19, 20]. Mortality seems to be an indicator for which both mathematical and statistical models can be used. While mortality can be measured directly with sensitive reporting systems, death counts are low and therefore unstable in small areas, and

Alba *et al. BMC Global and Public Health*   (2024) 2:77

Page 4 of 6

statistical modelling approaches can separate true differences in risk from stochastic noise [21].

It is interesting to note that most of the models described so far were developed with some intention to inform TB programs, but the outputs' accuracy was not always validated and their use for TB programming was not often reported. The original purpose of the models encompassed supporting subnational resource allocation, informing TB policies, guiding active case-finding activities and evaluating interventions. A further review of the models described so far reveals that only two, explicitly refer to their use by TB program planners either to steer case-finding activities [24] or for resource allocation and planning [28]. Furthermore, two TB Hackathon models were used as a basis for sample size calculations for the national TB prevalence survey [24]. However, we cannot discard the possibility that some models were used for TB program planning even if that was not reported in the publications we consulted. Several models reported some validation efforts to gauge the accuracy of their estimates, either using data not used for model building [27] or sample splitting methods [21, 24].

Usefulness of subnational estimates to find missing people with TB

From a programmatic perspective, the use of subnational TB estimates can serve two monitoring and evaluation (M&E) functions: learning and accountability. The models we presented here rather emphasized the learning function, as their primary purpose included supporting subnational resource allocation, informing TB policies, guiding active case-finding activities and evaluating the impact of future interventions. None of the models we reviewed made explicit reference to using the modelled estimates for an accountability function, i.e. to set *performance targets* for the absolute number of people with TB to be found through case-finding activities (as done for example in pay-for-performance schemes that use health outcome targets to incentivise health service delivery [31]. This is reassuring, given that most models did not report efforts to validate the accuracy of their estimates using independent data not used to develop the model (sample splitting methods seem to be more frequently used but provide a less reliable assessment of a models' performance). The TB hackathon offered a unique opportunity to gauge model validity and showed that even with similar predictors and comparable modelling approaches, estimates can vary substantially, thereby casting doubt on all models' accuracy. This is in line with other TB model comparison exercises that have shown that differences in results occur even when models evaluate the same policy alternatives in the same setting [32].

However, target setting aside, we believe subnational estimates can be useful in finding missing people by increasing the effectiveness of case-finding interventions. Ideally and in the long term the effectiveness of case-finding interventions would be measured in terms of their impact on TB incidence. However, given challenges in measuring incidence in settings with high TB burden (where it is assumed that many people remain undetected and untreated) [29] increases in notifications are often the most pragmatic metric. We believe subnational estimates can help increase case notifications and thereby identify effective interventions in two ways.

First of all, subnational TB estimates can be used for *spatial targeting* of interventions [33]. Indeed, while models may not have the level of accuracy needed to provide point estimates and ranges for each area, they may still enable the ranking of areas or discrimination between high and low case detection efficiency [27] deduced for example by high prevalence-to-notification-ratios [24, 30] or low case detection rates [19]. Ideally, reasons for under-notifications in the target areas should also be investigated to select the most effective intervention for the given target areas. Indeed, a TB program may be missing people with TB for several reasons along the care cascade [34]: there may be foci with high transmission in areas with particularly high numbers of key populations (meaning many individuals at risk of or with TB), areas with poor access to care or limited knowledge about TB (meaning few individuals accessed TB screening or testing), or issues with testing and screening or diagnostic capacity (meaning few individuals were tested and diagnosed). Selecting the most effective interventions to address the locally specific TB detection gap will maximise the chances of actually finding the missing people with TB in the target area (e.g. intensified screening in key populations, mobile chest camps and community sensitisation in areas with poor access to care, or setting up sputum transportation networks to improve testing). However, is important to realise that these types of strategies based on aggregated burden estimates may overlook foci involving fewer people, thus steering efforts away from structurally left-out populations.

In addition, subnational estimates can be useful as a starting point for *programmatic learning*, i.e. to provide insights for TB programmers on the real-life effectiveness of interventions. This is closely related to the application of the TIME model in South Africa [30] which provided a time series of estimates to estimate the impact of various TB interventions. The ability to make future projections and evaluate different hypothetical scenarios is one of the main strengths of mathematical models [7]. However, more real-time operational applications of programmatic learning can be achieved with all types

Alba *et al. BMC Global and Public Health*　　　(2024) 2:77

Page 5 of 6

of models when the roll-out of interventions is closely monitored with relevant program metrics (e.g. screening yield, number needed to screen, number needed to test, notification rate) to assess the match between target areas and selected interventions. Increases or decreases in program metrics could be indicative of a good match between a targeted area and a selected intervention. Conversely, no changes could mean either that the target area was wrongly classified as high burden by the model; or that the selected intervention was not appropriate for the area. Various iterations of the case-finding approach and field investigations may be necessary to understand whether the TB interventions need to be adapted whether the model needs to be improved, or both. The willingness of local program planners to use modelled estimates is a key pre-condition for this type of programmatic learning and often depends on the extent to which planners have been engaged in the generation of the estimates—as per various recommendations to increase the quality of modelled estimates [17, 18].

Conclusions

Subnational TB models may not be entirely correct but they can still be useful to find missing people with TB. Increasing their use may help increase their accuracy (and therefore usefulness) in the long term. A panoply of statistical and mathematical approaches are available to model subnational TB incidence, prevalence and mortality. These can help pinpoint areas that deserve more programmatic attention (spatial targeting) and better understand the effectiveness of interventions (programmatic learning). While ascertaining the validity of these models remains a challenge, it is important to consider that in the absence of a model, prioritisation would likely be based on experts' opinions (that may also be wrong); or there may be no prioritisation at all which is also judgement about subnational variation, namely that there is none (also unlikely to be true). On the other hand, increasing the use of subnational models can help improve both their accuracy and usefulness in the long run—if estimates are systematically compared against subnational programmatic and survey data and if models are increasingly improved to better capture reality on the ground. In a world of scarce and limited public health resources yet ample data and computational potential, subnational TB models represent an essential evidence-based tool to guide the search for the missing people with TB.

Authors' contributions
SA conceived the work and wrote the first draft of the manuscript based on several consultations with CM, MB and ER. CM conducted the literature search and reviewed all models together with ER. CM, MB and ER critically revised various iterations of the manuscript and provided important intellectual content. All authors provided substantial contributions and agree to be accountable for all aspects of the work.

Data availability
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Box GEP. Science and Statistics. J Am Stat Assoc. 1976;71(356):791–9. https://doi.org/10.1080/01621459.1976.10480949.
2. Global Tuberculosis Report 2022. Available from: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022. Cited 2023 Aug 22.
3. Chin DP, Hanson CL. Finding the missing tuberculosis patients. J Infect Dis. 2017;216(Suppl 7):S675–8.
4. van Gurp M, Rood E, Fatima R, Joshi P, Verma SC, Khan AH, 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 Infect Dis. 2020;20(1):490.
5. Rood E, Khan AH, Modak PK, Mergenthaler C, Van Gurp M, Blok L, et al. A spatial analysis framework to monitor and accelerate progress towards SDG 3 to end TB in Bangladesh. ISPRS Int J Geo-Inf. 2019;8(1):14.
6. Shaweno D, Karmakar M, Alene KA, Ragonnet R, Clements AC, Trauer JM, et al. Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. BMC Med. 2018;16(1):193.
7. World Health Organization. Compendium of data and evidence-related tools for use in TB planning and programming. Geneva; 2021.
8. Garnett GP, Cousens S, Hallett TB, Steketee R, Walker N. Mathematical models in the evaluation of health programmes. The Lancet. 2011;378(9790):515–25.
9. .World Health Organization. Electronic recording and reporting for tuberculosis care and control. 2012;(WHO/HTM/TB/2011.22). Available from: https://apps.who.int/iris/handle/10665/44840. Cited 2023 Sep 13.
10. Executive Board 134. Global strategy and targets for tuberculosis prevention, care and control after 2015: Report by the Secretariat. 2014. Report No.: EB134/12. Available from: https://apps.who.int/iris/handle/10665/172828. Cited 2023 Sep 1.
11. Lienhardt C, Espinal M, Pai M, Maher D, Raviglione MC. What Research Is Needed to Stop TB? Introducing the TB Research Movement. PLoS Med. 2011;8(11):e1001135.
12. Rylance J, Pai M, Lienhardt C, Garner P. Priorities for tuberculosis research: a systematic review. Lancet Infect Dis. 2010;10(12):889–92.

Alba *et al. BMC Global and Public Health*   (2024) 2:77

Page 6 of 6

13. Millennium Development Goals (MDGs). Available from: https://www. who.int/news-room/fact-sheets/detail/millennium-development-goals-(mdgs) . Cited 2023 Sep 13.

14. SDG Target 3.3 Communicable diseases. Available from: https://www. who.int/data/gho/data/themes/topics/sdg-target-3_3-communicable-diseases. Cited 2023 Sep 13.

15 Law I, Floyd K, African TB Prevalence Survey Group. National tuberculosis prevalence surveys in Africa, 2008–2016: an overview of results and lessons learned. Trop Med Int Health TM IH. 2020;25(11):1308–27.

16. Onozaki I, Law I, Sismanidis C, Zignol M, Glaziou P, Floyd K. National tuberculosis prevalence surveys in Asia, 1990–2012: an overview of results and lessons learned. Trop Med Int Health TM IH. 2015;20(9):1128–45.

17. Alba S, Rood E, Bakker MI, Straetemans M, Glaziou P, Sismanidis C. Development and validation of a predictive ecological model for TB prevalence. Int J Epidemiol. 2018;47(5):1645–57.

18. McQuaid CF, Clarkson MC, Bellerose M, Floyd K, White RG, Menzies NA. An approach for improving the quality of country-level TB modelling. Int J Tuberc Lung Dis. 2021;25(8):614–9.

19. Menzies NA, McQuaid CF, Gomez GB, Siroka A, Glaziou P, Floyd K, et al. Improving the quality of modelling evidence used for tuberculosis policy evaluation. Int J Tuberc Lung Dis. 2019;23(4):387–95.

20. Shaweno D, Trauer JM, Denholm JT, McBryde ES. A novel Bayesian geospatial method forestimating tuberculosis incidence reveals many missed TB cases in Ethiopia. BMC Infect Dis. 2017;17(1):662.

21. Chitwood MH, Alves LC, Bartholomay P, Couto RM, Sanchez M, Castro MC, et al. A spatial-mechanistic model to estimate subnational tuberculosis burden with routinely collected data: an application in Brazilian municipalities. PLOS Glob Public Health. 2022;2(9):e0000725.

22. Ross JM, Henry NJ, Dwyer-Lindgren LA, de Paula Lobo A, de Marinho-Souza F, Biehl MH, et al. Progress toward eliminating TB and HIV deaths in Brazil, 2001–2015: a spatial assessment. BMC Med. 2018;16:144.

23. Allorant A, Biswas S, Ahmed S, Wiens KE, LeGrand KE, Janko MM, et al. Finding gaps in routine TB surveillance activities in Bangladesh. Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis. 2022;26(4):356–62.

24. Alene KA, Python A, Weiss DJ, Elagali A, Wagaw ZA, Kumsa A, et al. Mapping tuberculosis prevalence in Ethiopia using geospatial meta-analysis. Int J Epidemiol. 2023;52(4):1124–36.

25. Alba S, Rood E, Mecatti F, Ross JM, Dodd PJ, Chang S, et al. TB Hackathon: development and comparison of five models to predict subnational tuberculosis prevalence in Pakistan. Trop Med Infect Dis. 2022;7(1):13.

26. Alba S. TB Hackathon: development and comparison of five models to predict subnational tuberculosis prevalence in Pakistan. 2021; Available from: https://zenodo.org/record/5112022. Cited 2023 Sep 1.

27. Faccin M, Rusumba O, Ushindi A, Riziki M, Habiragi T, Boutachkourt F, et al. Data-driven identification of communities with high levels of tuberculosis infection in the Democratic Republic of Congo. Sci Rep. 2022;12(1):3912.

28. Parwati CG, Farid MN, Nasution HS, Basri C, Lolong D, Gebhard A, et al. Estimation of subnational tuberculosis burden: generation and application of a new tool in Indonesia. Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis. 2020;24(2):250–7.

29. Dye C, Bassili A, Bierrenbach AL, Broekmans JF, Chadha VK, Glaziou P, et al. Measuring tuberculosis burden, trends, and the impact of control programmes. Lancet Infect Dis. 2008;8(4):233–43.

30. Hippner P, Sumner T, Houben RM, Cardenas V, Vassall A, Bozzani F, et al. Application of provincial data in mathematical modelling to inform sub-national tuberculosis program decision-making in South Africa. PLoS One. 2019;14(1):e0209320.

31. Kovacs RJ, Powell-Jackson T, Kristensen SR, Singh N, Borghi J. How are pay-for-performance schemes in healthcare designed in low- and middle-income countries? Typology and systematic literature review. BMC Health Serv Res. 2020;20(1):291.

32. Houben RMGJ, 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):e806–15.

33. Khundi M, Carpenter JR, Nliwasa M, Cohen T, Corbett EL, MacPherson P. Effectiveness of spatially targeted interventions for control of HIV, tuberculosis, leprosy and malaria: a systematic review. BMJ Open. 2021;11(7):e044715.

34. Subbaraman R, Nathavitharana RR, Mayer KH, Satyanarayana S, Chadha VK, Arinaminpathy N, et al. Constructing care cascades for active tuberculosis: a strategy for program monitoring and identifying gaps in quality of care. PLoS Med. 2019;16(2):e1002754.

Publisher's Note