Capturing heterogeneous infectiousness in transmission dynamic models of tuberculosis: a compartmental modelling approach

Yayehirad A Melsew1,2*, Romain Ragonnet1, Allen C Cheng1, Emma S McBryde3, James M Trauer1

1Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, 553 St Kilda Road, Melbourne, VIC 3004, Australia
2Department of Epidemiology and Biostatistics, Institute of Public Health, University of Gondar, Gondar, Ethiopia
3Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, QLD 4811, AUSTRALIA

*Corresponding Author (email: 078yayu@gmail.com or yayehirad.melsew@monash.edu, Tel: +61403676870)

Authors’ email addresses

YAM: 078yayu@gmail.com
RR: romain.ragonnet@monash.edu
ACC: allen.cheng@monash.edu
ESM: emma.mcbride@jcu.edu.au
JMT: james.trauer@monash.edu
Abstract

Infectiousness heterogeneity among individuals with tuberculosis (TB) is substantial and is likely to have a significant impact on the long-term dynamics of TB and the effectiveness of interventions. However, there is a gap in capturing heterogeneous infectiousness and evaluating its impact on the effectiveness of interventions.

Informed by observed distribution of secondary infections, we constructed a deterministic model of TB transmission using ordinary differential equations. The model incorporated assumption of heterogeneous infectiousness with three levels of infectivity, namely non-spreaders, low-spreaders and super-spreaders. We evaluated the effectiveness of dynamic transmission untargeted and targeted implementation of an intervention intended to represent active case finding with a point-of-care diagnostic tool. The simulated intervention detected 20% of all TB patients who would otherwise have been missed by the health system during their disease episode and was compared across four epidemiological scenarios.

Our model suggested that targeting the active case finding intervention towards super-spreaders was more effective than untargeted intervention in all setting scenarios, with more effectiveness in settings with low case detection and high transmission intensity. For instance, a targeted intervention achieved a 42.2% reduction in TB incidence, while the untargeted intervention achieved only a 20.7% reduction over 20 years, given the same number of people treated. Although the most marked impact on equilibrium TB incidence came from the rate of late reactivation, the proportion of super-spreaders and their relative infectiousness had shown substantial impact.

Targeting active case-finding interventions to highly infectious cases likely to be particularly beneficial in settings where case detection is poor. Heterogeneity-related parameters had an
equivalent effect to several other parameters that have been established as being very important to TB transmission dynamics.

Keywords: heterogeneous infectiousness; M. tuberculosis transmission; mathematical model;

Introduction

Tuberculosis (TB) is the world’s leading cause of death from a single infectious agent, in 2019, ranking above HIV/AIDS with an estimated 10.0 million cases and 1.2 million deaths worldwide [1]. The World Health Organization developed the new End TB Strategy for post-2015 TB elimination activities, with the ambition of a 95% reduction in TB deaths and a 90% reduction in TB incidence by 2035 by comparison to 2015 rates [2]. However, the natural history of TB remains poorly understood, and consequently, the uncertain potential impact of control interventions limits confidence about the possibility of its elimination. The heterogeneous transmission of Mycobacterium tuberculosis (Mtb) within populations is well-established, but its epidemiological impact is poorly understood. Moreover, up to a third of TB cases are not diagnosed, so finding and treating infectious cases is key to achieving TB control, and finding and treating highly infectious people is the key to TB control in high transmission settings [3, 4].

We define infectiousness heterogeneity as the variability in the capacity of infectious patients to produce secondary infections, which may be attributable to characteristics of the host, the agent and the environment [5-9]. Heterogeneity of infectious individuals in spreading Mtb infection is well-recognised, with a small group of highly infectious individuals producing a large proportion of secondary infections, while many others produce very few or none [10]. Previous TB genomic epidemiology has identified TB super-spreading events by quantifying heterogeneity in the infectiousness of TB patients through fitting a standard statistical distribution (the negative binomial distribution, NBD) to the distribution of secondary cases.
produced by each infectious patient (the “offspring distribution”) [11]. More recently, data from TB contacts in a low-transmission setting have allowed estimation of the proportion of all TB patients that can be categorised as super-spreaders, finding it to be approximately 10% [9]. In such studies, super-spreaders were typically defined as patients who produced a number of secondary “infections/cases” greater than the 99th centile of a standard Poisson offspring distribution, with distribution mean equal to the average number of secondary infections per index [9, 12]. This heterogeneity is likely to have implications for both the burden of disease and the effectiveness of control interventions. Thus, we propose that it is necessary to capture this heterogeneity when modelling TB transmission dynamics.

Some past compartmental models of TB transmission dynamics have attempted to capture heterogeneity in patients’ infectiousness by stratifying the active TB compartment into different levels of infectivity. Amongst the most typical approaches is stratification as either infectious (usually representing pulmonary TB) or non-infectious (extrapulmonary TB) [13-17]. This approach implies that all pulmonary TB patients are equally infectious, while all extrapulmonary patients are entirely non-infectious, and so does not capture the heterogeneity among pulmonary patients. Other TB models have considered sputum smear status as a factor in stratifying patients’ levels of infectiousness, considering both smear-negative and smear-positive pulmonary TB to be infectious, with the relative infectiousness of smear-negative patients compared to smear-positive typically set between 15 and 25% [18-33]. While stratifying pulmonary TB patients based on smear status captures an additional clinical attribute that is important in determining infectiousness, even these models still do not capture the full picture of TB patients’ infectiousness variation, since several behavioural and demographic factors other than smear status can affect the level of infectiousness [5].

An individual-based model simulated patterns of meetings and *Mtb* transmission between three different types of contact to determine the effects of variation in infectiousness and
susceptibility on transmission location. The study defined super-spreading defined based on the number of secondary cases instead of secondary infections and suggested that the majority of disease resulted from infection by a small proportion of people with TB or super-spreaders [34]. However, as TB disease activation may take long time and also depend on mainly on the characteristics of contact person, the definition of super-spreading used in the study may not show true heterogeneity in Mtb transmission [35]. In the current study, we present a deterministic compartmental Mtb transmission model that incorporates three levels of TB patients’ infectivity, namely non-spreaders, low-spreaders and super-spreaders. Using this framework, we analysed how infectiousness heterogeneity affects disease dynamics and evaluated the effectiveness of a hypothetical active case finding intervention and the impact of targeting super-spreaders.

Methods

Literature review

Before constructing our model, we systematically reviewed how previous TB models have approached heterogeneous infectiousness [36]. In the review, we found that TB models frequently stratified the active TB compartment according to one or more patient-related factors. We constructed a flexible model with three levels of infectiousness, without restricting to a given factor, and the ability to transition between these levels using empirical measures of heterogeneous infectiousness to parametrise the model.

Empiric data

Data from the VTP which are stored by the Victorian Department of Health and Human Services (DHHS) were used for the following analysis. Index patients were classified as confirmed cases of TB notified from 1 January 2005 to 31 December 2015 in residents of Victoria. The data set includes contact tracing information and results of testing for Mtb
infection, with cases of subsequent active TB disease linked to these contact episodes now extending to March 2017 (see [37] for earlier publication of linkage process). We constructed empirical offspring distributions from the detailed contact tracing data set of the VTP.

Ethical approval was obtained from Monash University, Human Research Ethics Committee (Project Number: 7776) and permission was given by the VTP and DHHS.

Model structure

Using ordinary differential equations (ODE), we constructed a deterministic model of \textit{Mtb} transmission in a hypothetical high TB burden setting. To represent heterogeneous infectiousness and super-spreading in \textit{Mtb} transmission, we stratified the active TB compartment into three sub-classes with different levels of infectiousness: non-spreaders (\(I_0\)), low-spreaders (\(I_1\)) and super-spreaders (\(I_2\)). We did not use any specific factor to discriminate \(I_0\), \(I_1\) and \(I_2\) a priori. Instead, we use the offspring data (number of secondary infections per index) and a Bayesian approach to find the most realistic parameterisation to implement heterogeneous infectiousness in our model.

Fig 1 presents the model structure. The model simulates a closed population, such that births \((\pi)\) accrue into the fully susceptible compartment \((S)\) to replace all deaths. Natural death occurs from all compartments at a constant rate \((\mu)\). Susceptible individuals are dynamically infected, with the force of infection \((\lambda)\) contributed by super-spreaders and low-spreaders.

Following infection, individuals enter a rapid-sojourn early latent compartment \((L_\lambda)\), from which they may progress to one of the three active TB compartments \((I_0, I_1\) or \(I_2)\) at a total rate of \(\epsilon\) or enter the low-risk late latent stage \((L_\kappa)\) at rate \(\kappa\). From the late latent state they may progress more slowly to disease (total rate \(\nu\)), also entering the active TB compartments \((\lambda)\). Re-infection may occur during late latency with a reduced force of infection \(\lambda_r\) \((\lambda_r = r \times \lambda)\) and, if re-infection occurs, these individuals similarly enter the early latent compartment,
progressing to the disease at the same rates as for newly infected persons. Individuals with active disease may either die due to TB (rate $\mu_i$) or background mortality ($\mu$); or spontaneously recover and return to the late latent state (rate $\gamma$); or be detected and treated by the health system (rate $\delta$) and return to the susceptible compartment ($S$).

Fig 1: Transmission dynamic model of TB with heterogeneous infectiousness. Boxes represent compartments in the model and arrows represent flows. Not shown here but included in the model are natural mortality from each compartment and TB-related mortality from each active TB compartment ($I_0$, $I_1$ and $I_2$). Red arrows represent transmission flows; black arrows represent disease progression processes; green arrows represent self-recovery, blue arrows represent detection and treatment cure.

Estimation of the infectiousness parameters using empiric evidence

The four model parameters that capture heterogeneous infectiousness in the model were the proportions of non-spreaders, low-spreaders and super-spreaders, and the relative infectiousness of super-spreaders compared to low-spreaders. In order to estimate these parameters, we first determined the statistical distribution of the number of secondary infections per index associated with the stochastic equivalent (continuous-time Markov) model of our ODE-based model. We obtained a mixture of three classic statistical distributions: two geometric distributions for low-spreaders and super-spreaders and one Dirac delta distribution for non-spreaders. A detailed demonstration of this result is provided in the Supplementary Material.

A Hamiltonian Monte Carlo (HMC) algorithm was used to generate 20,000 samples from the posterior distributions of the parameters. The 20,000 samples used for inference were obtained from 40,000 iterations of the HMC, burning the first 20,000 draws. Burn-in size and convergence were assessed by inspecting the trace plots of the estimated parameters. The
reported 95% credible intervals were obtained by computing the 2.5th and 97.5th percentiles of the parameters’ posterior distributions. The parameter estimates from the Bayesian process is provided in Supplementary fig.1.

The resulting statistical distribution was then compared in a Bayesian context with empiric data obtained from our previous analysis of TB contact tracing data in the Australian state of Victoria [9]. A comparison between the distribution of the number of secondary infections per index simulated by our calibrated model with the one observed empirically is presented in Fig 2.

The parameter values other than those pertaining to infectiousness introduced above were estimated from a review of relevant evidence, with the relevant sources presented in Table 1.

Fig 2. Comparison of the effective modelled secondary infection distribution with empiric data from Victoria, Australia (red points represent the modelled distribution while blue points are observed empiric data).

Model calibration

For the baseline equilibrium state, the ODE model was calibrated to a high TB burden settings with an incidence rate of 200 per 100,000 population per year by adjusting the transmission rate parameter for low-spreaders.

The model was programmed in R 3.4.0 (R Project for Statistical Computing), and we used a package called ‘deSolve’, solver for initial value problems of ODE [38, 39]. The Bayesian parameter estimation was done using the package ‘rstan’ (v.2.18.2) [40] and the source code for the model is publicly available from https://github.com/Yayehirad/TB_heterogeneity
Table 1: Model parameter values

| Corresponding model parameter | Symbol | Base model value | Uncertainty range | Source |
|-------------------------------|--------|------------------|-------------------|--------|
| Natural mortality rate, per year | $\mu$ | 0.0154 | 0.0133 - 0.0182 | [41] |
| Rate of early progression, per year | $\varepsilon$ | 0.401 | 0.307 - 0.548 | [42] |
| Rate of progression to late latency after infection, per year | $\kappa$ | 3.599 | 3.452 - 3.693 | [42] |
| Rate of late progression, per year | $\nu$ | 0.002 | 0.0009 - 0.004 | [42] |
| Average duration of infectiousness, years | $D$ | 3 | 2-4 | [43] |
| TB-related mortality among Super-spreaders, proportion | $\mu_{I2}$ | 0.534 | 0.327 - 0.78 | [43, 44] |
| TB-case fatality among low-spreaders, proportion | $\mu_{I1}$ | 0.077 | 0.077 - 0.096 | |
| TB-case fatality among non-spreaders, proportion | $\mu_{I0}$ | 0.077 | 0.077 - 0.096 | |
| Conversion from non-infectious to low-spreaders, proportion | $h$ | 0.014 | 0 - 0.06 | [32, 45] |
| Conversion from low-spreaders to super-spreaders, proportion | $j$ | 0.014 | 0 - 0.06 | Assumed |
| Relative reduction in force of re-infection compared to new infection | $r$ | 0.21 | 0.15 - 0.4 | [46] |

**Parameters calibrated from data**

| Parameter | Value | Source |
|-----------|-------|--------|
| Proportion of infectious among all TB | 0.42 | Estimated from data [9] |
| Proportion of in super-spreaders among infectious | 0.18 | Estimated from data [9] |
| Transmission rate per low-spreader per-year | $\beta_{I}$ | 12 | 10 - 15 | Calibrated |
| Relative infectiousness of super-spreaders compared to low-spreaders | $m$ | 4.67 | 3.69 - 5.88 | Estimated from data [9] |

**Intervention parameters**

| Parameter | Value | Source |
|-----------|-------|--------|
| Baseline case detection rate, proportion | CDR | 0.65 | 0.5 - 0.75 | Assumed |
| Intervention (proportion of missed cases detected) | $q$ | 0.2 | 0 - 0.4 | Assumed |
| Proportion of intervention targeted to super-spreaders | $d$ | 0.5 | 0 - 1.0 | Assumed |

Sensitivity analysis

We first performed a one-way sensitivity analysis on each parameter to understand the sensitivity of equilibrium incidence to parameter ranges in the absence of interventions. We then undertook multidimensional sensitivity analyses to assess the correlation between equilibrium incidence and each parameter by taking 1000 samples from each parameter ranges with Latin hypercube sampling (LHS) technique [47, 48].
We simulated an intervention intended to represent active case finding with a point-of-care diagnostic tool, accelerating successful diagnosis and treatment of individuals as early as possible before they start transmitting the infection. This could be conceptually considered as implemented by a highly sensitive point-of-care diagnostic technology following community awareness campaigns that target individuals with noticeable symptoms in order to increase their rate of presentation [49]. To target the intervention, clinical symptoms and behavioural characteristics can be used to identify those individuals who are more likely to be super-spreaders.

We implemented an intervention capable of finding \( q \) proportion of TB cases that would otherwise remain undetected by the passive case detection process:

\[
CDR_{\text{intervention}} = q \times (1 - CDR_{\text{baseline}}) + CDR_{\text{baseline}},
\]

where, \( CDR \) is the case detection rate, being the proportion of all cases detected during their disease episode and \( q \) is the proportion of cases detected under the intervention scenario that would otherwise be missed by the baseline CDR. We set the maximum attainable level of CDR in all intervention scenarios including targeting super-spreaders at 90%. In targeting the intervention towards super-spreaders, the maximum possible CDR targeted to super-spreaders under the intervention occurs when the CDR for non-spreaders and low-spreaders is kept at the baseline. With this restriction, we defined a single parameter \( (d) \) to represent the extent of intervention targeting relative to the maximum amount of targeting possible, which ranges from zero to one, with zero representing untargeted intervention and one representing 100% targeting. That is, the CDR targeted to super-spreaders under the intervention scenario was \( CDR_{\text{intervention}} + d \) multiplied by the difference between the maximum possible CDR and \( CDR_{\text{intervention}} \). This
implementation ensures equivalent effort or the number of people treated for untargeted and targeted interventions; details of these calculations are given in the *Supplementary Material*.

We then evaluated the effectiveness of targeting super-spreaders \((d>0)\) compared to untargeted intervention \((d = 0)\) in four independent scenarios. These four scenarios were settings with high transmission and high case detection, high transmission and low case detection, low transmission and high case detection, and low transmission and low case detection. These four scenarios were simulated by calibrating the transmission parameter for low-spreaders and CDR. In addition, we evaluated the impact of levels of targeting \((d)\) for different coverage levels of the intervention \((q)\) in these four setting scenarios, while other model parameters were sampled from the same plausible ranges as in the baseline sensitivity analysis using LHS.

**Results**

**Baseline sensitivity analyses**

The sensitivity analyses show that the most marked impact on equilibrium incidence arose from the rate of late TB progression followed by the CDR (*Fig 3*). The proportion of super-spreaders among persons with infectious TB and the relative infectiousness of super-spreaders compared to low-spreaders also had a substantial impact. *Supplementary Fig 2* presents a multidimensional sensitivity analysis, using the LHS method to sample 1000 parameter sets, which was consistent with the results of the one-way sensitivity analysis.

*Fig 3: One-way sensitivity analyses of equilibrium TB incidence to variation between extremes of plausible parameter values. Maximum and minimum parameter values are given in Table 1. Values on the x-axis represent the equilibrium TB incidence, and the vertical line indicates the equilibrium incidence obtained with the baseline parameter set.*
Targeting the active case finding intervention towards super-spreaders was more effective than mass intervention in all setting scenarios. However, its effectiveness was particularly marked in settings with low case detection, regardless of transmission rates. For example, in the first scenario of both high transmission and high CDR, a 20% untargeted active case finding intervention led to a 22.8% reduction in incidence over 20 years, while with an 80% targeting an equal active case finding intervention resulted in a 35.9% reduction in TB incidence over the same period. However, in the scenario of high transmission but low CDR, the targeted intervention is much more effective (achieving a 42.2% reduction) than the untargeted intervention, which only reduced incidence by 22.8% over 20 years, considering a 20% active case finding and an 80% level of targeting. Interventions in settings of low transmission and high case detection rates had a relatively minor impact on the burden.

We evaluated the impact of intensifying active case finding ($q$) from zero to 40% detection and varying the proportion of this intervention targeting the super-spreaders ($d$). As shown in Fig 4, both these quantities are essential – i.e. the increase in CDR and the extent of targeting provided the increase in CDR is more than negligible. The most significant impact of targeting was in the setting of high transmission and low CDR at all coverage levels of the intervention. Targeting super-spreaders was more effective at all levels of relative CDR increases in settings with high transmission and low case detection (Panel B), while in settings with high transmission and high case detection, targeting super-spreaders has a significant effect if the relative increase in CDR is low (Panel A).
Fig 4: Comparison of extent of active case finding and the proportion of targeting to super-spreaders on the 20-year projected relative reduction in TB incidence under the four scenarios. A) High transmission and high case detection setting; B) high transmission and low case detection; C) low transmission and high case detection; D) low transmission and low case detection.

Fig 5 presents the results of the one-way sensitivity analysis performed to observe the impact of high and low extremes of parameter values on the intervention’s 20-year projected relative incidence reduction. The sensitivity analyses show that the most marked impact on intervention effectiveness arose from the rate of late progression, followed by the rate of early progression and the proportion of super-spreaders. Notably, while higher rates of late progression were correlated with lower effectiveness of the active case finding intervention, higher rates of early progression resulted in greater reductions in the incidence.

Supplementary fig 4 presents the results of the multidimensional sensitivity analyses assessing the impact of parameter ranges on intervention effectiveness (relative reduction in annual incidence with 20-years projections). This multidimensional analysis supports the results of the one-way sensitivity analysis: that the proportion of super-spreaders had a considerable impact on intervention effectiveness, and was the most important parameter after the disease progression parameters.

Fig 5: One-way sensitivity analyses of 20-year projected relative reduction in incidence to variation between extremes of plausible parameter values. The maximum and minimum parameter values are given in Table 1. Values on the x-axis represent the relative reduction in incidence following the implementation of a 20% active case finding intervention with 80% targeting. The vertical line indicates the relative reduction in incidence obtained with the baseline parameter set.
Discussion

Our model suggested that targeted active case finding interventions directed toward people likely to be super-spreaders could have a substantial impact on the TB burden, particularly in settings with high transmission but low case detection rates. This model also showed that parameters related to infectiousness heterogeneity such as the proportion of super-spreaders and their relative infectiousness compared to other low-spreaders were as crucial as other previously recognised epidemiological parameters, such as the latency progression parameters, in determining the burden of TB. The current model is also able to incorporate a range of commonly used approaches to capturing heterogeneous infectiousness of active TB cases. The choice of three infectiousness levels for active TB and the values assigned to each compartment was based on empirical data estimates, strengthening the validity of our findings.

Active case finding has been described as “turning off the tap” in TB control intervention strategies since it represents a method of identifying individuals with TB and promptly initiating treatment to avoid further onward transmissions [50]. Previous TB modelling has suggested that active case finding with highly sensitive diagnostic tools can reduce delay to treatment and so significantly reduce TB transmission [51]. The use of point-of-care diagnostic technologies with high sensitivity to detect persons with TB soon after they begin to transmit, such as GeneXpert Omni [49], could enable the implementation of active case finding interventions similar to that considered in this study. In fact, any TB active case finding intervention necessarily incorporates some level of targeting towards the highly infectious individuals, because highly infectious patients’ disease characteristics (such as higher sputum-bacilli concentration and lung-cavitation) make them more easily detectable with existing microbiological and radiological diagnostic tools. This means that implementing more sensitive diagnostics may paradoxically decrease targeting of super-
spreaders – whereas traditional smear-based interventions might ensure that resources are targeted to the most infectious. In agreement with previous modelling [52], our model suggested that active case finding interventions are particularly valuable in high transmission and low case detection settings to limit onward transmission. In the implementation of the intervention, although we did not perform cost-effectiveness analyses, our study shows that finding super-spreaders (less than 10% of all TB cases) is equally as effective epidemiologically as finding 40% of general TB population. In addition, the cost of identifying and treating super-spreaders that only comprises less than 10% of total TB cases can be compared with the cost of treating all active TB cases in a population. Our findings concerning the effectiveness of interventions can provide broad directions to future TB modellers and policymakers, although predictions from our model are not intended to provide location-specific estimates.

Our sensitivity analyses showed that a higher late reactivation rate had a negative impact on the effectiveness of active case finding interventions, while higher rates of early progression increased the impact on the 20-year projected reduction of incidence following the intervention. This suggests that active case finding interventions are more effective if the TB epidemic is more dependent on early progression and intense recent transmission than late TB reactivation since the intervention would have a more significant effect on rapidly reducing transmission. A recent study showed that variation in the latency progression parameters had important impacts on model predictions around the effectiveness of preventive therapy interventions [53], while our analysis complements this finding by showing that these parameter variations also have indispensable effects on predictions regarding the effectiveness of case-finding interventions.

In our analyses, the rate of conversion from low-spreader to super-spreader had little significance, and impact of the rate of conversion from non-infectious to low-spreaders was
It is known that active TB cases’ ability to spread *Mtb* may increase as they progress clinically to more severe disease, e.g. from smear-negative to smear-positive or from non-cavitary to cavitary-TB [54]. In previous TB models, spontaneous conversion from less infectious to more infectious states such as smear-negative to smear-positive was included with rates around 1.5% per year [18, 25, 32, 45, 55-57]. These parameters do not affect model outputs substantially, such that they can be omitted for model simplicity if the actual values are consistent with what has previously been assumed. Nevertheless, the exact values of these quantities remain uncertain, and future research to refine these quantities may modify this conclusion.

The current model structure enhances flexibility around the assumption of infectiousness heterogeneity that allowed for reflection of a broad range of factors that can alter active TB cases infectivity level [5, 58]. Thus, the model is able to incorporate many previous compartmental modelling structural approaches, including those that stratified active TB cases’ infectivity into two levels, as non-infectious and infectious [13-17], or model structures that incorporate three levels of infectivity, as non-infectious, smear-negative infectious and smear-positive infectious [30, 31, 59]. However, in the application of this model to a particular epidemiological setting, data to estimate the proportion of super-spreaders is essential, since this information has a particularly marked impact on both baseline disease burden and intervention effectiveness.

We used empirical data to parameterise our model, but a noteworthy limitation is that these results may not be generalizable to other settings. However, we are not aware of any past work that has used empiric data to inform model parameters on the proportions of infectious TB and super-spreaders at all. As with any model-based analyses, our study has limitations that arise from its assumptions. In addition to those introduced above, our model is not intended to represent a specific setting, and as such does not incorporate stratification by age,
The other limitation is that we used a very simplified implementation of an active case finding intervention with a theoretical point-of-care diagnostic tool, without considering all the complexities in TB diagnosis and treatment implementation.

Conclusions

The approaches we used to inform model parameters related to infectiousness heterogeneity from the observed distribution of secondary infections can be useful for future modelling studies of TB, in particular, and other infectious diseases, in general. The principles of implementation could be used in future TB modelling studies that represent specific epidemiological settings, especially when TB contact investigation data are available to estimate the proportion of super-spreaders. In the usage of advanced point-of-care technologies, targeting active case-finding interventions to super-spreaders is likely to be especially beneficial in low CDR settings.

Acknowledgements

The authors would like to thank Monash University for providing the PhD scholarship to YAM.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

The R-code used to analyse this model is publicly available.

Authors’ contributions
YAM, RR and JMT conceived the study. YAM developed structure and ODE of the model, which was reviewed by JMT and RR. ESM, ACA, JMT conceived the strategy of the intervention in the model. YAM coded the model in R that RR and JMT reviewed and evaluated. YAM drafted the manuscript, and all authors provided input into revisions and approved the final draft for submission.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Supplementary material

Supplementary Material. doc

References

1. WHO. Global Tuberculosis Report 2019. Geneva: 2019.
2. WHO. The End TB Strategy: Global strategy and targets for tuberculosis prevention, care and control after 2015. Geneva, Switzerland 2014.
3. Herbert N, George A, Sharma V, Oliver M, Oxley A, Raviglione M, et al. World TB Day 2014: finding the missing 3 million. The Lancet. 2014;383(9922):1016-8.
4. Zumla A, George A, Sharma V, Herbert RHN, Oxley A, Oliver M. The WHO 2014 global tuberculosis report—further to go. The Lancet Global Health. 2015;3(1):e10-e2.
5. Melsew Y, Doan T, Gambhir M, Cheng A, McBryde E, Trauer J. Risk factors for infectiousness of patients with tuberculosis: a systematic review and meta-analysis. Epidemiol Infect. 2018:1-9.
6. Turner RD, Chiu C, Churchyard GJ, Esmail H, Lewinsohn DM, Gandhi NR, et al. Tuberculosis infectiousness and host susceptibility. The Journal of infectious diseases. 2017;216(suppl_6):S636-S43.
7. Issarow CM, Mulder N, Wood R. Environmental and social factors impacting on epidemic and endemic tuberculosis: a modelling analysis. Royal Society open science. 2018;5(1):170726.
8. Trauer JM, Dodd PJ, Gomes MGM, Gomez GB, Houben RMGJ, McBryde ES, et al. The Importance of Heterogeneity to the Epidemiology of Tuberculosis. Clin Infect Dis. 2018;69(1):159-66. doi: 10.1093/cid/ciy938.

9. Melsew YA, Gambhir M, Cheng AC, McBryde ES, Denholm JT, Tay EL, et al. The role of super-spreading events in Mycobacterium tuberculosis transmission: evidence from contact tracing. BMC Infect Dis. 2019;19(1):244. doi: 10.1186/s12879-019-3870-1.

10. Lloyd-Smith JO, Schreiber SJ, Getz WM, editors. Moving beyond averages: Individual-level variation in. Mathematical Studies on Human Disease Dynamics: Emerging Paradigms and Challenges: AMS-IMS-SIAM Joint Summer Research Conference on Modeling the Dynamics of Human Diseases: Emerging Paradigms and Challenges, July 17-21, 2005, Snowbird, Utah; 2006: American Mathematical Soc.

11. Ypma RJ, Altes HK, van Soolingen D, Wallinga J, van Ballegooijen WM. A sign of superspreading in tuberculosis: highly skewed distribution of genotypic cluster sizes. Epidemiology. 2013;24(3):395-400.

12. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. Nature. 2005;438(7066):355-9.

13. Aparicio JP, Castillo-Chavez C. Mathematical modelling of tuberculosis epidemics. Mathematical Biosciences & Engineering: MBE. 2009;6(2):209-37. PubMed PMID: 19364150.

14. Hickson R, Mercer G, Lokuge K, editors. Sensitivity analysis of a model for tuberculosis. 19th international congress on modelling and simulation; 2011.

15. Hickson RI, 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.

16. Korthals Altes H, Kloet S, Cobelens F, Bootsma M. Latent tuberculosis infection in foreign-born communities: Import vs. transmission in The Netherlands derived through mathematical modelling. [Erratum appears in PLoS One. 2018 May 24;13(5):e0198376; PMID: 29795693]. PLoS ONE [Electronic Resource]. 2018;13(2):e0192282. PubMed PMID: 29444122.

17. Thomas EG, Barrington HE, Lokuge KM, Mercer GN. Modelling the spread of tuberculosis, including drug resistance and HIV: a case study in Papua New Guinea’s Western Province. The ANZIAM Journal. 2010;52(1):26-45.

18. Dowdy DW, Chaixson RE, Maartens G, Corbett EL, Dorman SE. Impact of enhanced tuberculosis diagnosis in South Africa: a mathematical model of expanded culture and drug susceptibility testing. Proceedings of the National Academy of Sciences. 2008.

19. Dowdy DW, Chaixson RE, Moulton LH, Dorman SE. The potential impact of enhanced diagnostic techniques for tuberculosis driven by HIV: a mathematical model. 45AIDS. 2006;20(5):511-62. PubMed PMID: 16514306.

20. Dowdy DW, Davis JL, den Boon S, Walter ND, Katamba A, Cattamanchi A. Population-level impact of same-day microscopy and Xpert MTB/RIF for tuberculosis diagnosis in Africa. PLoS One. 2013;8(8):e70485.

21. Lin H-H, Dowdy D, Dye C, Murray M, Cohen T. The impact of new tuberculosis diagnostics on transmission: why context matters. Bull World Health Organ. 2012;90:739-47.

22. Oxlade O, Schwartzman K, Benedetti A, Pai M, Heymann J, Menzies D. Developing a tuberculosis transmission model that accounts for changes in population health. Med Decis Making. 2011;31(1):53-68.

23. Pandey S, Chadha V, Laxminarayan R, Arinaminpathy N. Estimating tuberculosis incidence from primary survey data: a mathematical modeling approach. The International Journal of Tuberculosis and Lung Disease. 2017;21(4):366-74.
Sun AY, Pai M, Salje H, Satyanarayana S, Deo S, Dowdy DW. Modeling the impact of alternative strategies for rapid molecular diagnosis of tuberculosis in Southeast Asia. Am J Epidemiol. 2013;178(12):s1740-9.

Houben R, Lalli M, Sumner T, Hamilton M, Pedrazzoli D, Bonsu F, et al. TIME Impact—a new user-friendly tuberculosis (TB) model to inform TB policy decisions. BMC Med. 2016;14(1):56.

Menzies NA, Cohen T, Hill AN, Yaesoubi R, Galer K, Wolf E, et al. Prospects for tuberculosis elimination in the United States: Results of a Transmission Dynamic Model. Am J Epidemiol. 2018.

Menzies NA, Cohen T, Lin H-H, Murray M, Salomon JA. Population health impact and cost-effectiveness of tuberculosis diagnosis with Xpert MTB/RIF: a dynamic simulation and economic evaluation. PLoS Med. 2012;9(11):e1001347.

Garcia A, Maccario J, Richardson S. Modelling the annual risk of tuberculosis infection. Int J Epidemiol. 1997;26(1):190-203. PubMed PMID: 9126520.

Dowdy DW, Andrews JR, Dodd PJ, Gilman RH. A user-friendly, open-source tool to project impact and cost of diagnostic tests for tuberculosis. Elife. 2014;3:e02565.

Arregui S, Iglesias MJ, Samper S, Marinova D, Martin C, Sanz J, et al. Data-driven model for the assessment of Mycobacterium tuberculosis transmission in evolving demographic structures. Proc Natl Acad Sci U S A. 2018;115(14):E3238-E45. PubMed PMID: 29563223.

Dowdy DW, Lotia I, Azman AS, Creswell J, Sahu S, Khan AJ. Population-level impact of active tuberculosis case finding in an Asian megacity. PLoS One. 2013;8(10):e77517.

Salomon JA, Lloyd-Smith JO, Getz WM, Resch S, Sánchez MS, Porco TC, et al. Prospects for advancing tuberculosis control efforts through novel therapies. PLoS Med. 2013;10(8):e273.

Vynnycky E, Sumner T, Fielding KL, Lewis JJ, Cox AP, Hayes RJ, et al. Tuberculosis control in South African gold mines: mathematical modeling of a trial of community-wide isoniazid preventive therapy. Am J Epidemiol. 2015;181(8):619-32.

McCreesh N, White RG. An explanation for the low proportion of tuberculosis that results from transmission between household and known social contacts. Sci Rep. 2019;8(15):5382.

Melsew YA, Cheng AC, McBryde ES, Denholm JT, Tay E, Ragonnet R, et al. Profiles of tuberculosis disease activation among contacts of patients with tuberculosis. Eur Respir J. 2019;54(4):1900353.

Melsew YA, Adekunle AI, Cheng AC, McBryde ES, Ragonnet R, Trauer JM. Heterogeneous infectiousness in mathematical models of tuberculosis: A systematic review. EBioMedicine. 2020;50:102676.

Moyo N, Tay E, Denholm J. Evaluation of tuberculin skin testing in tuberculosis contacts in Victoria, Australia, 2005–2013. Public health action. 2015;5(3):188-93.

RCORE TEAM. R: A Language and Environment for Statistical Computing. Vienna, 2014.

WHO. ETHIOPIA: Factsheet of Health Statistics 2016. 2016.
43. Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJ. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV-negative patients: a systematic review. PLoS One. 2011;6(4):e17601.

44. Trauer JM, Denholm JT, McBryde ES. Construction of a mathematical model for tuberculosis transmission in highly endemic regions of the Asia-Pacific. J Theor Biol. 2012;358:74-84.

45. Hughes GR, Currie CS, Corbett EL, editors. Modeling tuberculosis in areas of high HIV prevalence. Simulation Conference, 2006 WSC 06 Proceedings of the Winter; 2006: 521-524 IEEE.

46. Andrews JR, Noubary F, Walensky RP, Cerda R, Losina E, Horsburgh CR. Risk of progression to active tuberculosis following reinfection with Mycobacterium tuberculosis. Clin Infect Dis. 2012;54(6):784-91.

47. Drake JM, Rohani P. Sensitivity analysis of deterministic models through Latin hypercube sampling: A model for the spread of Ebola virus disease2015. Available from: https://daphnia.ecology.uga.edu/drakelab/wp-content/uploads/2015/07/sensitivity-ebola.pdf.

48. Gomero B. Latin Hypercube Sampling and Partial Rank Correlation Coefficient Analysis Applied to an Optimal Control Problem: University of Tennessee; 2012.

50. FIND Ca. World's most portable molecular diagnostics system unveiled at AACC. 2015 [cited 2019 May 14]. Available from: https://www.prnewswire.com/news-releases/worlds-most-portable-molecular-diagnostics-system-unveiled-at-aacc-300119213.html.

51. Yuen CM, Amanullah F, Dharmadhikari A, Nardell EA, Seddon JA, Vasilyeva I, et al. Turning off the tap: stopping tuberculosis transmission through active case-finding and prompt effective treatment. The Lancet. 2015;386(10010):2334-43.

52. Uys PW, Warren R, van Helden PD, Murray M, Victor TC. Potential of rapid diagnosis for controlling drug-susceptible and drug-resistant tuberculosis in communities where Mycobacterium tuberculosis infections are highly prevalent. J Clin Microbiol. 2009;47(5):1484-90.

53. Dowdy DW, Basu S, Andrews JR. Is passive diagnosis enough? The impact of subclinical disease on diagnostic strategies for tuberculosis. Am J Respir Crit Care Med. 2013;187(5):543-51.

54. Sumner T, White RG. Modelling the impact of tuberculosis preventive therapy: the importance of disease progression assumptions. bioRxiv. 2019:666669.

55. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. Nature Reviews Disease Primers. 2016;2:16076. doi: 10.1038/nrdp.2016.76.

56. Knight GM, Griffiths UK, Sumner T, Laurence YV, Gheorghe A, Vassall A, et al. Impact and cost-effectiveness of new tuberculosis vaccines in low-and middle-income countries. Proceedings of the National Academy of Sciences. 2014;111(43):15520-5.

57. Legrand J, Sanchez A, Le Pont F, Camacho L, Larouze B. Modeling the impact of tuberculosis control strategies in highly endemic overcrowded prisons. PLoS One. 2013;8(5):e62100.

58. Mellor GR, Currie CS, Corbett EL. Incorporating household structure into a discrete-event simulation model of tuberculosis and HIV. ACM Transactions on Modeling and Computer Simulation (TOMACS). 2011;21(4):26.

59. Grassly NC, Fraser C. Mathematical models of infectious disease transmission. Nature Reviews Microbiology. 2008;6(6).

60. Abu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini IM, Dye C, et al. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. Proceedings of the National Academy of Sciences. 2009;106(33):13980-5.
Fig 1: Transmission dynamic model of TB with heterogeneous immunity.
Fig 2: Comparison of the effective modelled secondary infections
Fig 3: One-way sensitivity analyses of equilibrium TB incidence to

- Rate of late TB progression, $v$
- Case detection rate, CDR
- Rate of early TB progression, $\epsilon$
- Proportion super-spreaders, among infectious
- Transmission rate per low-spreader, $\beta_1$
- Relative infectiousness of super-spreaders, $m$
- Proportion of infectious TB
- Rate of conversion from low-spreaders to super-spreaders, $j$
- Natural mortality rate, $\mu$
- Rate of conversion from non-infectious to low-spreaders, $h$
- TB-related mortality, super-spreaders, $\mu_{l2}$
- TB-related mortality, non-spreaders, $\mu_{l0}$
- TB-related mortality, low-spreaders, $\mu_{l1}$

Equilibrium incidence per 100,000 population

[Chart showing sensitivity analyses with high and low value bars]
Fig 4: Comparison of extent of active case finding and testing for relative CDR increase.
Fig 5: One-way sensitivity analyses of 20-year projected relative reduction incidence.