Modeling the impact of case finding for tuberculosis: The role of infection dynamics

Authors: Theresa S Ryckman[1], Sourya Shrestha[2], Anthony T Fojo[3], Parastu Kasaie[2], David W Dowdy[2], Emily A Kendall[1]

[1]Department of Medicine, Division of Infectious Diseases, Johns Hopkins School of Medicine, Baltimore, Maryland 21287

[2]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205

[3]Department of Medicine, Division of General Internal Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland 21287

Corresponding Author: Theresa S Ryckman; 601 North Wolfe Street, Carnegie Building #380, Baltimore, Maryland, 21287, United States; +1-920-840-4123; tryckma1@jh.edu

Abstract Word Count: 244/250 words

Manuscript Word Count: 2942/3500 words

Abstract

Background: Our understanding of *Mycobacterium tuberculosis (Mtb)* natural history and infection dynamics is evolving, including recognition that many individuals previously infected with *Mtb* may clear their infections or experience substantially reduced progression risks with time since infection. Such dynamics suggest that recent transmission is more important in driving TB incidence in high-burden settings than previously estimated; thus, the impact of interventions to reduce transmission (e.g., community-based active case-finding) may also be greater than previously thought.

Methods: We constructed two models of *Mtb* transmission that differed only in that one model included a clearance mechanism while the other did not. We then calibrated these models independently to the same set of epidemiological data representative of a high-TB-burden setting (India). Finally, we used the calibrated models to project the impact of illustrative biennial active case-finding campaigns (achieving 75% population coverage with 65% screening sensitivity).

Results: The model that included a clearance mechanism projected a greater impact of case-finding on the incidence of TB disease: 45% [95% uncertainty interval 28-57%] reduction versus no intervention after 10 years, versus 11% [6-18%] in the model without a clearance mechanism. The estimated annual risk of *Mtb* infection and prevalence of recent infection were both substantially higher in the model that allowed for *Mtb* clearance, despite being fit to the same data.

Conclusions: Models that allow for *Mtb* clearance are supported by biological and epidemiological evidence and project greater impact from active case-finding than models that do not include these dynamics.

Keywords: Tuberculosis, Mathematical Modeling, Transmission, Screening, Infection Dynamics

Background

Interventions to reduce the global burden of tuberculosis (TB; 10.8 million new cases and 1.3 million deaths in 2023) are urgently needed.[1] In prioritizing between interventions to find, treat, or prevent TB, it is important to understand that underlying dynamics of infection and progression affect the relative impact of these interventions. Interventions that detect people with undiagnosed TB and link them to treatment (such as community-based "active case-finding") are likely to have greater impact in epidemics where most TB disease is caused by recent transmission events. If most TB incidence instead arises from reactivation of *Mycobacterium tuberculosis* (*Mtb)* infections that were transmitted in the remote past, then finding and treating people who have TB disease now is unlikely to have major effects on subsequent TB incidence, because TB will continue to develop from within the current pool of remote infections.

Community-based active case-finding programs are globally recommended and are being increasingly prioritized in high-burden settings.[2–5] Empirical evidence on the impact of active case-finding has been mixed and is challenging to obtain, owing to the need for very large sample sizes to demonstrate effects at a population level.[6] At least one case-finding study, the ACT3 randomized trial in Vietnam, detected a large and significant reduction in prevalence.[7] Mathematical models, which are often used to bridge the gap between empirical evidence and policy decisions, have yielded varied estimates of the long-term impact of case-finding. Several modeling studies have projected that case-finding will have only limited and temporary effects on TB incidence, [8–11] while others predict greater reductions.[12]

Model-driven estimates of the impact of case-finding may depend on how these models represent underlying *Mtb* infection dynamics – and our understanding of these dynamics is evolving. It is increasingly understood that *Mtb* infection is more complex and dynamic than the conventional concept of a state of lifelong "latency", and that immunoreactivity to *Mtb* antigens may only weakly reflect current infection.[13–16] Specifically, numerous pieces of evidence suggest that the risk of progression to active disease continues to decline substantially beyond the first 5-10 years after infection.[17] For example, data on progression from longitudinal cohorts with *Mtb* infection in low-burden settings, along with models fit to cross-sectional data, are consistent with substantially declining progression rates over time.[18,19] Furthermore, high rates of immunoreactivity reversion and low-to-moderate rates of progression among

immunoreactive people after immunosuppression suggest that a meaningful proportion of *Mtb* infections clear over time. In people with cleared infections, immunoreactivity may indicate immunological memory rather than persistent infection with viable bacteria capable of causing reactivation disease.[20,21]

In this study, we sought to use mathematical models to understand the implications of our updated, more nuanced conceptualization of *Mtb* infection - in terms of estimating the importance of recent transmission in perpetuating TB epidemics in high-burden settings and projecting the impact of TB active case-finding.

Methods

Overview

We took a previously published dynamic compartmental *Mtb* transmission model ("Comparator Model") and created an alternative version of the model structure, which differed only in the addition of a mechanism by which *Mtb* infections can clear ("Clearance Model").[8] We then calibrated both models to the same empirical epidemiological data from a high-burden setting (India). We subsequently used the calibrated models to project the impact of an illustrative active case-finding intervention, covering 70% of the population with a 65%-sensitive screening algorithm every 2 years. We compared simulated infection dynamics and projected intervention impact between the two models.

Key Terms

We use the following definitions in this exercise:

- ***Mtb* Infection:** Exposure to *Mtb* at a previous point in life, resulting in a positive tuberculin skin test (TST) or interferon gamma release assay (IGRA) if tested, with ongoing risk of progressing to TB disease.
- **TB Disease:** Disease resulting in bacteriologically positive sputum, regardless of symptoms.
- **Immunoreactivity:** A positive TST or IGRA result, regardless of progression risk, prior TB disease, or infection status.
- **Cleared Infection:** Prior *Mtb* infection that is no longer at risk of progressing to TB disease. Cleared infections are assumed not to still generate immunoreactivity (an assumption we tested in sensitivity analysis).

Model Structure

The Comparator Model followed a typical *Mtb* transmission model structure (Figure 1, Supplementary Methods, Supplementary Figure 1), in which a population transitions between six TB-related states. In this model, progression risk is highest immediately after *Mtb* infection (represented by a “recent infection” state with a higher progression rate) but persists for life among all who have been infected (represented by a “remote infection” state with a lower progression rate). In the alternative Clearance Model, we allowed for transition from the two infection states to a seventh, “cleared” state with no probability of progression (unless they experience reinfection). This clearance rate was calibrated.

In both models, we assumed that individuals with new TB disease start in a non-care-seeking phase, which can subsequently resolve to the recovered state or progress to symptom-driven care seeking. Individuals in the care-seeking phase can subsequently be diagnosed and treated/cured or die from TB; for simplicity in this model, we assumed that individuals who enter the care-seeking phase do not return to the non-care-seeking phase. Both models were further stratified into two age groups: children <15 years and adults ≥15 years, who face differential transmission, progression and mortality risks. Active case-finding and TPT interventions were simulated via movement to the recovered state from all TB disease states and all *Mtb* infection states, respectively.

Model Calibration

Model parameters were calibrated to prevalence-survey-based and WHO-reported disease burden data from India (selected as an illustrative but representative high-burden country; Supplementary Methods, Supplementary Table 1), plus historical cohort data on progression over time after new *Mtb* infection (Table 1, Supplementary Tables 2-3). India-specific targets included WHO-estimates of TB incidence, the rate of change in incidence over time, and the case fatality ratio[22], as well as estimates from India’s 2019-21 national TB prevalence survey of the prevalence of TB disease (in absolute terms and relative to annual notifications) and the prevalence of *Mtb* immunoreactivity.[23]

Additional calibration targets were based on empirical estimates of time to progression to TB disease after infection, from two historical cohorts of TST-positive individuals with presumed new infection (based on recent exposure).[24–26] We extracted estimates of cumulative progression in the first 2 years of infection and in years 3 through 10 after infection (chosen to inform recent and remote progression rates). Since these studies were conducted in low-prevalence settings, we assumed that all TB in these cohorts resulted from progression rather than reinfection.

After a burn-in period to reach a steady state, we induced declines in incidence over 20 years to match data on temporal trends in TB burden. In the main analysis, we used an annual linear increase in the treatment initiation rate to generate these declines, reflecting expansions in treatment that occurred over recent decades. In secondary analysis, we considered instead a declining transmission rate, reflecting development-related changes in opportunities for transmission or susceptibility to infection. The target for annual reductions in TB incidence was compared to the average rate of change over this 20-year period as estimated by WHO, and all other calibration targets were compared to model outputs at the end of these 20 years of incidence decline.

The models were coded in R v4.5.1 and calibrated using Bayesian Incremental Mixture Importance Sampling with uninformed priors (Supplementary Methods), resulting in 20,000 posterior parameter set samples per model.[27] The model code is available at www.github.com/rycktessman/tb-model-clearance.

Infection and Disease Dynamics

We first assessed key metrics of infection dynamics absent intervention, including the proportion of prevalent *Mtb* infections acquired recently, the proportion of TB incidence arising from recently versus remotely acquired infections, and the proportion of a given year's TB incidence that could be attributed to transmission from the current population with prevalent TB disease. The latter was estimated by instantaneously moving all individuals with TB disease to the recovered compartment and running the model for one year. We also examined how key quantities not used in calibration, such as the annual risk of infection and the duration of TB disease, varied across the two models.

Interventions

We used the calibrated models to simulate the impact of community-wide active case-finding campaigns conducted every 2 years, covering 70% of the population during each campaign with a 65% sensitive screening algorithm (Supplementary Methods; Supplementary Table 4). We evaluated intervention impact on TB incidence over a 10-year horizon and compared the projected impacts of case-finding between models. We also evaluated additional outcomes (prevalence, mortality) and considered a scenario in which TPT provision was added to the case-finding campaign.

Sensitivity Analysis

We conducted probabilistic sensitivity analyses to evaluate uncertainty across all model parameters. Main results are presented as mean values, with 95% uncertainty intervals (UI) representing the 2.5th and 97.5th percentiles across 20,000 parameter set samples. We also evaluated correlations between parameters and modeled outcomes, examined variation in projected intervention impact between the lowest and highest deciles of each parameter's distribution (as a form of multivariable one-way sensitivity analysis), and conducted two-way sensitivity analysis on the percentages of TB disease and *Mtb* infection cured via intervention.

Results

Model Calibration

Both models achieved reasonable fits to most of the nine empirical calibration targets (Table 1, Supplementary Figure 2). The Comparator Model was unable to replicate pre-intervention declines in incidence and tended to overestimate immunoreactivity prevalence and late progression, while the Clearance Model (when assuming no immunoreactivity among people with cleared infections) estimated somewhat lower immunoreactivity prevalence than measured by the prevalence survey.

The calibrated effective contact rate ("beta") was 4 times higher in the Clearance Model than the Comparator Model, suggesting more transmission is required to achieve consistency with epidemiological targets when some individuals clear their infections (Supplementary Table 5). In the Clearance Model, the clearance rate corresponded to approximately half of people who had not yet progressed clearing their

130 infection within 5 years. In both models, the recent infection period lasted an average of 5-6 months, but
131 the proportion of progression that occurred during that period was greater in the Clearance Model.

132

133 *Infection and Disease Dynamics*

134 Recent infection accounted for a greater share of total *Mtb* infection prevalence under the Clearance
135 Model (8% [4-15%]) versus the Comparator Model (1% [1-2%]; Figure 2). Since progression rates from
136 recent infection are higher than remote reactivation rates, this resulted in a substantially greater
137 percentage of TB incidence attributable to progression from recent infection in the Clearance Model: 61%
138 [35-79%], versus 28% [14-46%] in the Comparator Model. Similarly, under the Clearance Model, 32% [20-
139 44%] of incidence in a given year resulted from transmission from the current population with prevalent
140 TB (i.e., at the beginning of that year), versus 14% [8-23%] under the Comparator Model.

141 Consistent with infection from recent transmission playing a greater role when clearance is considered,
142 the annual risk of infection was also higher under the Clearance Model (4.8% [2.5-7.3%]) than the
143 Comparator Model (1.1 % [0.8-1.5%]) (Supplementary Table 6). The duration of TB disease, estimated
144 both with current treatment intensity and absent any treatment, was similar under both models
145 (Supplementary Table 7).

146 *Intervention Impact*

147 Screening 70% of the population every 2 years with a 65% sensitive screening algorithm was projected to
148 yield a significantly greater reduction in incidence by year 10 under the Clearance Model (45% [28-57%])
149 than the Conventional Model (11% [6-18%]) (Figure 3). For the combined intervention of both active case-
150 finding and TPT, impact was similar across models (all 95% UIs within 79-85%; Supplementary Figure 3).
151 Corresponding prevalence and mortality estimates over time are shown in Supplementary Figures 4-5.

152 *Sensitivity Analyses*

153 We conducted sensitivity analyses that varied the mechanism of pre-intervention incidence declines (from
154 treatment- to susceptibility-mediated) and the proportion of cleared infections that remain immunoreactive

155 (from 0 to 100%; calibration data shown in Supplementary Figures 6-7). Across these analyses, infection
156 from recent transmission continued to make up a greater share of *Mtb* infection prevalence and TB
157 disease incidence under the Clearance Model than the Comparator Model, and annual risk of infection
158 was consistently higher (Supplementary Tables 6-7). The greater projected impact from active case-
159 finding under the Clearance Model was also consistent across these sensitivity analyses (Supplementary
160 Fig 8-9), and was similarly consistent when varying the coverage and efficacy of case-finding
161 (Supplementary Figure 10). The impact of case-finding on incidence was most sensitive to the
162 progression rate from remote infection; other influential parameters included the effective contact rate, the
163 spontaneous resolution rate, and the treatment rate (Supplementary Figures 11-12).

164 **Discussion**

165 A more nuanced conceptualization of *Mtb* infection that allows for the possibility of infection clearance is
166 starting to be incorporated into mathematical models,[12] and studies have explored its implications for
167 natural history trajectories[16], lifetime risk of disease,[28] and vaccine trials[14]. Here, we expand on this
168 understanding by demonstrating the implications for disease dynamics and the impact of interventions.
169 We found that incorporating clearance into a typical *Mtb* transmission model is consistent with a greater
170 role of recent transmission in sustaining TB epidemics, implying that interventions which interrupt that
171 transmission – for example, by finding more people with TB and linking them to effective treatment –
172 could have more substantial and lasting epidemiological impact than previously estimated. Importantly,
173 mass TPT provision enhanced impact in both models, but was not necessary to achieve large reductions
174 in incidence under the Clearance Model.

175 To understand these results, one must consider how models that do not include infection clearance must
176 operate to fit observed data. Since a large fraction of the population is assumed to have ongoing, lifelong
177 reactivation risk, the majority of incident TB must derive from the pool of remote infections. This pool
178 cannot be reduced by transmission-reducing interventions like case-finding. By contrast, if a substantial
179 fraction of infections clear over time, a greater fraction of TB incidence arises from recent transmission -
180 which can be interrupted by interventions like case-finding. This behavior also explains why the Clearance
181 Model was better able to replicate empirical estimates of declines in incidence that have occurred over

182 the last 20 years than the Comparator Model, under which it is difficult to achieve incidence reductions
183 absent interventions that reduce progression. Importantly, while we chose to model infection clearance,
184 any mechanism that continues to reduce the risk of progression to disease beyond the first few years of
185 infection, or shifts the balance of progression from remote to recent infection in other ways, is expected to
186 have a similar effect.

187 This exercise was not intended to provide precise estimates of the impact of any specific case-finding
188 intervention. Findings regarding the greater importance of recent transmission and greater impact of
189 case-finding when clearance is considered were consistent across a range of intervention coverages and
190 efficacies, demonstrating that our conclusions are not sensitive to intervention-related parameters. As
191 anticipated, we projected a greater impact from case-finding than previous models that did not include a
192 clearance (or similar) mechanism.[8–11] The annual risk of infection estimated by the Clearance Model
193 exceeded an estimate from India based on TST surveys among children.[29] However, a recent review
194 found that cross-sectional surveys among children likely underestimate the true annual risk of infection in
195 high-burden settings, which may be closer to 5-10% (consistent with our estimates from the Clearance
196 Model).[30]

197 Although we calibrated to data from India, these findings are expected to generalize to other high-burden
198 settings with low HIV prevalence. The balance of key targets that influenced model behavior -
199 immunoreactivity (or infection) prevalence, incidence, and progression by time since infection – are
200 similar in other high-TB-burden countries (Supplementary Table 1). Some favor recent infection even
201 more than India's data – with higher incidence, a lower ratio of infection prevalence to incidence, and/or
202 greater annual reductions in incidence (where adequate data exist to inform trend estimates).[22,31] Other
203 unique epidemiologic features of India's TB epidemic, such as a high prevalence-to-notification ratio, were
204 not major drivers of calibration fit or projected intervention impact. On the other hand, these results may
205 not generalize to settings with high TB-HIV co-prevalence. A setting-specific model that represents TB-
206 HIV dynamics would be needed to apply the structural questions considered here to a high HIV-
207 prevalence setting.

208 Our findings should be interpreted in light of certain limitations. First, our main analysis assumed that
209 cleared infections no longer exhibit immunoreactivity. While the percentage of cleared infections that
210 maintain immunoreactivity remains uncertain, we found consistently higher (>3x) impact of case-finding
211 under the Clearance Model regardless of this percentage (varied from 0% to 100% in sensitivity analysis).
212 Second, our model's TB disease states only captured individuals with sputum-positive TB. Alternative
213 model structures that explicitly represent a sputum-negative TB disease state exhibit different underlying
214 dynamics (such as a higher clearance rate than estimated here)[16] but similarly project a greater impact of
215 active case-finding than many other modeling exercises.[12] Finally, all models are inherently simplified
216 representations of complex dynamics - but we believe such simplified representations are useful when
217 trying to isolate the effect of a single mechanism, such as that of infection clearance explored here.

218 Numerous pieces of evidence suggest that, rather than *Mtb* infection representing a protracted state of
219 persistent reactivation risk, many people either clear their *Mtb* infections or stabilize to a very low risk of
220 reactivation within a relatively short time after infection. This fundamental reconceptualization of *Mtb*
221 infection dynamics has important implications for the impact of transmission-reducing interventions.
222 Specifically, remote infections likely play less of a role in driving the ongoing burden of TB in high-
223 incidence settings than has previously been thought, implying that interventions that reduce transmission
224 (e.g., case-finding) may have greater impact than previously estimated. These findings support the
225 continued rollout of setting-specific community case-finding efforts, along with targeted programs
226 delivered to people at high risk of progression (such as household contacts, people with HIV, and people
227 with undernutrition), as part of a comprehensive intervention package to reduce the global burden of TB.

References

1 World Health Organization. Global Tuberculosis Report 2024. 2024 https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2024 (accessed Dec 17, 2024).

2 Biermann O, Tran PB, Viney K, Caws M, Lönnroth K, Sidney Annerstedt K. Active case-finding policy development, implementation and scale-up in high-burden countries: A mixed-methods survey with National Tuberculosis Programme managers and document review. *PloS One* 2020; **15**: e0240696.

3 Pande T, Vasquez NA, Cazabon D, *et al.* Finding the missing millions: lessons from 10 active case finding interventions in high tuberculosis burden countries. *BMJ Glob Health* 2020; **5**. DOI:10.1136/bmjgh-2020-003835.

4 Nguyen T-A, Teo AKJ, Zhao Y, *et al.* Population-wide active case finding as a strategy to end TB. *Lancet Reg Health – West Pac* 2024; **46**. DOI:10.1016/j.lanwpc.2024.101047.

5 World Health Organization. Consolidated guidelines on tuberculosis module 2: systematic screening for tuberculosis disease. 2021; published online March 22. https://www.who.int/publications/i/item/9789240022676 (accessed Oct 28, 2025).

6 MacPherson P, Shanaube K, Phiri MD, *et al.* Community-based active-case finding for tuberculosis: navigating a complex minefield. *BMC Glob Public Health* 2024; **2**: 9.

7 Marks GB, Nguyen NV, Nguyen PTB, *et al.* Community-wide Screening for Tuberculosis in a High-Prevalence Setting. *N Engl J Med* 2019; published online Oct 2. DOI:10.1056/NEJMoa1902129.

8 Shrestha S, Kendall EA, Chang R, *et al.* Achieving a “step change” in the tuberculosis epidemic through comprehensive community-wide intervention: a model-based analysis. *BMC Med* 2021; **19**: 244.

9 Ragonnet R, Williams BM, Largen A, *et al.* Estimating the long-term effects of mass screening for latent and active tuberculosis in the Marshall Islands. *Int J Epidemiol* 2022; **51**: 1433–45.

10 Houben RMGJ, Menzies NA, Sumner T, *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**: e806–15.

11 Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health* 2013; **34**: 271–86.

12 Schwalb A, Horton KC, Emery JC, *et al.* Potential impact, costs, and benefits of population-wide screening interventions for tuberculosis in Viet Nam: A mathematical modelling study. *PLOS Glob Public Health* 2025; **5**: e0005050.

13 Behr MA, Kaufmann E, Duffin J, Edelstein PH, Ramakrishnan L. Latent Tuberculosis: Two Centuries of Confusion. *Am J Respir Crit Care Med* 2021; **204**: 142–8.

14 Scarponi D, Clark RA, Weerasuriya CK, *et al.* Is neglect of self-clearance biasing TB vaccine impact estimates? *BMJ Glob Health* 2023; **8**. DOI:10.1136/bmjgh-2023-012799.

15 Dale KD, Schwalb A, Coussens AK, *et al.* Overlooked, dismissed, and downplayed: reversion of Mycobacterium tuberculosis immunoreactivity. *Eur Respir Rev Off J Eur Respir Soc* 2024; **33**: 240007.

16 Horton KC, Richards AS, Emery JC, Esmail H, Houben RMGJ. Reevaluating progression and pathways following Mycobacterium tuberculosis infection within the spectrum of tuberculosis. *Proc Natl Acad Sci U S A*; **120**: e2221186120.

17 Behr MA, Edelstein PH, Ramakrishnan L. Revisiting the timetable of tuberculosis. *BMJ* 2018; **362**: k2738.

18 Dale KD, Karmakar M, Snow KJ, Menzies D, Trauer JM, Denholm JT. Quantifying the rates of late reactivation tuberculosis: a systematic review. *Lancet Infect Dis* 2021; published online April 20. DOI:10.1016/S1473-3099(20)30728-3.

19 Menzies NA, Swartwood N, Testa C, *et al.* Time Since Infection and Risks of Future Disease for Individuals with Mycobacterium tuberculosis Infection in the United States. *Epidemiology* 2021; **32**: 70.

20 Behr MA, Edelstein PH, Ramakrishnan L. Is Mycobacterium tuberculosis infection life long? *BMJ* 2019; **367**: l5770.

21 Wang Z, Posey DL, Brostrom RJ, Morris SB, Marano N, Phares CR. US Postarrival Evaluation of Immigrant and Refugee Children with Latent Tuberculosis Infection Diagnosed Overseas, 2007-2019. *J Pediatr* 2022; **245**: 149-157.e1.

22 World Health Organization. Global Tuberculosis Report 2023. 2023; published online Nov. https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023 (accessed Nov 30, 2023).

23 Indian Council of Medical Research (ICMR), ICMR-National Institute for Research in Tuberculosis, Ministry of Health and Family Welfare, Government of India, Central TB Division and National Tuberculosis Elimination Programme, World Health Organization, India Office, State TB Cells of all States and Union Territories, India. National TB Prevalence Survey in India (2019-2021). 2022; published online March. https://tbcindia.gov.in/showfile.php?lid=3659 (accessed April 19, 2023).

24 Menzies NA, Wolf E, Connors D, *et al.* Progression from latent infection to active disease in dynamic tuberculosis transmission models: a systematic review of the validity of modelling assumptions. *Lancet Infect Dis* 2018; **18**: e228–38.

25 Ferebee SH, Mount FW. Tuberculosis Morbidity in a Controlled Trial of the Prophylactic Use of Isoniazid among Household Contacts1. *Am Rev Respir Dis* 1962; **85**. https://www.atsjournals.org/doi/epdf/10.1164/arrd.1962.85.4.490?role=tab (accessed April 19, 2023).

26 Sutherland I. The ten-year incidence of clinical tuberculosis following ‘conversion’ in 2550 individuals aged 14 to 19 years. 1968.

27 Raftery AE, Bao L. Estimating and Projecting Trends in HIV/AIDS Generalized Epidemics Using Incremental Mixture Importance Sampling. *Biometrics* 2010; **66**: 1162–73.

28 Emery JC, Richards AS, Dale KD, *et al.* Self-clearance of Mycobacterium tuberculosis infection: implications for lifetime risk and population at-risk of tuberculosis disease. *Proc R Soc B Biol Sci* 2021; **288**: 20201635.

29 Chadha VK, Sarin R, Narang P, *et al.* Trends in the annual risk of tuberculous infection in India. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2013; **17**: 312–9.

30 Dowdy DW, Behr MA. Are we underestimating the annual risk of infection with Mycobacterium tuberculosis in high-burden settings? *Lancet Infect Dis* 2022; **22**: e271–8.

31 Houben RMGJ, Dodd PJ. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLOS Med* 2016; **13**: e1002152.

Funding: This work was supported by the US National Institutes of Health, National Institute of Allergy and Infectious Diseases (NIH K01AI182503 to TSR), National Heart Lung and Blood Institute (NIH R01HL153611 to EAK), and the Johns Hopkins University Tuberculosis Research Advancement Center (NIH P30AI168436 to PK and TSR). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest: The authors do not have commercial or other associations that might pose a conflict of interest.

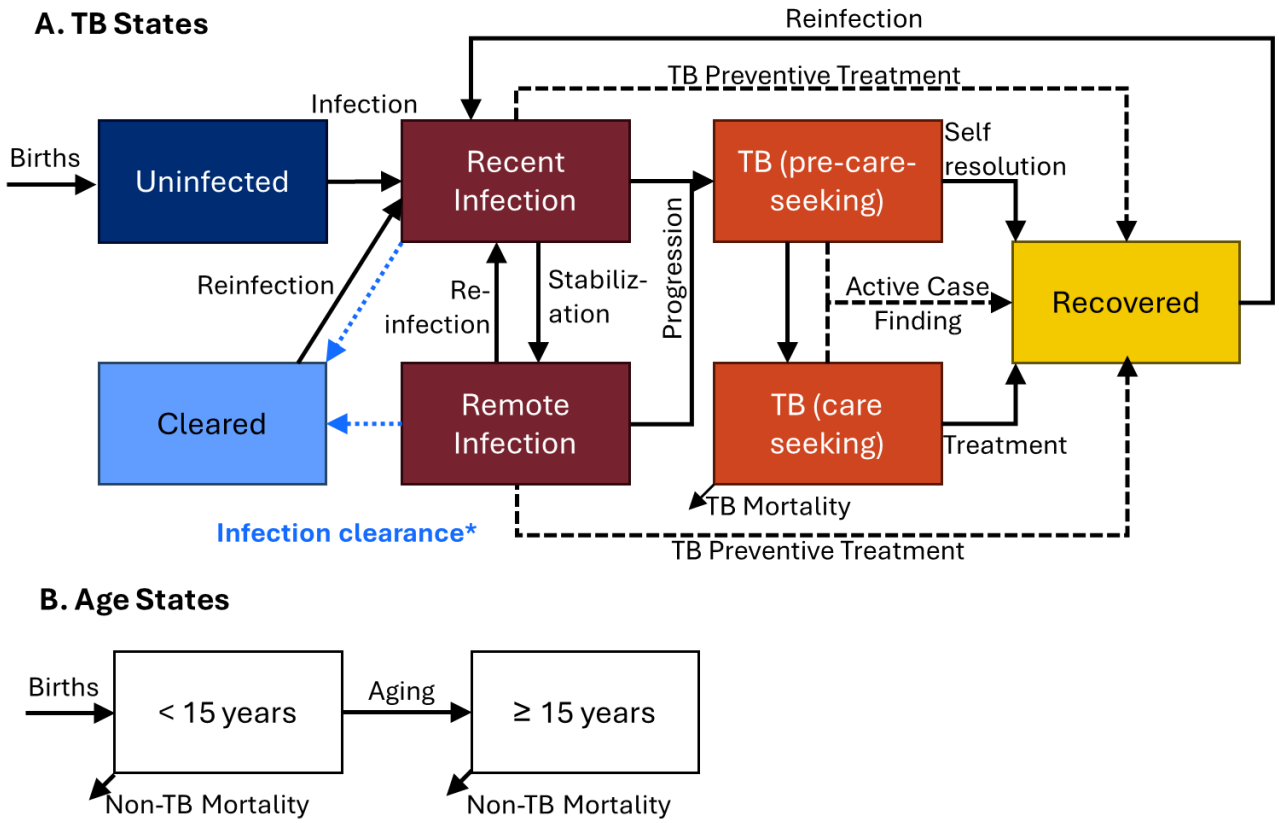
Acknowledgements: We would like to thank Nicolas Menzies for sharing a copy of the Sutherland et al. study (reference 26), data from which was used to quantify the early and late progression targets to which we calibrated the model.

Author Contributions: TSR, EAK, and DWD conceptualized the study, and SS, ATF, and PK contributed to initial discussions on motivation and scope. TSR curated and cleaned the data, ran all analyses, and wrote the first draft of the manuscript. SS provided initial model code that was adapted by TSR. EAK provided ongoing input and suggestions on methods and preliminary analyses. EAK, DWD, SS, ATF, and PK contributed to discussions of preliminary results and subsequent model drafts. EAK and TSR verified the study data. All authors had full access to all the data in the study and accept responsibility to submit for publication.

Data Availability: Model code is published online at www.github.com/rycktessman/tb-model-clearance. All data used in the study are displayed and cited in Table 1, Supplementary Table 3, and Supplementary Table 4.

Meetings: Findings from this study were presented at the Union World Conference on Lung Health, Copenhagen, November 20, 2025.

Figure 1: TB model structure

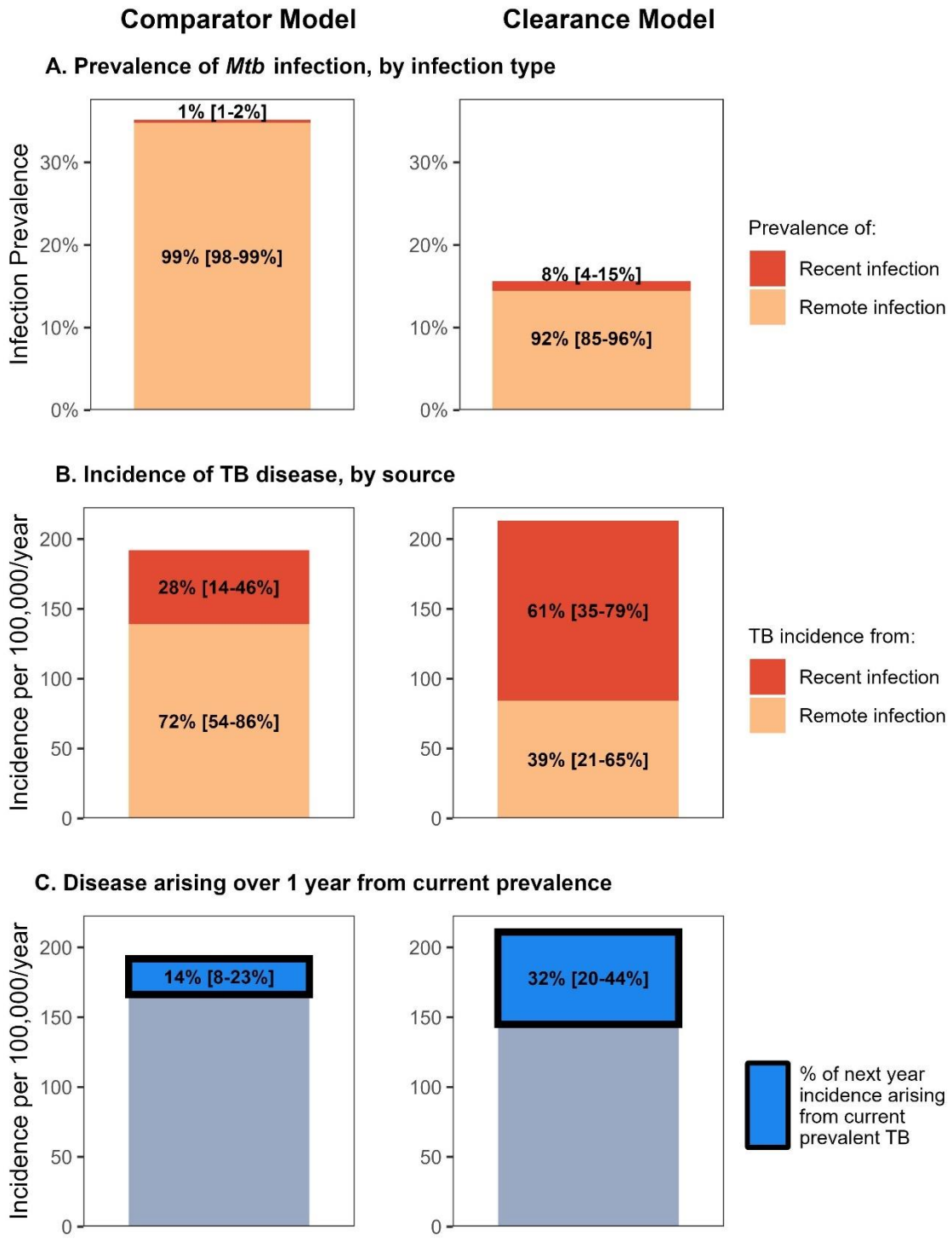


*Infection clearance (blue dashed arrows from Recent Infection and Remote Infection to Cleared) was included only in the Clearance model, not the Comparator model.

In this model diagram, boxes represent TB-related (panel A) and age-related (panel B) states, while arrows represent transitions between states. Intervention-related transitions (resulting from active case-finding or provision of TB preventive treatment) are shown via black dashed lines. All individuals enter the model uninfected and in the younger (< 15 years) age stratum. TB mortality occurs from the care-seeking TB state only, while non-TB mortality occurs from all states. Apart from the birth and non-TB mortality rates and intervention-related parameters, all parameters were calibrated (details in the text). Additional details are available in the Supplementary Methods and Supplementary Figure 1.

Alt text: Mathematical model diagram with two panels, in which boxes depict TB compartments (panel 1) and age compartments (panel 2) and arrows depict transitions between compartments.

Figure 2: Effect of considering immune clearance on modeled TB infection dynamics



This figure illustrates how inclusion of *Mtb* infection clearance affects modeled infection and disease dynamics. Panel A demonstrates that, without a mechanism for immune clearance ("Comparator Model"), people who were infected remotely account for a larger absolute population, and also a larger fraction of those with *Mtb* infection. Because calibration data shared by both models results in similar progression

rates after infection, a greater prevalence of recent *Mtb* infection results in the Clearance model estimating that a larger proportion of incidence arises from recent (vs. remote) infection than the Comparator model (Panel B). Since interventions to diagnose and treat people with current TB disease will primarily affect the burden of recent infection, the Clearance model estimates a larger impact from such interventions (Panel C). In panel C, the “disease arising over 1 year from current prevalence” represents the fraction of new incident disease in a given year that would be averted if all people with TB disease at the start of that year were instantaneously diagnosed and treated at year-start. All results show means across 20,000 model simulations, with corresponding 95% uncertainty intervals in parentheses.

Alt Text: Bar chart with 3 panels (organized horizontally) and 2 columns (first column shows results under the Comparator model without clearance; second column shows results under the Clearance model). The top panel displays the prevalence of *Mtb* infection, with shading indicating the proportion of that prevalence that is recent vs. remote infection. The middle panel displays TB incidence, with shading indicating the proportion of that incidence that is from progression of recent infection vs. reactivation of remote infection. The bottom panel again displays TB incidence, with the proportion of incidence that could be averted over the next year by finding and treating all people with current prevalent TB emphasized via bright blue shading and a thick black outline. Text labels show numerical estimates on each panel (means, with 95% uncertainty intervals in brackets).

Figure 3: Projected impact of active case-finding under models with and without infection clearance

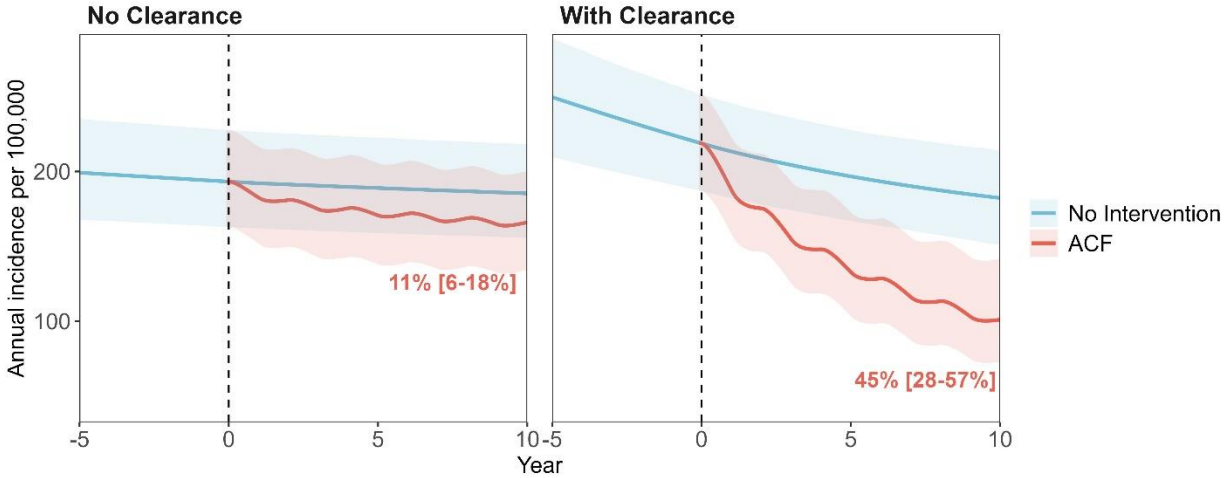


Figure shows the projected the mean annual TB incidence per 100,000 people (solid lines), with 2.5th and 97.5th percentiles (shaded ribbons), over ten years of simulated interventions (years 0 through 10), including no intervention (blue) and Active Case-Finding/ACF waves every 2 years (red). Despite overlapping uncertainty ribbons due to variation in baseline incidence across posterior parameter sets, all simulations projected a reduction in incidence from ACF (vs. no intervention). Text labels indicate the percent reduction in incidence at year 10, relative to no intervention (text shows means, with 95% uncertainty intervals in brackets). Panels differ by whether the model structure includes infection clearance (panel A = Comparator Model without clearance; panel B = Clearance Model). Incidence trends in the 5 years preceding the intervention period are also shown in each panel (years -5 through 0).

Alt text: Line graphs in 2 panels; the left panel shows results under the Comparator Model without clearance and the right panel shows results under the Clearance Model. Each panel shows incidence (vertical axis) over time (horizontal axis) with no additional interventions (blue line) or with active case-finding campaigns every 2 years (red line). Text labels on each panel indicate the reduction in incidence after 10 years from the active case-finding campaigns compared to no intervention (means with 95% uncertainty intervals in brackets).

Table 1: Empirical data used as model calibration targets

Target	Empirical Estimate [95% CI]	Comparator Model Posterior [95% UI]	Clearance Model Posterior [95% UI]	Sources for empirical estimates
Annual incidence per 100,00 (2021)	210 [178-244]	193 [163-228]	219 [187-251]	WHO Global TB Report [22]
Mortality to incidence ratio (2021)	0.17 [0.14-0.21]	0.17 [0.14-0.20]	0.16 [0.13-0.20]	
Annual % reduction in incidence (2002-21)	2.9% [1.9-3.9%]	0.6% [0.3-1.0%]	2.6% [1.6-3.6%]	
Prevalence per 100,000, ≥15 years (2019-21)	316 [290-342]	314 [285-341]	312 [284-342]	India prevalence survey [23]*
Prevalence-to-notification ratio, ≥15 years (2019-21)	2.84 [2.61-3.10]	2.8 [2.6-3.1]	2.8 [2.6-3.1]	
Percent of prevalence pre-care-seeking, ≥ 15 years (2019-21)	80% [73-88%]	78% [69-85%]	78% [71-84%]	
Immunoreactivity prevalence, ≥ 15 years (2019-21)	31% [20-44%]	48% [40-57%]	20% [14-29%]	
Cumulative progression (first 2 years after infection)	5.6% [4.3-7.1%]	5.9% [4.6-7.2%]	5.2% [4.0-6.5%]	Ferebee & Mount 1962; Sutherland 1968, reported in Menzies et al. 2018 [24–26]
Cumulative progression (years 3-10 since infection)	1.5% [0.6-2.8%]	2.8% [1.7-4.2%]	1.8% [0.8-3.3%]	

Likelihood distributions for each calibration target are available in Supplementary Table 2. Distributions for each posterior output are shown in Supplementary Figure 2.

*The prevalence-to-notification ratio was calculated from a combination of prevalence survey data and national notification data.