# Effectiveness of contact investigations for tuberculosis control in Arkansas

**Giorgio Guzzetta**[a,b], **Marco Ajelli**[a], **Zhenhua Yang**[c], **Leonard N. Mukasa**[d,e], **Naveen Patil**[d,e], **Joseph H. Bates**[d,e], **Denise E. Kirschner**[f], and **Stefano Merler**[a]

[a]Fondazione Bruno Kessler, via Sommarive 18, 38123, Povo (TN), Italy

[b]Trento RISE, via Sommarive 18, 38123, Povo (TN), Italy

[c]University of Michigan, 500 South State Street, Ann Arbor, MI 48109, USA

[d]Arkansas Department of Health, Slot 8, 4815 West Markham Street, Little Rock, AR 72205, USA

[e]University of Arkansas for Medical Sciences, 4301 West Markham Street, Little Rock, AR 72205, USA

[f]University of Michigan Medical School, 1500 East Medical Center Drive, Ann Arbor, MI 48109, USA

## Abstract

Comprehensive assessment of the effectiveness of contact investigations for tuberculosis (TB) control is still lacking. In this study, we use a computational model, calibrated against notification data from Arkansas during the period 2001–2011, that reproduces independent data on key features of TB transmission and epidemiology. The model estimates that the Arkansas contact investigations program has avoided 18.6% (12.1–25.9%) of TB cases and 23.7% (16.4–30.6%) of TB deaths that would have occurred during 2001–2014 if passive diagnosis alone were implemented. If contacts of sputum smear-negative cases had not been included in the program, the percentage reduction would have been remarkably lower. In addition, we predict that achieving national targets for performance indicators of contact investigation programs has strong potential to further reduce TB transmission and burden. However, contact investigations are expected to have limited effectiveness on avoiding reactivation cases of latent infections over the next 60 years.

## Keywords

## Introduction

In high-income countries, the incidence of tuberculosis (TB) has seen a sustained decline since 1993, thanks to general improved health conditions and to control interventions applied by public health authorities [1]. In the United States, the backbone of TB control is based on three core elements: case detection and management, active investigation of contacts of infectious patients and treatment of latent TB infection (LTBI) [2]. The cost-effectiveness of contact tracing in high-income countries has been evaluated in several epidemiological studies, yielding conflicting results, possibly due to limitations in study designs [1]. Randomized control trials would provide conclusive evidence of the effectiveness of contact investigation studies [3], but no such studies are available to date [4]. Mathematical models can supplement epidemiological studies and aid the design and assessment of alternative strategies, but modeling studies for TB contact tracing are also lacking [5].

In this work, we apply a computational model of TB transmission dynamics in Arkansas, USA, as an example of a low-burden, high-income setting, to assess the effectiveness of current contact investigation protocols and the potential impact of alternative intervention scenarios.

## Methods

The epidemiological model used in this study is based on a previously published one, featuring a realistic, spatially explicit and time-evolving representation of the Arkansas socio-demographic structure by means of an individual based approach [6]. TB transmission within households, schools, workplaces, and casual, distance-dependent contacts is considered. Schools and workplaces are enclosed settings where individuals spend much of their day interacting with a significant number of contacts [7, 8] thus representing transmission network hubs [9]. Indeed, TB outbreaks in schools are relatively common (e.g. [10]), and both schools and workplaces are important settings for screening potential contacts of index cases [9]. The population in the model is spatially distributed according to estimates of the population density [11] over 219 square cells covering the state territory. The demography and the network of contacts among individuals are updated at the end of each year: individuals are generated, grow older, create new households, procreate and die; the processes of school enrolment (following the educational career), new employment, job loss and retirement are also modeled. The evolution of the socio-demographic structure was validated in the original paper [6]. The model captures immigration via a yearly influx of a foreign-born population based on available age-specific data [12]. Foreign-born individuals are characterized by a higher rate of LTBI prevalence [13], and the model can accommodate a small fraction of new immigrants being infectious at the time of entry in the US. Since we do not explicitly model targeted TB screening during the immigration process, this fraction of infectious immigrants accounts for infectious individuals who may avoid the screening process (e.g. illegal immigrants), or with LTBI identified at screening who do not complete treatment and progress rapidly to TB disease.

Following the original model [6], whose structure is proposed in Figure 1, individuals are born uninfected and move to the recently infected class upon contact with an infectious (Active TB) individual. Recent infections are asymptomatic and progress to either of the following epidemiological outcomes: i) complete healing, with pathogen clearance via immune response; ii) active TB disease within a few years from infection episode [14] ("primary TB"); iii) LTBI. Individuals with LTBI are asymptomatic and non-infectious, and may develop endogenous TB reactivation [14], even several decades after the infection episode. Individuals with LTBI can be re-infected by contacts with infectious hosts. Re-exposed individuals may develop active TB ("exogenous reinfection") within a few years, or revert to the latent class, just like recently infected individuals. However, the probability of exogenous reinfection is lower than that of primary TB, to account for the probability of protection from TB-specific immune memory arising after first exposure to TB [6, 14, 15]. The model considers extrapulmonary (non-infectious) and pulmonary (infectious) TB and assigns a higher infectiousness to smear-positive cases compared to smear-negative [16, 17]. Irrespectively of the site of TB and smear status, individuals with active TB may be diagnosed and cured and are at an increased risk of death.

The original model [6] was expanded to include contact investigation activities as per guidelines of the Centers for Disease Control (CDC) [18]. Full details on implementation are given in the Technical Appendix. A summarizing graphical representation of the contact investigation procedures is provided in Figure 2. More specifically, Figure 2A represents a case management procedure, which includes a decision model for initiation of contact investigation and, where appropriate, elicitation of a list of contacts from the index case. We note here that the CDC does not prioritize investigation of contacts of smear-negative cases (although it is recommended when resources are available) [18]. In Arkansas, the Department of Health applies the same protocol to smear-positive and smear-negative patients with equal priority [19]. Figure 2B depicts our procedure of tracing, testing and, where applicable, treating contacts on the list. Completed treatment of LTBI is assumed to have 100% and instantaneous efficacy in controlling the reactivation of LTBI. The contact investigation program applied in Arkansas (termed "Arkansas contact tracing" in the rest of the paper), was modeled by using yearly state-specific performance estimates from the Aggregate Reports on TB Program Evaluation [19] on the program coverage (proportion of index cases for which contacts are elicited), proportion of missed contacts, rates of treatment initiation and completion against LTBI.

The model is calibrated using surveillance data [20] and validated against data from four independent sources: a molecular epidemiology study in Arkansas [21], surveillance data on TB in the foreign-born [20] and nation-wide epidemiological studies on contact tracing in households [22] and workplaces [9]. The 9 free model parameters are calibrated by a procedure minimizing the error between data and model output. Latin Hypercube Sampling [23] was used to explore the parameter space, with three rounds of nested sampling with size 10,000 at each round. Full details on model implementation and calibration, including references for the range of exploration of parameter values and best estimates for free model parameters, are reported in the Technical Appendix. Model outputs were calculated for the 100 parameter sets yielding the lowest error during calibration and each parameter set was run 100 times to account for the intrinsic stochasticity of the modeled processes. All results

are shown as the average (with 95 percentile intervals) over the $100 \times 100 = 10,000$ simulations. The sensitivity of results with respect to each individual parameter was evaluated through a multivariate analysis (Technical Appendix).

Simulation results under the Arkansas contact investigation are compared to several hypothetical programs where different choices for control are applied. In particular, we considered the following programs:

- "passive diagnosis": TB cases are only found by passive diagnosis, i.e. no contact investigation is in place; this program serves as a reference for evaluating the overall effectiveness of the Arkansas contact investigation program;

- "no LTBI treatment": contact investigation is performed and active TB cases diagnosed within the program are treated; however, there is no treatment of LTBI; this program allows the estimation of the relative contribution to TB control of prevented LTBI reactivations;

- "2015 targets": we estimate what would have happened if targets for performance indicators set by the CDC for 2015 [24] had been achieved in 2001 and maintained throughout the study period, to show the potential impact of improving the performance of contact investigation; these targets include an increased coverage, a reduced proportion of missed contacts and higher rates of treatment initiation and completion against LTBI identified through contact tracing (see Technical Appendix);

- "smear+ only": contact tracing of smear-negative individuals is not done.

## Results

### Quality of model calibration

Figure 3 shows the quality of model calibration relative to TB incidence data over time [20] (Figure 3A, years 2001–2011, blue bars) and age-group specific TB incidence (Figure 3B). The model reproduces the observed steady decline in TB incidence from about 6 cases per 100,000 in 2001 to about 3 per 100,000 in 2011. This decline results from a combination of ongoing control strategies, which keep transmission at low levels, and cohort effects. Older cohorts are characterized by higher LTBI prevalence [13], due to higher TB transmission in past epochs: as they die naturally, the overall LTBI prevalence declines in the population and consequently also endogenous reactivations and secondary cases recently transmitted by reactivated cases. This effect was previously observed in a US-based epidemiological study [25]. Model predictions also follow closely the profile of TB incidence by age [20] (Figure 3B), particularly the progressive increase at older ages (with 12 cases per 100,000 individuals over 85 years, compared to about 4.5 in the 45–64 age class) that is typical of high income, low burden countries. This increase is mainly due to endogenous reactivation cases, as both LTBI prevalence and the risk of endogenous reactivation are higher at older ages [13, 26].

## Model validation

Before the model was applied to ask questions regarding contact tracing, we ensured it reproduces key features of TB dynamics in Arkansas by comparing its predictions with several epidemiological measures from independent studies.

Figures 4A and 4B compare model outputs with results from a recent molecular epidemiology study in which the large majority (87%) of culture-confirmed isolates from TB cases in Arkansas during the period 2004–2010 were genotyped (spoligotype and 12-locus MIRU-VNTR) to identify transmission clusters [21]. In that study, a cluster is defined as at least two cases diagnosed within a year from each other, caused by *Mycobacterium tuberculosis* isolates sharing the same molecular typing, independently of known epidemiological links [21]. The Arkansas molecular epidemiological study found that 61% of cases were non-clustered, a figure well below the national average of 77% [27] because of the time-restricted cluster definition [21]. The fraction of clustered cases is often considered as a proxy of recently transmitted cases (e.g. [16]); however, the accuracy of this interpretation may depend on the adopted definition of cluster. For this reason, we do not compare directly model estimates on the fraction of recently transmitted cases with results of the study. Instead, we reproduce criteria for cluster definition of the molecular epidemiology study within the model and compare model-predicted clusters to epidemiological observations in [21]. Figure 4A shows the percentage of cases identified as "non-clustered" for the entire population and disaggregated by age group. The model predicts there exists a proportion of non-clustered cases of 58.2% (56.3–60.2%) in the general population, and a clear heterogeneity by age groups, with 44.3% (38.7–50.0%) in 20–64 years old versus 63.4% (55.2–70.6%) in 65+ years old. While the latter two figures underestimate observations from the molecular epidemiology study, the model still recapitulates the observed significant difference between the two age groups [21]. Of note, the "all ages" category includes cases from the age class 0–19 years, which we do not show since the molecular epidemiology study only focused on age classes above 20 years [21]. Figure 4B shows the relative proportion of clusters of different sizes. About 79.0% (70.9–86.3%) of clusters in the simulation have less than 4 cases, 16.6% (9.7–23.9%) are intermediate size clusters with 5 to 9 cases, and 4.4% (0.9–8.8%) are larger than 9. Figure 4C compares the predicted and observed proportion of TB cases in foreign-born individuals over time. The model predicts an increase in this proportion from 7.5% (4.7–10.5%) in 2001 to 21.0% (16.6–29.9%) in 2011, reflecting the observed pattern of growth of TB incidence in the foreign-born during the last decade. The general slight underestimation by the model may be attributable to specific demographic [28], sociological [29] and biological risk factors [30] for increased TB transmission and reactivation in the foreign-born, which are not included in the model. Figure 4D shows total number of cases from foreign-born patients disaggregated by time of sojourn to the United States. The model predicts that about 25.5% (15.1–35.2%) of cases occurred within 1 year of arrival, 18.8% (9.7–30.7%) after 1 to 4 years and 56.8% (40.5–78.9%) among long-term immigrants (5 years or more). Additionally, 86.4% (81.4–91.0%) of TB cases in foreign-born are predicted to be non-clustered, consistently with the 85% figure reported in [27]. Figures 4E and 4F show model predictions for two further quantities that are particularly important to contact investigation programs, i.e. the prevalence of secondary TB cases (termed "yield" in contact tracing studies) and the

prevalence of LTBI in contacts of smear-positive TB cases, disaggregated by transmission setting (household or workplace) [22, 9]. For household contacts, the estimated yield is about 2.42% (0.95–4.21%) and the estimated LTBI prevalence is about 33.0 (25.3–40.8%). Estimates for the corresponding quantities in workplace contacts are 0.45% (0.21–0.71%) and 17.1% (12.5–22.2%) respectively.

The model also captures the geographic distribution of cases, identifying 5 of the 6 Arkansas counties with the highest number of TB cases ( 8) between 2011 and 2013 (Benton, Craighead, Desha, Pulaski, Sebastian and Washington - L.N. Mukasa, unpub. data). The only county not identified in this list was Desha, which has witnessed an unusually high TB incidence due to a recent cluster of cases (L.N. Mukasa, pers. comm.). A Pearson correlation coefficient between cases observed and predicted by county of 0.760 (0.644–0.842) and a non-significant mean difference (paired t-test p-value: 0.86). Finally, the model estimated an average 70 (95% percentile interval: 51–91) deaths due to TB in 2001–2010, closely reflecting the 71 deaths registered in Arkansas in the same period [31].

Overall, the model reproduces existing TB transmission patterns, and is therefore suitable for assessing contact investigation programs.

### Effectiveness of contact investigation

A key question the model can be used to explore is the impact of implementing alternative contact investigation programs on the burden of TB disease. Figure 5A and B show respectively the number of TB cases and deaths due to TB in the different considered programs. Percentage difference between each program and the passive diagnosis program are also reported. According to the model, the program implemented in Arkansas during 2001–2014 avoided about 372 cases (173–559), and 35 deaths (20–53), corresponding to about 18.6% and 23.7 % of the total that would have occurred with passive diagnosis only. This result was afforded mainly by the reduced transmission and mortality caused by the early identification of TB cases in contacts, and marginally by avoiding progression to disease in individuals with LTBI or recent infection. More specifically, treatment of LTBI contacts who tested positive for protein-purified derivative (PPD) avoided 2.9% of cases and 2% of deaths, as can be seen by comparing estimates for the Arkansas program versus the one with no LTBI treatment. The figure also shows that reaching the performance targets set for 2015 has strong potential to further reduce the burden of TB, and this stems from the proportion of TB transmissions that are still overlooked by contact investigation programs. If we consider TB cases from recent transmissions (i.e., those not derived from endogenous reactivation of an LTBI case), the model predicts that only about 17.5% (14.1 – 19.9%) of these are identified through contact investigation. This can be explained by considering that, when a recently transmitted TB case is missed by an investigation, it has the potential to infect many other individuals before being passively diagnosed. Therefore, even small improvements in performance indicators (e.g., the percentage of contacts traced), will result in remarkable improvements in effectiveness. For example, in the 2015 targets program, the proportion of TB cases from recent transmission who are found through contact tracing rises to 21.3% (18.0 – 24.4%). Even more important, Figure 5 highlights model predictions regarding the effect of tracing contacts of smear-negative TB cases: considering smear-

positive cases only has a strong negative impact on the effectiveness of the program, reducing the percentages of potentially avoided cases and deaths to 10.4 and 13.2%, respectively.

### LTBI treatment, prevalence and reactivation

Contact investigations reduce LTBI prevalence directly, through treatment of LTBIs, and indirectly, through reduced TB transmission. In this section, we use the model to assess estimates on the amount of this reduction and its effect on future reactivation cases. Figure 6A shows the LTBI prevalence in 2011 for the different programs, and the reduction in percentage with respect to passive diagnosis. These reductions are relatively small, ranging from 640 less LTBIs for the "smear+ only" program to 2,450 in the "2015 targets" program. The predicted number of endogenously reactivated TB cases up to year 2061 is shown in Figure 6B. Even in the absence of LTBI treatment, about 1.5% of reactivated cases are avoided, due to reduced TB transmissions afforded by case identification and treatment. With the addition of LTBI treatment (as in the Arkansas contact tracing program), this proportion increases to 2.8%.

## Discussion

TB continues to be a worldwide health burden and great steps are needed to understand many aspects of disease dynamics and control at both the individual and population scales. At the individual scale, reducing time to diagnosis through targeted interventions is a critical factor for control of TB transmission and decisions on how and whom to treat greatly impact the population trends of disease. At the population level, one intervention in this direction has been the investigation of contacts with each index case of TB identified. Herein we use a published individual based model of TB in Arkansas [6] to examine current policies and predict the impact of alternative strategies. The model was designed to reproduce heterogeneities in contact patterns due to the population's socio-demographic structure.

Although these heterogeneities complicate the structure of the model, they are a necessary component to realistically capture the dynamics and impact of contact investigation [5]. A similar model structure has been recently employed to evaluate contact tracing protocols to contain the accidental release of potential pandemic influenza viruses from biosafety laboratories [32]. Calibration of complex models is difficult and potentially prone to problems in parameter estimation, although using appropriate techniques such as LHS sampling may mitigate this issue [23]. Our calibrated model reproduces key features of TB in Arkansas, including the balance between recent transmission and reactivation [21], characteristics of TB cases in foreign-born TB [20], the amount of TB transmission in households and workplaces [22, 9] and the geographical distribution of TB cases. Model predictions were robust across optimal parameter sets, and a multivariate sensitivity analysis [23] showed little impact on predictions from uncertainties in parameter estimations (see Technical Appendix).

Model results suggest that the current contact tracing program implemented in Arkansas significantly reduced TB incidence and deaths. A key component of the program's success in reducing the burden of TB was the systematic investigation of contacts of smear-negative

TB cases, which almost doubled program effectiveness when compared to one where only smear-positive cases are subject to contact investigation. Despite a lower infectiousness [16, 17], the share of transmission directly attributable to smear-negative cases is estimated to be between 17% and 41% in two studies based in San Francisco and Vancouver respectively [16, 33]. In addition, the inclusion of smear-negative cases allows rapid identification of smear-positive secondary cases among contacts, which otherwise could have caused larger clusters of transmission indirectly due to the original smear-negative case.

According to our predictions, only about one fifth of recently transmitted TB cases are identified by contact investigation. The high fraction of missed transmissions is a multiplicative result of imperfections in each sequential step of the contact investigation process (Figure 2): failure to elicit contact investigation for some of the index cases, incomplete identification of contacts by the patient, incomplete tracing and screening of identified contacts, and missed diagnoses. Thus, our model suggests that improvements in each of these areas can have a large and synergistic impact on reducing disease burden.

Model results also predict a limited potential for the contact investigation program towards reducing TB incidence from endogenous reactivation in the coming decades. Demographic factors are known to be major drivers of the decline of LTBI prevalence in the past decades [25]: our simulations show that the additional relative contribution of contact investigation has a residual impact, even under scenarios where initiation and completion of LTBI treatments is significantly improved, such as that envisioned by the 2015 CDC targets. In this regard, since July 2012, the large majority of LTBI cases identified in Arkansas are offered a 3-month regimen based on 12 doses of isoniazid-rifapentine instead of the standard 9-month isoniazid course. With the shorter course, treatment completion rates have increased to 83% (unpublished data). In addition, the deployment of the more specific Interferon-Gamma Release Assay substituting for the tuberculin skin test for the detection of LTBI should allow a better focus of resources by avoiding false-positive results [34].

The present study does not explicitly model specific TB transmission patterns and risks in foreign-born individuals, and this may have caused a slight underestimation of the proportion of TB in the foreign-born. Socio-demographic heterogeneities (e.g. household sizes and age composition, school attendance and employment rates by age) [28] can alter contact patterns within foreign-born subpopulations and with the resident population. The existence of genetic and socio-economic risk factors [28–30] for TB transmission and reactivation specific to the foreign-born subpopulation complicate the issue, as they expand the number of unknown parameters that the model would need to estimate. These aspects should be explicitly included when evaluating the impact of control interventions specifically targeted to these subpopulations, such as LTBI screening of immigrants [35], to correctly account for the indirect advantages of prevented transmission. In this study, we analyzed a context with low immigration [12] and a limited contribution of TB cases from foreign born (<20% in the considered period) [20]. Thus, these issues are not expected to affect our conclusions significantly, especially because recently transmitted TB cases are a minority in the foreign-born. However, we caution against a straightforward extension of these results to settings where heterogeneities of the foreign-born population and interventions targeted to this risk group play a larger role in the local TB epidemiology.

Other socio-demographic groups also exist that are at an increased risk of TB transmission and reactivation [36], including those associated with correctional facilities [37], hospitals, long-term care facilities [38], shelters for homeless persons, and alcohol or drug consumption sites. Outbreaks of TB transmission have been confirmed within all of these settings, for which specific investigation protocols may apply [2]. Currently, we coarsely include high-risk transmission sites within the random component of transmission, due to the lack of data that are needed to fine grain these details into our model. Thus, additional effort that is occasionally applied during outbreak investigations in these settings [2] is also not explicitly considered. However, the total contribution of TB cases from these settings is only around 10% of the total TB occurring in Arkansas in the considered period [20]; therefore, model recommendations should still hold true for the general population. Targeted testing of populations at high-risk, a prominent control strategy in Arkansas, has great potential to contribute to the TB elimination goal [36, 39]. However, evaluation of its effectiveness goes beyond the scope of this study, due to specific limitations of the socio-demographic model. Finally, the model also does not include HIV co-infection, as this currently is not epidemiologically relevant in Arkansas (TB cases in HIV positive individuals comprise only 3 to 5% of all cases 2001–2011 [20]).

Several recent reviews have highlighted the lack of studies evaluating the effectiveness of contact investigation programs [1, 3, 4], also with specific reference to mathematical models [5]. A recent study [40] used an agent-based model to provide useful predictions on the impact of household contact investigations in a non-specific moderate-burden setting. Here, we attempt to provide quantitative estimates of contact investigation activities in an actual setting characterized by low TB burden, which may be useful to the debate on contact investigation protocols. It is however important to remark that the low level of foreign-born TB and HIV coinfection in Arkansas may limit the generalizability of our conclusions to other settings. On one hand, the increased amount of reactivated TB in foreign-born individuals should reduce the relative contribution of recently transmitted cases, which are the main target of contact investigation programs. On the other hand, the increased risk of developing TB in HIV-positive individuals is expected to at least partially counterbalance this effect in settings with higher rates of HIV/TB coinfection. Further studies are needed to quantitatively elucidate the combined effect of these two factors on the effectiveness of contact tracing. Similar studies of this kind can be of particular importance for geographic settings with high TB burden and low resources [5], in order to support efficient resource allocation. The conclusions of this study cannot be extended straightforwardly to make predictions on the impact of contact investigation in high-burden countries. Adaptation of the model to local socio-demographic and epidemiological data will be needed for this purpose, and specific local issues need to be considered, including high HIV burden [41], emergence of TB drug resistance [42] and the effect of strategies for actively finding TB cases [43].

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

## References

1. Verdier JE, de Vlas SJ, Baltussen R, Richardus JH. A systematic review of economic evaluation studies of tuberculosis control in high-income countries. Int J Tuberc Lung Dis. 2011; 15(12):1587–1597. [PubMed: 21740647]

2. Coordinating Center for Health Information and Service, Centers for Disease Control and Prevention (CDC). Morbidity and Mortality Weekly Report. Vol. 54. Atlanta, GA: U.S. Department of Health and Human Services; 2005. Controlling Tuberculosis in the United States.

3. Fox, GJ.; Dobler, CC.; Marks, GB. The Cochrane Library. 2011. Active case finding in contacts of people with tuberculosis.

4. Fox GJ, Barry SE, Britton WJ, Marks GB. Contact investigation for tuberculosis: a systematic review and meta-analysis. Eur Respir J. 2013; 41:140–156. [PubMed: 22936710]

5. Begun M, Newall AT, Marks GB, Wood JG. Contact Tracing of Tuberculosis: A Systematic Review of Transmission Modelling Studies. PLoS One. 2013; 8(9):e72470. [PubMed: 24023742]

6. Guzzetta G, Ajelli M, Yang Z, Merler S, Kirschner DE, Furlanello C. Modeling socio-demography to capture tuberculosis transmission dynamics in a low burden setting. J Theor Biol. 2011; 289:197–205. [PubMed: 21906603]

7. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. PLoS Med. 2008; 5(3):e74. [PubMed: 18366252]

8. Fumanelli L, Ajelli M, Manfredi P, Vespignani A, Merler S. Inferring the Structure of Social Contacts from Demographic Data in the Analysis of Infectious Diseases Spread. PLoS Comput Biol. 2012; 8(9):e1002673. [PubMed: 23028275]

9. Davidow AL, Mangura BT, Wolman MS, Bur S, Reves R, Thompson V, et al. Workplace contact investigation in the United States. Int J Tuberc Lung Dis. 2003; 7:S446–S452. [PubMed: 14677836]

10. Faccini M, Codecasa LR, Ciconali G, Cammarata S, Borriello CR, De Gioia C, et al. Tuberculosis outbreak in a primary school, Milan, Italy. Emerg Infect Dis. 2013; 19(3):485–7. [PubMed: 23621942]

11. Balk, D.; Yetman, G. The global distribution of population: evaluating the gains in resolution refinement. Palisades, NY: Center for International Earth Science Information Network; 2004.

12. Capps, R.; McCabe, K.; Fix, M.; Huang, Y. A Profile of Immigrants in Arkansas: Changing Workforce and Family Demographics. Vol. 1. Little Rock, AR, and Washington, DC: Winthrop Rockefeller Foundation and Migration Policy Institute; 2013.

13. Bennett DE, Courval JM, Onorato I, Agerton T, Gibson JD, Lambert L, et al. Prevalence of tuberculosis infection in the United States population. Am J Respir Crit Care Med. 2008; 177:348–355. [PubMed: 17989346]

14. Vynnycky E, Fine PEM. The natural history of tuberculosis: the implications of age- dependent risks of disease and the role of reinfection. Epidemiol Infect. 1997; 119:183–201. [PubMed: 9363017]

15. Guzzetta G, Kirschner D. The roles of immune memory and aging in protective immunity and endogenous reactivation of tuberculosis. PLoS One. 2013; 8(4):e60425. [PubMed: 23580062]

16. Behr MA, Warren SA, Salamon H, Hopewell PC, Ponce de Leon A, Daley CL, et al. Transmission of Mycobacterium tuberculosis from patients smear-negative for acid-fast bacilli. Lancet. 1999; 353(9151):444–449. [PubMed: 9989714]

17. Tostmann A, Kik SV, Kalisvaart NA, Sebek MM, Verver S, Boeree MJ, et al. Tuberculosis Transmission by Patients with Smear-Negative Pulmonary Tuberculosis in a Large Cohort in The Netherlands. Clin Infect Dis. 2008; 47(9):1135–1142. [PubMed: 18823268]

18. Coordinating Center for Health Information and Service, Centers for Disease Control and Prevention (CDC). Morbidity and Mortality Weekly Report. Vol. 54. Atlanta, GA: U.S.

Department of Health and Human Services; 2005. Guidelines for the Investigation of Contacts of Persons with Infectious Tuberculosis.

19. ARPE – Aggregate Reports for Program Evaluations. http://www.healthindicators.gov/Resources/DataSources/Aggregate-Reports-for-Tuberculosis-Program-Evaluation_220/Profile

20. Online Tuberculosis Information System (OTIS), National Tuberculosis Surveillance System, United States. US Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), Division of TB Elimination, CDC WONDER Online Database. Apr. 2014 http://wonder.cdc.gov/tb-v2012.html

21. Berzkalns A, Bates J, Ye W, Mukasa L, France AM, Patil N, et al. The road to tuberculosis (Mycobacterium tuberculosis) elimination in Arkansas; a re-examination of risk groups. PLoS One. 2014; 9(3):e90664. [PubMed: 24618839]

22. Moran-Mendoza O, Marion SA, Elwood K, Patrick DM, FitzGerald JM. Tuberculin skin test size and risk of tuberculosis development: a large population-based study in contacts. Int J Tuberc Lung Dis. 2007; 11:1014–1020. [PubMed: 17705981]

23. Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol. 2008; 254:178–196. [PubMed: 18572196]

24. Coordinating Center for Health Information and Service, Centers for Disease Control and Prevention (CDC). Morbidity and Mortality Weekly Report. Vol. 59. Atlanta, GA: U.S. Department of Health and Human Services; 2010. Monitoring Tuberculosis Programs — National Tuberculosis Indicator Project, United States, 2002–2008.

25. Winston CA, Navin TR. Birth cohort effect on latent tuberculosis infection prevalence, United States. BMC Infect Dis. 2010; 10:206. [PubMed: 20626871]

26. Horsburgh CR, O'Donnell M, Chamblee S, Moreland JL, Johnson J, Marsh BJ, et al. Revisiting rates of reactivation tuberculosis. Am J Respir Crit Care Med. 2010; 182(420):425.

27. Ricks PM, Cain KP, Oeltmann JE, Kammerer JS, Moonan PK. Estimating the Burden of Tuberculosis among Foreign-Born Persons Acquired Prior to Entering the U.S., 2005–2009. PLoS One. 2011; 6(11):e27405. [PubMed: 22140439]

28. Cantwell MF, McKenna MT, McCray E, Onorato IM. Tuberculosis and race/ethnicity in the United States: impact of socioeconomic status. Am J Respir Crit Care Med. 1998; 157(4):1016–1020. [PubMed: 9563713]

29. Stead WW, Senner JW, Reddick WT, Lofgren JP. Racial differences in susceptibility to infection by Mycobacterium tuberculosis. N Engl J Med. 1990; 322(7):422–427. [PubMed: 2300105]

30. Velez DR, Wejse C, Stryjewski ME, Abbate E, Hulme WF, Myers JL, et al. Variants in toll-like receptors 2 and 9 influence susceptibility to pulmonary tuberculosis in Caucasians, African-Americans, and West Africans. Hum Genet. 2010; 127(1):65–73. [PubMed: 19771452]

31. CDC Wonder – Underlying Cause of Death, Detailed Mortality. http://wonder.cdc.gov/

32. Merler S, Ajelli M, Fumanelli L, Vespignani A. Containing the accidental laboratory escape of potential pandemic influenza viruses. BMC Medicine. 2013; 11:252. [PubMed: 24283203]

33. Hernández-Garduño E, Cook V, Kunimoto D, Elwood RK, Black WA, FitzGerald JM. Transmission of tuberculosis from smear negative patients: a molecular epidemiology study. Thorax. 2004; 59:286–290. [PubMed: 15047946]

34. Coordinating Center for Health Information and Service, Centers for Disease Control and Prevention (CDC). Morbidity and Mortality Weekly Report. Vol. 59. Atlanta, GA: U.S. Department of Health and Human Services; 2010. Updated Guidelines for Using Interferon Gamma Release Assays to Detect Mycobacterium tuberculosis Infection — United States.

35. Pareek M, Baussano I, Abubakar I, Dye C, Lalvani A. Evaluation of immigrant tuberculosis screening in industrialized countries. Emerg Infect Dis. 2012; 18(9):1422–9. [PubMed: 22931959]

36. Abubakar I, Stagg HR, Cohen T, Mangtani P, Rodrigues LC, Pimpin L, et al. Controversies and unresolved issues in tuberculosis prevention and control: a low-burden-country perspective. J Infect Dis. 2012; 205(Suppl 2):S293–300. [PubMed: 22448025]

37. Ijaz K, Yang Z, Templeton G, Stead WW, Bates JH, Cave MD. Persistence of a strain of Mycobacterium tuberculosis in a prison system. Int J Tuberc Lung Dis. 2004; 8(8):994–1000. [PubMed: 15305483]

38. Ijaz K, Dillaha JA, Yang Z, Cave MD, Bates JH. Unrecognized tuberculosis in a nursing home causing death with spread of tuberculosis to the community. J Am Geriatr Soc. 2002; 50(7):1213–8. [PubMed: 12133015]

39. Mukasa, L.; Bates, J.; Abernathy, R.; Phillips, J.; Karpoff, E.; Bakhtawar, I. TB Notes. Vol. 4. Atlanta, GA: U.S. Division of Tuberculosis Elimination - National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention; 2009. Addressing the Challenges of Missed Opportunities for Finding TB in Arkansas, 2008; p. 5-6.

40. Kasaie P, Andrews JR, Kelton WD, Dowdy DW. Timing of tuberculosis transmission and the impact of household contact tracing. An agent-based simulation model. Am J Respir Crit Care Med. 2014; 189(7):845–52. [PubMed: 24559425]

41. Houben RMGJ, Dowdy DW, Vassall A, Cohen T, Nicol MP, Granich RM, et al. How can mathematical models advance tuberculosis control in high HIV prevalence settings? Int J Tuberc Lung Dis. 2014; 18(5):509–514. [PubMed: 24903784]

42. Jenkins HE, Plesca V, Ciobanu A, Crudu V, Galusca I, Soltan V, et al. Assessing spatial heterogeneity of multidrug-resistant tuberculosis in a high-burden country. Eur Respir J. 2013; 42(5):1291–301. [PubMed: 23100496]

43. Yaesoubi R, Cohen T. Identifying dynamic tuberculosis case-finding policies for HIV/TB coepidemics. Proc Natl Acad Sci USA. 2013; 110(23):9457–62. [PubMed: 23690585]

## Highlights

- We evaluate the effectiveness of TB contact investigations in Arkansas in 2001–2014

- We estimate that contact investigations have avoided about 20% of TB cases and deaths

- Treatment of latent TB infections contributed marginally (2–3%) to this purpose

- We show the importance of tracing contacts of sputum smear-negative index patients

- Achieving national performance targets can significantly improve TB control
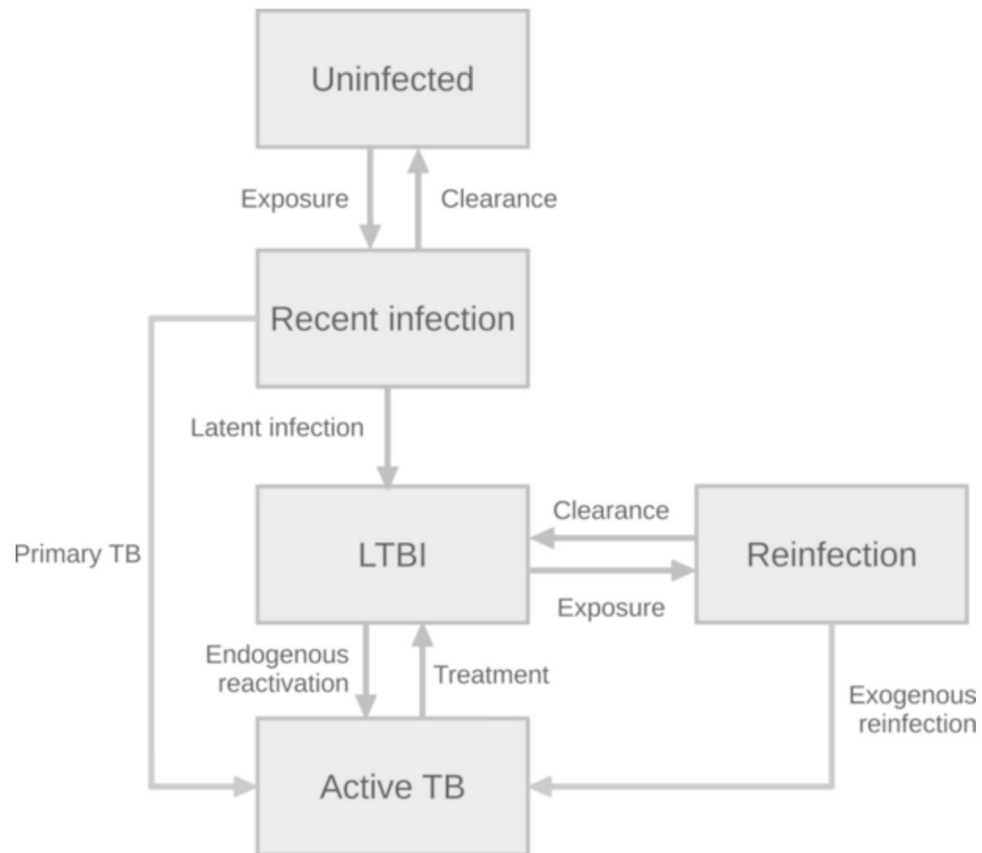
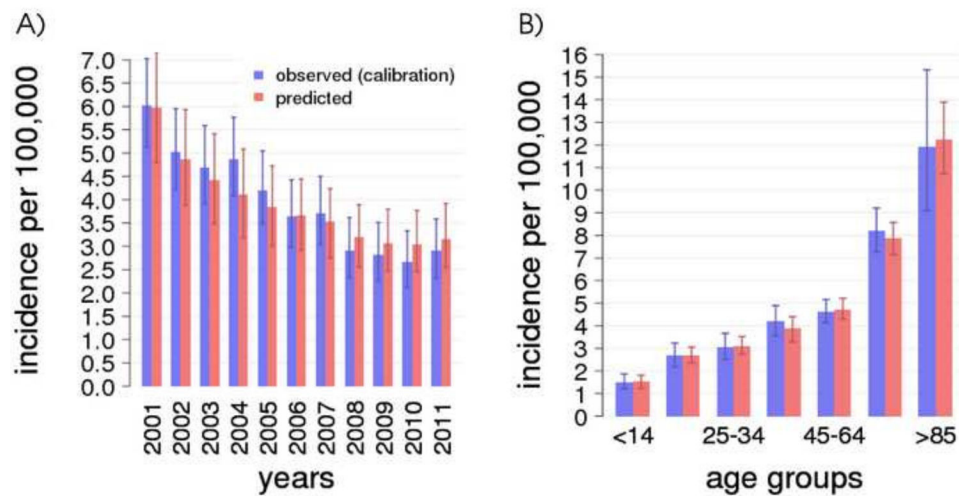

**Figure 1.**
Epidemiological workflow of the individual-based model. Individuals are born susceptible (Uninfected), may be exposed to TB infection (Recent infection) and either heal, develop primary TB or latent TB infection (LTBI). Latently infected individuals can develop endogenous reactivation or be re-exposed (Reinfected). Re-exposed individual can revert to the latent compartment or develop TB from exogenous reactivation. Individuals with TB (Active TB) can be diagnosed and treated, reverting to the latent compartment.
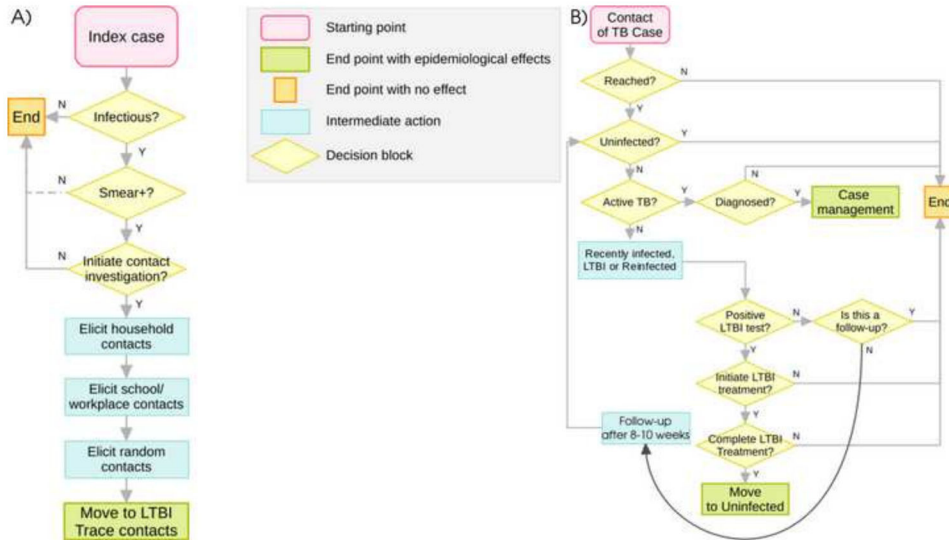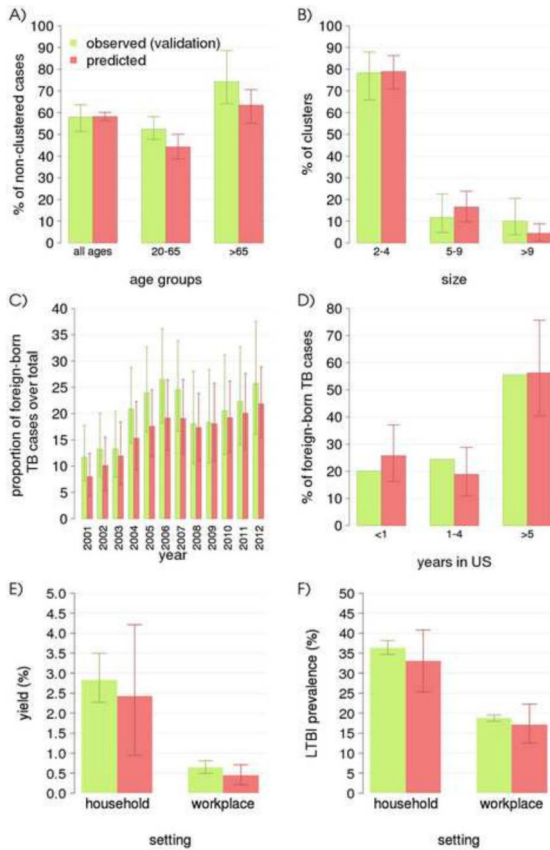
**Figure 2.**
Contact investigation program as implemented in the model; A) case management procedure. "Infectious" refers to pulmonary active TB; depending on the contact investigation program considered, a negative smear status does not necessarily imply a conclusion of the case management procedure: for this reason the arrow corresponding to the N option in the block labeled "Smear+?" is represented with a dotted line. B) Contact tracing procedure. For a fraction of contacts in the list, the procedure ends with no further action, representing contacts which could not be reached. If the contact is actually traced, action is taken according to its epidemiological status. For uninfected individuals, the algorithm will end with no further action. In a case of active disease, individuals may be correctly diagnosed and cured (in this case, cure is assumed instantaneous) and a new case management procedure for elicitation of its contacts is started. If the contact was recently or remotely (LTBI) infected, a protein-purified derivative (PPD) test is simulated. If the test is positive the contact can initiate and complete LTBI treatment with given probabilities. The contact is assumed to heal completely and moves to the susceptible compartment only if LTBI treatment is completed. If the PPD result is negative, a repetition of the test is performed in 8–10 weeks, during which time the individual's epidemiological status may change. If a second PPD test is negative, the procedure ends with no further action. All details are provided in the Technical Appendix.
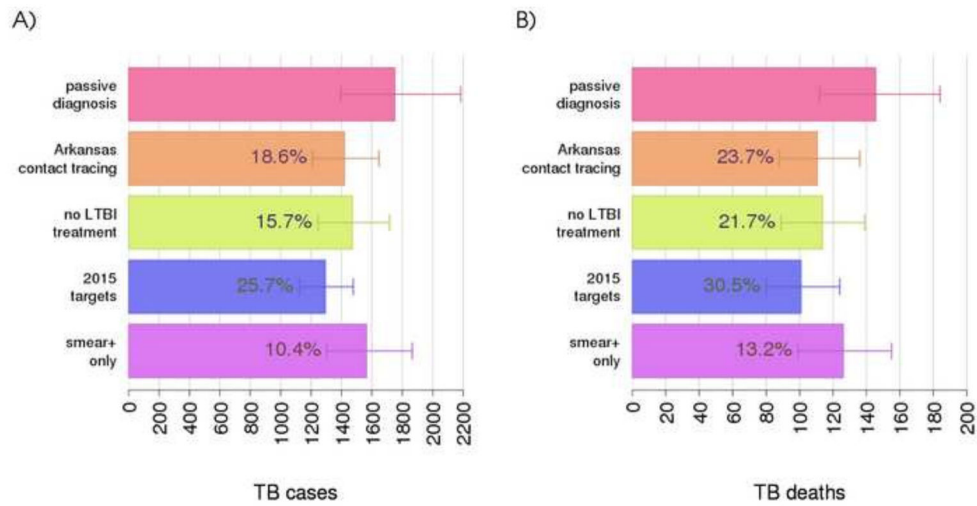
**Figure 3.**
A) TB incidence over time (2001–2011) in Arkansas comparing model and data. B) TB incidence by age groups in Arkansas, average 2001–2011. For both panels, error bars for the data are calculated as the 95-percentile confidence interval of a binomial distribution with size equal to the age-group specific Arkansas population at each time point and probability equal to the observed incidence. Average and 95-percentile error bars for the model are calculated over 100 best parameter sets and 100 stochastic simulations of the model.
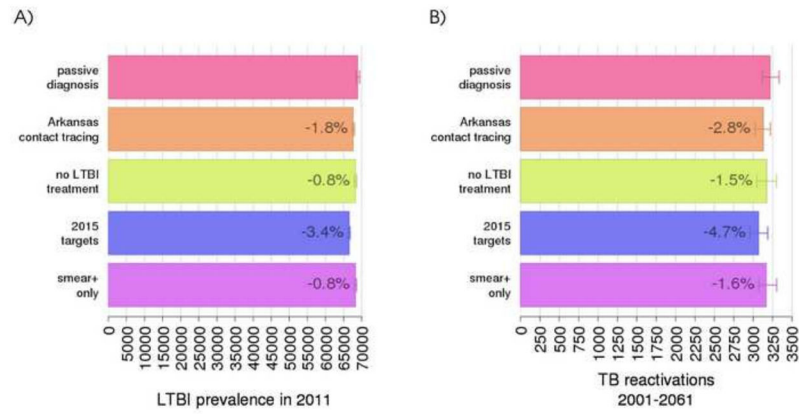
**Figure 4.**
Model validation comparing data and model predictions. A) fraction of clustered TB over total of cases for the whole population and restricted by age group as in [21]; B) distribution of cluster sizes [21]; C) proportion of TB in foreign-born individuals over time [20]; D) proportion of TB in foreign-born individuals by time of sojourn in the US [20] for years 2005–2011; E) prevalence of secondary TB cases in household [22] and workplace [9] contacts of smear positive index cases; F) prevalence of LTBI in household [22] and workplace [9] contacts of smear positive index cases. Average and 95-percentile error bars for the model are calculated over 100 best parameter sets and 100 stochastic simulations of the model.

**Figure 5.**
Impact of contact investigation programs on TB cases (A) and TB deaths (B). Numbers within bars indicate the average percentage of avoided TB cases and deaths with respect to a passive diagnosis program. An expanded description of each program is reported at the end of section Methods. Average and 95-percentile error bars for the model are calculated over 100 best parameter sets and 100 stochastic simulations of the model.

**Figure 6.**
Impact of contact investigation programs on LTBI prevalence cases in 2011 (A) and cases of endogenous TB reactivation between 2001 and 2061 (B). Numbers within bars indicate the average percent reduction with respect to a passive-only diagnosis program. An expanded description of each program is reported at the end of section Methods. Average and 95-percentile error bars for the model are calculated over 100 best parameter set and 100 stochastic simulations of the model.