Web Appendix 1. Demographic and epidemiological processes

Modeled processes (TB transmission, aging, immigration, etc) were simulated by transitions between model compartments, represented by arrows in Figure 1. This section describes modeled processes, how these were operationalized, and evidence used to define transition rates. Model parameters (such as those describing transition rates) were based on the best available evidence, yet in many cases substantial uncertainty exists as to the most appropriate parameter values to use. This uncertainty was operationalized as probability distributions constructed for each parameter. In the text below parameter uncertainty is summarized as a mean value plus the equal-tailed 95% interval for the underlying probability distribution (mean value [lower bound of interval, upper bound of interval]). Unless indicated otherwise, all rates are presented per person-year.

1. Initial population

The model was initialized to represent a 1950 population. The model was populated to match the population size, age distribution, and nativity in 1950 using data from the US Census Bureau (1, 2).

1.2. Births


Parameter format

Total births by are specified by month, from 1950 through 2100. A fixed fraction ($p^{HR}$) of total births are allocated to the high-risk compartment, the remainder (1-$p^{HR}$) enter low-risk.

Data sources

Historical birth data 1900-2001 were derived from the US Census Bureau Statistical Abstract (3) and then National Vital Statistics Reports for 2002-2013 (4). Projections for future years were taken from US Census Bureau National Population Projections 2014 (5). Annual estimates were divided by 12 and smoothed with a cubic spline smoother to obtain monthly estimates. It was assumed that the fraction of the population born into the high-risk group is approximately 0.005 [0.0025, 0.01] (further discussion of the high-risk group is given in Section 1.8).

1.3. Mortality

Mortality rates were calculated as the sum of background mortality and disease-specific mortality. Background mortality was assumed to vary by calendar year, age group, and risk group. Disease-specific mortality was assumed to be the sum of TB- and HIV-specific mortality rates, and applied to the active TB and HIV positive compartments respectively.

Parameter format

Background mortality rates ($\mu_{bg}$) for the general population were operationalized as monthly values from 1950 though 2100, stratified by age group. Background mortality rates for the high-risk group were assumed to be a multiple of the general population mortality rate for a given month and age group, operationalized as a rate ratio ($RR^{HR}_{bg}$). TB-specific mortality rates ($\mu_{TB}$) were described for active TB (smear-negative, smear-positive) as well as TB treatment compartments. HIV-specific mortality rates ($\mu_{HIV}$) were specified for each HIV-positive compartment. Total mortality rate for a given model compartment was calculated as the sum of background mortality, TB-specific mortality, and HIV-specific mortality rates.

Data sources

Background mortality

Historical mortality rates were derived from the life tables provided in the National Vital Statistics Report series (4). Background mortality rates were calculated from survivorship estimates ($l_x$), providing mortality rates that reproduce reported total mortality in each age band. A cubic spline smoother was used to interpolate between estimates from decennial life tables. Future trends in mortality rates were extrapolated based on Bell and Miller (6),
who report annual percentage reductions in central death rates by age group for the US social security area through 2100.

**Excess mortality due to homelessness**

The rate ratio describing additional mortality in the high-risk group was informed by studies in Canada, Rotterdam and Boston (7-9), with the parameter value (3.4 [2.3, 4.8]) chosen to produce a 15 year (range: 10-20 years) reduction in life expectancy.

**Excess mortality due to TB**

Recent data are not available on TB mortality in the absence of treatment. However, Tiemersma and colleagues report a qualitative review of historical data that provide indirect evidence about TB mortality rates (10). They estimate an average duration of TB disease of 3 years, with case fatality of 70% and 24% for smear-negative and smear positive respectively. If self-cure rates are assumed to be in the range of 20% per year, the estimates reported by Tiemersma are consistent with mortality rates in the range of 0.08-0.12 and 0.20-0.25 for smear-negative and smear-positive TB respectively. However, these rates apply to high-burden settings, and it is likely that mortality rates are lower in a setting such as the United States, due to better nutrition as well as rapid diagnosis and treatment for symptomatic disease. We assumed a mortality rate centered at 0.1 with a broad uncertainty interval [0.02, 0.25] for model compartments <15 years of age with smear-positive TB, and assumed a risk ratio of 0.5 [0.3, 0.8] for smear-negative TB relative to smear-positive disease. No TB-specific mortality risks were assumed for latent infection or following recovery from active disease. TB-specific mortality risks for TB treatment compartments were calibrated against reported treatment outcomes. We assumed increasing TB mortality risks as a function of age (mortality rate 1.7 [1.3, 2.3] times higher for each additional age group above 15 years old), based on reporting data showing a monotonic upward trend in the ratio of TB mortality to TB cases as a function of age.

**Excess mortality due to HIV**

HIV-specific mortality rates were assumed to differ by CD4 cell count category and receipt of ART. Reported estimates of untreated HIV mortality (11-14) suggest a mortality rate of 0.1-0.2 for CD4<350 and minimal excess mortality for CD4>350. We adopted a mortality rate of 0.010 [0.005, 0.040] for untreated HIV with CD4<350, and 0.10 [0.05, 0.20] for untreated HIV with CD4<350. Survival for HIV-infected individuals declines with age (15), which was modeled via elevated HIV progression rates for age groups >25 years of age (Section 1.15), and increasing HIV-specific mortality rates for successive age groups (mortality rate 1.7 [1.3, 2.3] times higher for each additional age group). ART produces rapid and sustained mortality rate reductions (16-20). We assumed a mortality rate of 0.010 [0.005, 0.030] for individuals initiating ART with CD4<350 (ART2 compartment). We assumed no HIV-specific mortality applied to the ART1 compartment, though allowed a transition rate of 0.023 [0.005, 0.05] from the ART1 to the ART2 compartment (Section 1.15) to allow for eventual regimen failure for individuals initiating ART with high CD4 cell count.

**Comorbid TB-HIV**

High mortality rates have been observed among individuals with comorbid active TB and advanced HIV. van der Sande et al (21) report a mortality rate of 0.44 [0.36, 0.52] based on a cohort with a median CD4 of 200 and incident TB. We allowed for an additional increment of 0.2 [0.1, 0.4] to the disease-specific mortality rate for compartments describing active TB and untreated HIV with CD4<350.

1.4. Immigration

A proper accounting of immigration is important for US TB epidemiology, given the role of TB in non-US-born individuals in driving the domestic TB epidemic. Immigration was modeled as an additional source of entrants to the model (in addition to births), with these new entrants added to the ‘recent arrivals’ compartment and distributed across all other model subdivisions. All new migrants were assumed to be HIV negative.

**Parameter format**

**Total immigration by month and age**

The absolute number of new migrants was described as monthly values from 1950 through to 2100. For each month, the total number of new migrants was stratified across modeled age groups using a fixed age distribution.
Active TB / rapidly progressing M. tb infection

Historical data suggest high TB incidence rates among recent migrants. In the model we allowed for two hypotheses that could explain this outcome: (i) that some individuals were entering the country with active TB disease (despite screening prior to entry), and (ii) that some individuals entering the country experienced recent M. tb infection prior to entry, and thus were at high risk of rapid progression. To operationalize this, a fraction of all migrants were assumed to have either active TB or rapidly progressing M. tb infection (‘Latent fast’ in Figure 1) upon arrival in the United States. A time trend was described for the prevalence of these individuals as a fraction of all arriving migrants, with changes in prevalence of this group assumed to be proportional changes in global TB prevalence. Prevalence for each age group were assumed to be proportional to this overall time trend, with the largest migrant age group (25-34 year olds) used as the reference category. A fixed fraction of these individuals entered the active TB compartments, and the rest entered the Latent Fast compartments. Among new migrants entering the active TB compartments, a fraction were assumed to be treatment experienced, with this fraction stratified by age group. The distribution of active TB disease across drug resistance subdivisions was assumed to be a function of calendar time and treatment history.

Latent TB infection

A fraction of all new migrants were assumed to be latently infected with TB (‘Latent slow’ in Figure 1), with this fraction varying by age and month. A functional relationship for LTBI prevalence was assumed, with prevalence at age \( y \) and year \( t \) assumed equal to \( 1 - e^{-a(t+b)y+c(t+y)} \). The time trend \( a \) was based on the time trend for active TB prevalence, while parameters \( b \) and \( c \) were fitted to empirical data to describe age trends in LTBI prevalence. All latently infected individuals were assumed to be treatment-naïve and infected with a pansensitive TB strain.

Data sources

Time trends in immigration volume

Monthly estimates of the total number of new migrants entering the US were derived from several sources. For the period 1950-2000, the number of new migrants entering the US each year was derived from the time series of new Legal Permanent Residents (LPRs) reported by the Department of Homeland Security’s Office of Immigration Statistics (DHS-OIS) (22), stratified by age and country of origin. Several adjustments were made to these estimates to account for known deviations between the reported number of LRP entrants and actual immigration volume:

1. Temporary visitors and undocumented immigration: By definition, the LPR data exclude individuals legally present in the US on temporary visas, as well as undocumented migrants. For the purposes of this analysis short-term visitors were not modeled with explicit entry into and exit from the model. Instead, their effect on US TB epidemiology was captured through an exogenous component of the TB force of infection, described in Section 1.9. To account for undocumented migrants, the absolute number of new LPRs was inflated on a country-by-country basis. To do so, estimates of the undocumented resident population in 2012, stratified for major sender countries (23), were compared to estimates of the total resident population (legal and undocumented) for the same time period and countries, derived from the American Community Survey (24). This comparison was used to create country-specific inflation factors for the top 10 sender countries (by number of undocumented migrants), and then a common inflation factor applied to all other countries.

2. Discrepancy between date of entry to the country and date of gaining LPR status: Approximately half of all LPR migrants enter the United States several states before gaining LPR status (representing ‘change-of-status’ LPR applications). This time-lag could influence projections if there were major short term fluctuations in ‘change-of-status’ LPR applications. This is the case with the Simpson-Mazzoli Act of 1986, which led to a large number of previously-undocumented residents gaining LPR status over a short time period producing large spike in the LPR time series around 1990-91 (Web Figure 1, Panel A), particularly for Mexico and some other Latin-American countries. To create a time series for immigration volume we first estimated the additional number of new LPR migrants reported for the 1990 decade\(^1\) over and above a trend line created by assuming linear change in immigration volume between the 1980 decade and 2000/01. This additional immigration volume was redistributed equally over the period 1960-2000. This adjustment was implemented on a country-

---

\(^1\) For the datasets reporting a detailed disaggregation by sender country, data prior to 2000 are summarized by decade.
by-country basis, with no adjustment made if the number of new migrants reported for the 1990 decade fell below the linear 1980-2000/01 trend line.

Web Figure 1: Total Volume and Composition of New Migrants Entering the United States.
For each country, a cubic spline smoother was used to interpolate between the decennial immigration estimates provided for earlier years (1950-2000).

For more recent years (2000-2015) we used data from the American Community Survey (25) to estimate the annual number of new migrants (operationalized as the number of non-US-born individuals reporting residence outside of the US the previous year).

Projections of future immigration volume were based on US Census Bureau immigration projections for Mexico and major world regions (26). These projections cover the period 2015 through 2060, and were extended to 2100 with linear extrapolation. Finally, estimates of immigration volume from each sender country were summed to obtain an estimate of overall immigration volume for each year. Panel A of Web Figure 1 presents the annual estimates of total immigration volume used for the model, overplotted with the LPR and ACS estimates used to construct this time trend.

Age distribution of new migrants

The age distribution of new migrants was assumed to be equal to the age distribution of new LPRs. This age distribution has shown little fluctuation over the period for which detailed data are available (1986-2013), and so the age distribution of new migrants (Panel B of Web Figure 1) was calculated from the combined LPR data over this period.

Latent TB infection

Estimates of LTBI prevalence derived from NHANES for 2011 (27) were used to parameterize the function for LTBI prevalence in new migrants as a function of age. The fitted curve produced by this analysis is shown in Panel E of Web Figure 1, over-plotted with the survey estimates for comparison.

Active TB infection

A time trend in active TB prevalence among migrants was calculated by assuming that TB prevalence in the sender country and TB prevalence in migrants from that country are related by a risk ratio ($RR_{TB prev}$), common across countries and years. With this assumption, a prevalence trend for new migrants can be calculated by averaging across reported TB prevalence values reported for all sender countries, weighted by the fraction of total immigration volume contributed by that country in a particular year:

$$Prevalence_{migrants} = \frac{\sum_{gen\, pop\, prev\, \mu, t} \times total\, migrants_{\mu, t} \times RR_{TB\, prev}}{\sum_{total\, migrants_{\mu, t}}} \text{, for year } t \text{ and country } \mu$$

Even if the assumption that $RR_{TB\, prev}$ is constant across countries is violated, given the number of countries included in the weighted average the estimation strategy should be somewhat robust to these violations.

Estimates of general population TB prevalence for sender countries were obtained from the WHO TB database (28), covering the period 1990-2014. Prevalence values for earlier years were generated by extrapolating backwards from the 1990 level for each country assuming that prevalence fell by 0.55% per year over the period 1950-1990 (this 0.55% value was estimated from the average annual reduction in TB prevalence across all countries over the period 1990-2000). While there is little reason to be confident about this backwards extrapolation the analysis should be relatively robust to uncertainty in these values given that absolute migration volume was much lower in the earlier years. For example, assuming that global TB prevalence declined at a rate of 2% per year over the period 1950-1990 (a rate 3.6 times higher than assumed under the base case, and producing 1950 prevalence values over two times higher) only results in a 15% increase in the total number of active TB cases entering the US between 1950 and 2013. While recent years have seen global TB prevalence decline at 3-4% per year, for the base case it was conservatively assumed that in the future TB prevalence in sender countries would decline at 1.5% per year, consistent with other projections (29).

Most migrants to the US (all persons applying for US refugee or immigration status from an overseas location) undergo screening for active TB disease prior to receiving clearance to travel. In 2007 the CDC revised the technical instructions for this screening from an approach based on chest x-ray and sputum smear (30) to a more sensitive approach requiring diagnosis via sputum culture (31). This new screening approach was introduced between 2007 and 2012, over which time period the number of new TB cases reported among individuals within 1 year of arrival dropped by –40% (32). To operationalize this change in screening practice, the model allowed for a linear change in the prevalence of active TB among the migrant population between 2007 and 2012, to a final value.
that is some multiple of the value at the start of 2007, on top of the other trends already described. This multiple was given a weakly-informative prior centered at 1.0 and bounded at 0.0 and 2.0.

The distribution of the migrant pool across sender countries in each year was already generated as part of the estimation of total immigration volume (described above under ‘Time trends in immigration volume’), and these estimates were combined with the country prevalence estimates to calculate the time trend in overall TB prevalence in the combined migrant population. This TB prevalence time trend (weighted average of prevalence in sender countries) is shown in Panel C of Web Figure 1. This time trend was used to provide TB prevalence values for the largest migrant age group, 25-34 year-olds. Relative TB prevalence in other age groups, described by risk ratios, were estimated by dividing the number of TB cases reported among recent (<4 years in the US) migrants for each age group over the period 1993-2013 (33) by the total migrants over the same period and age group, using 25-34 year-olds as the reference category. These risk ratios are shown in Panel D of Web Figure 1.

**History of prior TB**

The fraction of individuals with active TB entering the US with a history of prior TB treatment was estimated using the national TB case reporting data (33): for each age group, the number of TB cases among recent migrants over 1993-2013 who reported prior TB treatment was divided by the total number of TB cases for that age group and time period (excluding those for whom TB treatment history was not reported). The resulting estimates are shown in Panel A of Web Figure 2, with the fraction with prior treatment experience rising from less than 1% in the 0-4 age group to greater than 8% for those over 55 years of age. Finally it was assumed that 10% of individuals entering the country with active TB had smear-positive disease. This fraction is much lower than would be expected for TB in the general population (34), given the higher likelihood that individuals with active TB will be identified during pre-immigration screening (applies to legal migrants).

**TB drug resistance**

The distribution of TB drug resistance in new migrants with active disease was estimated from a number of data sources. The national TB case reporting data (33) contain information on MDR-TB status for individuals receiving drug sensitivity testing (~99% of those with culture-confirmed TB), and these data were analyzed to estimate the prevalence of MDR-TB for TB cases among recent (<4 years in the US) migrants, stratified by treatment history. This produced estimates of 2.2% and 15.3% MDR-TB prevalence for treatment-naive and treatment-experienced cases respectively. To obtain estimates for the prevalence of INH-resistant/RIF-sensitive and INH-sensitive/RIF-resistant disease, it was assumed the prevalence of these strains relative to MDR-TB matched the results of Espinal and colleagues (35). This study synthesized the results of a large number of drug resistance surveys in different countries, finding that 5.2%, 0.2%, and 1.0% of all treatment naïve TB cases exhibited INH-resistant/RIF-sensitive TB, INH-sensitive/RIF-resistant TB, and MDR-TB respectively. For treatment experienced cases these same fractions were 10.3%, 2.7%, and 9.3%. When these results are scaled to match the results from the US case reporting data they produce estimates of 11.2% and 0.4% prevalence for INH-resistant/RIF-sensitive TB and INH-sensitive/RIF-resistant TB among treatment naïve cases, and 16.9% and 4.4% for these strains among treatment experienced patients. Analyses reported by Zignol et al (36) were used to estimate the fraction of all MDR-TB that was also XDR-TB. This percentage (9.4%) was used to subdivide MDR-TB estimates for current migrants into ‘MDR-TB, non-XDR’ and XDR-TB fractions. Given the lack of accurate information about past years, it was assumed that the time trend in the prevalence of drug-resistant strains followed a smooth monotonic increase. This was operationalized as a logistic curve increasing from 0% prevalence to current values between 1970 and 2010 for first-line drug resistance, and between 1995 and 2010 for 2nd-line drug resistance. The consequences of these assumptions are shown in Panel B (treatment naïve cases) and Panel C (treatment experienced cases) of Web Figure 2. The fraction of cases with MDR-TB estimated for each year from the TB case reports is plotted against the estimates prepared for the model.
Uncertainty

A parsimonious approach was taken for expressing the uncertainty associated with these different assumptions describing migrants:

i. Uncertainty in the volume of historical immigration (1950-2015) was implemented using a parameter \([1.0 \ [0.5, 1.7]\) that multiplied the time series of immigration estimates (blue bars in Panel A, Web Figure 1), shifting the entire trend-line up or down proportionally. This prior distribution represents substantial uncertainty, and to ensure reasonable estimates of the foreign-borne population the model was also calibrated to cross-sectional data on the resident foreign-born population derived from the decennial census and American Population Survey (Web Figure 2).

ii. Uncertainty in future immigration volume was introduced as variability in the slope of the projection for total immigration volume post-2015, implemented by adding a random term to the slope of the trend line (distributed Normal with mean zero and 95% interval \([-10, 10]\) thousand per year). We also allowed for
temporal variation around this overall trend, implemented via a Gaussian process that added multiplicative noise to the trend line. Web Figure 3 shows the resulting range of trajectories for future immigration volume.

iii. Uncertainty in the prevalence of active TB / rapidly-progressing infection in the migrant population was expressed through $RR_{TB}^{pred}$, which was given a broad prior distribution (1.0 [0.08, 3.0]). We adopted a diffuse prior for the fraction of this group entering the US with active TB (0.25 [0.13, 0.50]).

iv. Uncertainty in the prevalence of LTBI in the migrant population was expressed through uncertainty in the parameters used to describe LTBI prevalence as a function of age (Section 1.4: Latent TB infection), with $b = 0.2 \ [0.1, 0.4]$ and $c = 0.5 \ [0.3, 0.9]$.

v. Uncertainty in the future trends of active TB / rapidly-progressing infection as well as LTBI in migrants was operationalized as uncertainty in the annual rate of change assumed for the TB prevalence time trend in sender countries (prevalence in year $t+1$ as a multiple of prevalence in year $t$), assumed to drop by 2.0% per year in the base case. This parameter was given a relatively narrow prior (0.98 [0.96, 0.995]), given the effects of this parameter are magnified by annual compounding. Similar to immigration volume we allowed for temporal variation around the overall trend for each of these inputs, implemented via a Gaussian process that added multiplicative noise to the trend line. Web Figure 3 shows the resulting range of trajectories for active TB / rapidly-progressing infection as well as LTBI in migrants entering the United States in future years.

vi. Uncertainty in TB drug resistance among migrants was introduced as multiplicative uncertainty in the point estimates shown in Web Figure 2. For future projections we allowed additional uncertainty in the slope of the trend line as well as temporal variation, implemented via a Gaussian process. Web Figure 3 shows the resulting range of trajectories for MDR-TB among treatment naïve and treatment experienced migrants in future years.

1.5. Emigration

A rate of emigration was assumed for non-US-born individuals. In reality all population groups will have some rate of emigration, but these rates are generally low for US-born individuals and are ignored for the purposes of this analysis. Emigration rates are known to be higher for non-US-born individuals, and explicitly modeling the permanent departure for these individuals allows for the total number of non-US-born residents to follow a realistic time trend without reducing immigration volume.

**Parameter format**

A fixed emigration rate was applied to individuals in the non-US-born compartment, stratified by duration of residence.

**Data sources**

The US Census estimates emigration rates of 0.0024-0.003 for non-US-born individuals resident in the county for over 10 years, and 0.0138-0.0143 for those resident for 0-10 years (37). This 10-year threshold does not align with the definition of ‘recent migrant’ assumed for this model, but if it is assumed that two rates operate (i.e. one for our recent entrants compartment and one thereafter) the appropriate rates can be calculated. With this approach, an emigration rate of 0.05 [0.03, 0.07] was adopted for recent migrants and 0.003 [0.002, 0.004] for long-term non-US-born residents.

1.6. Aging

A fixed transition rate was applied to move individuals between successive age groups. This rate was calculated so that expected sojourn time in a given age compartment (ignoring other transitions) was equal to the number of years in the age band.

1.7. Progression from recent migrant to long-term non-US-born resident

A fixed transition rate was applied to move individuals in the recent migrant compartment to the long-term non-US-born resident compartment. Similar to the approach used for aging, this rate was calculated so that expected sojourn time in the recent migrant compartment (ignoring other transitions) was equal to the number of years specified in the definition of recent migrant.
1.8. Entry and exit from the high-risk population

Individuals were assumed to enter the high-risk compartment from the low-risk compartment (comprising the US-born general population), and for this reason the high-risk/low-risk distinction excludes the non-US-born population. The high-risk group was defined to include the homeless populations and associated marginalized groups who defining features (in the context of TB epidemiology) include poor access to care, higher background
mortality and HIV risks, higher TB effective contact rates, and assortative mixing with respect to the general population.

Parameter format

Rates of entry and exit from the high-risk group were defined as a function of age, with exponentially decreasing rates of entry/exit as a function of age used to represent increasing duration of residence in the high-risk group for older age groups. Entry rates were assumed to be a multiple of exit rates for each age group. Higher rates of entry to the high-risk population were assumed for HIV-positive individuals, given observed correlation of homelessness and HIV infection (38). We assumed a rate ratio of 2.0 [1.0, 3.0] for entry into the high-risk group for HIV-positive individuals as compared to the HIV-negative population, which was applied to each age group separately. In combination with higher rates of HIV incidence in the high-risk group (see Section 1.14), these assumptions will produce HIV prevalence levels in the high-risk group approximately 5-8 times higher than the low risk group.

Data sources

Given the definition of the high-risk group (including homeless populations plus similar marginalized groups) a single population size is difficult to define. Most recent estimates suggest that approximately 0.6 million individuals are homeless at a given point in time (39), yet many more will live in the conditions of unstable, overcrowded housing, poor access to services, and greater exposure to other health risks that characterize this group. For the purposes of this analysis it was assumed that the high-risk population would represent approximately 0.5% of the total population. Rates of birth into this population (Section 1.2) and rates of transition between low- and high-risk groups were defined to reflect this assumption, as well as other known features of homelessness: an age distribution that approximately matches that of the general US population, and longer average duration of episodes in older age groups (39-41). It was assumed that average duration of residence in the high-risk group was 3 years for the youngest age group, rising to 20 years for the 45-55 year-old age group. Exit rates were defined as the inverse of these residence times (i.e. 1/3 for the 0-4 age group, 1/20 for the 45-55 age group), with rates for other groups following an exponential decline defined by these two points. Rates of entry into the high risk group were calculated as the rates of exit for each age group divided by $p_{HR} \times 1.3$, with this factor calculated to ensure the overall fraction of the population in the high risk group is approximately 1%, given higher background mortality rates.

While the approach taken for modeling high-risk abstracts from many of the complex issues faced by the homeless population and associated groups, and ignores any temporal trends in the size and composition of these populations (41), it serves the purpose of allowing for heterogeneity of transmission risk and treatment access within the population, which has the potential to be important in the context of declining population-level TB risks (42).

1.9. TB transmission

The force of infection was calculated assuming frequency-dependent contact rates. Mixing was assumed to be assortative by risk group and by HIV status (dichotomized as HIV negative and HIV positive), and homogeneous within other subdivisions. To operationalize this assortative mixing, it was assumed that some fraction of contacts by members of a risk group or HIV status are reserved for ‘with-like’ interactions, while the remainder occur in a setting with homogeneous mixing. This follows the approach described by Jacquez et al (43) and Busenberg and Castillo-Chavez (44), and applied by multiple subsequent models. For simplicity, assortative behavior was modeled explicitly for non-US-born (both recent and long term combined), high-risk, and HIV positive groups (i.e. each is assumed to withhold a fraction of contacts for with-like interactions) but not for low-risk and HIV negative groups. The fraction withheld is denoted $\sigma_{NUSB}$, $\sigma_{HR}$, and $\sigma_{HIV}$ for non-US-born, high risk, and HIV positive respectively. It should be noted that the assumption of assortative mixing by these groups induces assortative mixing in the low risk and HIV negatives. One complication of the current analysis is the crossing of risk group with HIV status. For simplicity it was assumed that individuals with two such characteristics (e.g. high risk and HIV positive) form their own mixing group. This produces six groups within that mixing was assumed to be homogeneous: (A) the common pool, exclusive mixing for (B) non-US-born, (C) high risk, and (D) HIV positive respectively, and exclusive mixing for (E) non-US-born and HIV positive, and (F) high risk and HIV positive, as shown in Web Figure 4. The probability of transmission per effective contact in each of these groups is given as $Y_A, Y_B, Y_C, Y_D, Y_E, Y_F$, respectively.

These assumptions allowed for a parsimonious approach to mixing, governed by the parameters $\sigma_{NUSB}$, $\sigma_{HR}$, and $\sigma_{HIV}$. Varying these parameters between 0.0 and 1.0 represents extremes of homogeneous mixing and complete
with-like assortative mixing respectively. In addition, the three risk groups were allowed to have different effective contact rates operationalized as $\beta_k$ for risk group $k$ ($k = 1$ for low risk, 2 for high risk, and 3 for non-US-born). In the mixing calculations given below it is assumed that differences in $\beta_k$ represent differences in the average number of contacts per unit time, rather than differences in the relative infectiousness of each contact, thus individuals with a high effective contact rate contribute a greater number of contacts to a mixing group. Effective contact rates were assumed to be constant over time.

Web Figure 4: Schematic of Mixing Groups*

* Different colors represent different mixing groups (A through F). Where a particular HIV and risk group participates in multiple mixing groups, the fraction of their contacts contributed to each mixing group is shown.

Each TB strain was assumed to contribute separately to the overall force of infection, which was calculated as the sum of the force of infection for each modeled strain. Strains were assumed to have differential fitness, operationalized as $r_j$ for strain $j$, assumed to be equal to 1.0 for pansensitive TB and <1.0 for drug resistant strains. While the biological mechanism for reduced fitness likely involves differential ability to produce clinically relevant infection following transmission, for modeling convenience this differential fitness was operationalized as adjustments in the probability of transmission between strains. Finally, individuals with smear-negative disease were assumed to have substantially reduced (but non-zero) transmission risks compared to individuals with smear-positive disease. Individuals receiving TB treatment were assumed to exhibit ongoing transmission risks in the case of ineffective treatment, calculated as the transmission risk for untreated disease multiplied by the fraction of patients receiving ineffective treatment (1.0 minus the parameter $TxEff_t$, Section 3.4), multiplied by a tuning parameter that allowed the transmission rate for ineffectively treated patients to vary between zero and the risk observed for untreated patients (with a prior of 0.5 [0.1, 0.9]). These differences in transmission risk were operationalized as $q_i$, equal to the infectivity of individuals in TB compartment $i$ relative to those with smear-positive active disease.

The equations below express the assumptions made about TB transmission dynamics, and are used to calculate the force of infection by risk group, HIV status and TB strain.
For a susceptible individual, the probability of infection by TB strain $j$, per effective contact with individuals in risk group $k$ (1 for low risk, 2 for high risk, and 3 for non-US-born) and HIV status $l$ (categorized as 1 for HIV negative and 2 for HIV positive) is given by $\gamma_{jkl}$, where

$$\gamma_{jkl} = \frac{\sum X_{ijkl} q_{ij}}{N_{kl}}.$$  

$X_{ijkl}$ is the number of individuals in TB compartment $i$ with TB strain $j$ in risk group $k$ and HIV status $l$ (summing across other subdivisions), and $N_{kl} = \sum_i \sum_j X_{ijkl}$.

As mixing is homogeneous within each of the mixing groups A through F, the probability of infection by strain $j$ within each group was calculated as a weighted average of $\gamma_{jkl}$ based on the number of effective contacts each risk and HIV category contributes to that mixing group:

$$\gamma_{jA} = \frac{\beta_1 N_{12} Y_{12} + (1 - \sigma_{HR}) N_{12} Y_{12}}{(1 - \sigma_{HIV}) N_{12}}$$
$$\gamma_{jB} = \frac{(1 - \sigma_{HIV}) N_{22} Y_{22} + N_{22} Y_{22}}{(1 - \sigma_{HIV}) N_{22}}$$
$$\gamma_{jC} = \frac{(1 - \sigma_{HIV}) N_{32} Y_{32} + N_{32} Y_{32}}{(1 - \sigma_{HIV}) N_{32}}$$
$$\gamma_{jD} = \frac{\beta_1 N_{12} Y_{12} + (1 - \sigma_{HR}) N_{12} Y_{12} + (1 - \sigma_{NUSB}) N_{12} Y_{12}}{(1 - \sigma_{HIV}) N_{12}}$$
$$\gamma_{jE} = \gamma_{j32}$$
$$\gamma_{jF} = \gamma_{j22}$$

The force of infection for each TB strain, risk group and HIV status ($\lambda_{jkl}$) can be calculated as the product of the effective contact rate for that group and an average of $\gamma_A$ through $\gamma_F$, weighted by the proportion of contacts devoted to each mixing pool. In addition, it was assumed that non-US-born individuals are exposed to an additional exogenous risk of infection due to contact with individuals in their sender country (either through travel back to sender country or contact with short-term visitors from that country). Such risks have been demonstrated in a number of settings (45–48). This exogenous infection risk was modeled as an additional time-dependent term ($\pi_t$) in the force of infection.

$$\lambda_{j11} = \beta_1 Y_A$$
$$\lambda_{j12} = \beta_1 [\sigma_{HIV} Y_D + (1 - \sigma_{HIV}) Y_A]$$
$$\lambda_{j21} = \beta_2 [\sigma_{HR} Y_C + (1 - \sigma_{HR}) Y_A]$$
$$\lambda_{j22} = \beta_2 [\sigma_{HIV} (\sigma_{HR} Y_F + (1 - \sigma_{HR}) Y_D) + (1 - \sigma_{HIV}) (\sigma_{HR} Y_C + (1 - \sigma_{HR}) Y_A)]$$
$$\lambda_{j31} = \beta_3 [\sigma_{NUSB} Y_B + (1 - \sigma_{NUSB}) Y_A] + \pi_t$$
$$\lambda_{j32} = \beta_3 [\sigma_{HIV} (\sigma_{NUSB} Y_E + (1 - \sigma_{NUSB}) Y_D) + (1 - \sigma_{HIV}) (\sigma_{NUSB} Y_B + (1 - \sigma_{NUSB}) Y_A)] + \pi_t$$

The overall force of infection is equal to the sum of the force of infection for each TB strain:

$$\lambda_{kl} = \sum_j \lambda_{jkl}.$$  

Individuals in both ‘uninfected’ compartments of the core TB model are at risk of (re)infection. Newly-infected individuals transition to latent infection or active disease, as described in Section 1.10.

Individuals in the ‘latent slow’ compartment were assumed to be at risk of reinfection. As the model does not explicitly allow for mixed infection, it was assumed that any reinfection results in transition to the compartment of the reinfecting strain, with the same force of infection experienced by the uninfected individuals.
Parameter format

Strain fitness ($r_f$) was stratified by strain (equal to 1.0 for the pansensitive strain, <1.0 otherwise). Relative infectiousness ($q_f$) was stratified by TB and smear status (equal to 1.0 for smear positive active disease, <1.0 otherwise). The effective contact rate ($\beta_f$) was assumed to differ by risk group, with a rate ratio ($RR^\beta_{HR}$) used to specify a higher rate for the high-risk group compared to the rest of the population. For non-US-born individuals, an exogenous component of their force of infection ($\pi_{jt}$) was specified, assumed to vary by time and drug-resistance category.

Data sources

Direct evidence on the mixing parameters is not available, and for $\sigma_{FB}$, $\sigma_{HR}$, and $\sigma_{HIV}$ (defined between 0 and 1) we assumed weak prior distributions (0.5 [0.1, 0.9]). Similarly, the overall effective contact rate and rate ratio for the high-risk group were given broad priors (10.0 [5.0, 15.0]) and 3.0 [2.0, 5.0] respectively. The relative infectiousness of smear-negative as compared to smear-positive TB (0.23 [0.12, 0.38]) was taken from studies analyzing TB cluster data (49, 50). Transmission risks were assumed to drop to a low level immediately upon initiation of TB treatment, with the relative risk of transmission calculated as $1 - T_xEff_{it}$, where $T_xEff_{it}$ is the fraction of TB patients receiving effective treatment (Section 2.4). Strain fitness estimates were synthesized from a set of modeling studies (51-54), and assumed to be 0.95 [0.90, 0.97], 0.85 [0.74, 0.92], and 0.73 [0.57, 0.84] for mono-INH resistant, mono-RIF resistant, and MDR- or XDR-TB respectively.

A prior distribution for the exogenous infection risk $\pi_{jt}$ is difficult to construct. While individual cases of infection acquired abroad have been documented (45-48), it is unclear how these reports can be translated into a population-level incidence rate. Similarly, while it is likely that short-term visitors with undiagnosed TB contribute to the force of infection, a quantitative estimate of this effect is not available. We assigned a diffuse prior to the value of $\pi_t$ ($\pi_{jt}$ summed across drug resistance categories) of 5.0 [1.0, 10.0] per 100,000 person-years. Given the approximately 42 million non-US-born individuals currently resident in the United States, and assuming ~5% of all newly infected individuals progress rapidly to active disease, this rate is consistent with roughly 100 new TB cases per year, or 1.5% [0.3, 3.0] of total non-US-born TB cases diagnosed in 2014. We assumed $\pi_t$ was proportional to the average prevalence of TB infection in sender countries (Panel C of Web Figure 1), obtained $\pi_{jt}$ by dividing $\pi_t$ across TB strains based on the distribution of drug resistance in imported cases (25-34 year-old age group) in each year.

1.10. Progression to active TB following TB infection

The risk of developing TB is known to decline as a function of time since infection. To provide flexibility in modeling this decline, we assumed some fraction of newly infected individuals progress to active TB over a relatively short time frame (primary progressive TB) while the rest enter a ‘latent slow’ compartment with low progression rates (endogenous reactivation). Those exhibiting primary progressive TB either progress instantaneously to active disease or enter a ‘fast latent’ compartment with a rapid rate of progression, which effectively spreads their progression to active TB over the 2-4 years following infection. We also allowed a slow rate of clearance of latent TB infection for individuals in the ‘latent slow’ compartments. The consequence of this model structure and parameterization is that a cohort of newly infected individuals will exhibit a monotonically declining risk of developing active TB as a function of time since infection.

The details of progression to active TB were assumed to differ by age and HIV status. In addition, prior TB exposure was assumed to provide partial protection against progression to active disease, operationalized as a reduction in the risk of fast progression for individuals reinfected with TB. This reduced risk was applied to individuals in the ‘uninfected, partially immune’ and ‘latent slow’ compartments. Individuals with advanced HIV disease were assumed to experience no reduction in progression risk due to prior TB exposure.

Parameter format

A parameter ($p^{fast}$) was defined for the fraction of all newly infected individuals who experience primary progressive TB. Additional parameters define the fraction of this group progressing instantaneously ($p^{immediate}$), and the rate of progression for the fraction of this group not progressing instantaneously ($r^{fast}$). Parameters were specified for the rate of endogenous reactivation from ‘latent slow’ to active disease ($r^{slow}$), and rate of clearance of from latent slow’ to ‘uninfected, partially immune’ ($r^{clear}$) (55).
The fraction of experiencing primary progressive TB \( (p_{\text{fast}}) \) was assumed to differ by age (different values assumed for the 0-4 and 5-14 year old age groups), HIV status, and prior TB exposure. These differences were operationalized as odds-ratios \( (OR_{0-4}, OR_{5-14}, OR_{\text{HIV}}, OR_{\text{fast prior \, TB}}) \). Higher progression rates were assumed for the ‘Latent slow’ compartment in the case of HIV infection \( (c_{\text{slow}}^{\text{HIV}}) \), and we also allowed for the potential for faster progression rates among the non-US-born \( (OR_{FB}^{\text{slow}}) \), and for older adults \( (OR_{age}^{\text{slow}}) \) (56).

Parameter values for individuals with early HIV (CD4=350) were calculated as a linear combination of the values for HIV negative individuals and individuals with advanced HIV disease, governed by a parameter \( f_{\text{early HIV}} \). For example, the fraction experiencing fast progression was calculated as \( f_{\text{early HIV}} \cdot p_{\text{HIV}}^{\text{fast}} + (1 - f_{\text{early HIV}}) \cdot p_{\text{HIV}}^{\text{fast}} \), such that higher values of \( f_{\text{early HIV}} \) produce parameter values for early HIV similar to those for advanced HIV, and lower values of \( f_{\text{early HIV}} \) produce parameter values similar to those for HIV negative individuals. A similar approach was used to operationalize the effect of ART in reducing the impact of HIV on TB natural history, with a parameter \( f_{\text{ART effect}} \) used to calculate parameter values for individuals on ART as a linear combination of the values for HIV negative individuals and HIV positive individuals not receiving ART. Lower values of \( f_{\text{ART effect}} \) produce parameter values for individuals on ART similar to those for untreated HIV infection, and higher values of \( f_{\text{ART effect}} \) produce parameter values similar to those for HIV negative individuals. This approach—utilizing \( f_{\text{early HIV}} \) and \( f_{\text{ART effect}} \) to calculate parameter values for early HIV and ART as a linear combination of the parameters for HIV negative and advanced HIV—was used for all parameters describing TB natural history, with the exception of the mortality rates (Section 1.3).

Data sources

The three pathways to active disease described above (via instantaneous progression, via the fast latent compartment, via the ‘latent slow’ compartment) cannot be distinguished empirically, and direct evidence on individual parameter values is not available. Instead, parameters were chosen by fitting the model to observational data on the distribution of new cases as a function of time since infection, so that the combination of parameters matched the available evidence. These observational data are included in the calibration likelihood, and include data from the early UK MRC BCG trials (57, 58), data from the placebo arm of the US Public Health Service isoniazid trials (59, 60), and more recent data from long-term follow-up of contacts of TB cases in the Netherlands (61). Priors for these parameters were 0.064 \([0.003, 0.2]\) for \( p_{\text{fast}} \), 0.25 \([0.05, 0.54]\) for \( p_{\text{immediate}} \), 0.55 \([0.19, 1.1]\) for \( r_{\text{fast}} \), 0.0007 \([0.00004, 0.002]\) for \( r_{\text{slow}} \), and 0.003 \([0.01, 0.06]\) for \( r_{\text{clear}} \).

Odds ratios describing the elevated risk of fast progression in the 0-4 year old age group \( (OR_{0-4}^{\text{fast}} : 2.5 \, [0.8, 5.1]) \), and the lower risk in the 5-14 year old age group \( (OR_{5-14}^{\text{fast}} : 0.5 \, [0.2, 1.0]) \), were derived from a review of observational studies of pediatric TB (62). Individuals with advanced HIV infection have much higher rates of TB incidence compared to HIV negative individuals. There is evidence to suggest that in high-burden settings recent infection is a major source of this increased incidence (63, 64), and it is likely that rates of endogenous reactivation are also higher given the loss of immune function. The odds ratio describing the elevated risk of fast progression for individuals with advanced HIV \( (OR_{HIV}^{\text{fast}}) \) was assumed to be 10 \([4, 19]\), with a wide confidence interval. The elevated rate of endogenous reactivation for individuals with HIV \( (r_{HIV}^{\text{slow}} : 0.042 \, [0.001, 0.146]) \) was taken from observational studies of HIV in the era before widespread availability of ART (65, 66). As the TB cases identified in these HIV studies likely included recent reinfection in addition to endogenous reactivation, progression rates from the studies were adjusted down in order to develop endogenous reactivation rates. The odds ratio describing elevated risk in the non-US-born (as compared to US-born) was given a prior of 1.2 \([1.0, 1.5]\), and we assumed a 1.2 \([1.1, 1.4]\) factor increase in the endogenous reactivation rate for each successive age-group above 65 years of age.

The odds ratio describing the reduced risk of fast progression in individuals with prior TB exposure has been estimated by a number of studies. Andrews et al report a meta-analysis of 23 cohorts including individuals with and without latent TB infection, and report an incidence rate ratio of 0.21 \([0.14, 0.30]\) for active TB in the previously infected group as compared to the uninfected group, with non-trivial heterogeneity between studies. Other analyses using indirect estimation techniques have generally found higher values (67-69). For this study we centered the parameter at the point-estimate reported by Andrews et al but with an expanded uncertainty interval \( (0.21 \, [0.08, 0.40]) \).
A prior distribution for $f_{\text{early HIV}}$ (0.19 [0.05, 0.38]) was based on the relative rates of TB incidence between CD4 strata in Antonucci et al. (66). Prior distributions for $f_{\text{ART effect}}$ were based on the results of a recent meta-analysis (70), with values of 0.57 [0.42, 0.72] for individuals with early HIV and 0.75 [0.57, 0.90] for individuals with advanced HIV.

1.11. Smear status for incident TB cases

Individuals developing active TB are divided between smear-negative or smear-positive disease, with this fraction depended on age and HIV status. Given the advent of improved diagnostics (e.g. NAAT) the distinction between smear-negative/positive cases is primarily used to model the heterogeneity in mortality and transmission risk among those with active TB.

Parameter format

A parameter ($p_{\text{smr pos}}$) is defined for the fraction of all incident cases of active TB among adult HIV negative individuals who are smear positive. Odds ratios are defined to calculate values for the youngest age group ($OR_{0-4}^{\text{smr pos}}$) and for HIV positive individuals ($OR_{\text{HIV}}^{\text{smr pos}}$) given lower rates of smear-positivity observed in these groups.

Data sources

The rate of smear positivity in incident TB cases is difficult to measure empirically. Consequently, this parameter given a weak prior distribution (0.50 [0.25,0.75]) and then calibrated to data on the fraction of prevalent TB found to be smear-positive in a population with low treatment access, based on results from the first 2 years (1960-61) of the studies on TB epidemiology undertaken by Karel Stýblo and colleagues in Kolin, Czechoslovakia (34). This fraction (38% of all prevalent adult cases smear-positive) was assumed to apply to the US population in 1950. The odds ratio for smear positivity in the youngest age group (0.07 [0.02, 0.17]) were derived from the results of Vynnycky and Fine (67), and the odds ratio for HIV positive individuals (0.28 [0.14, 0.47]) was based on reported differences in smear positivity in diagnosed TB cases (71-73).

1.12. Smear conversion

Individuals with smear-negative active TB have a slow rate of transition to smear-positive disease. It was assumed that this rate is zero for individuals with advanced HIV infection.

Parameter format

A parameter ($p_{\text{smear conversion}}$) is defined as the rate at which individuals with smear-negative TB convert to smear-positive TB.

Data sources

Few recent data are available on smear conversion rates (as this would require observation of a cohort of smear-negative TB not initiated on treatment). This rate was estimated from a study conducted by the Hong Kong Chest Service, as part of which a cohort of individuals with active TB but initially testing bacteriologically negative were followed for up to 60 months for bacteriological deterioration (74). The incidence of bacteriological progression in this cohort was consistent with a rate in the range of 0.09-0.24 depending on how the data are interpreted. These data also relate to culture conversion rather than smear conversion, but do provide some guidance on conversion rates. For the present study a rate of 0.15 [0.06, 0.29] was adopted.

1.13. Self-cure

Individuals with active TB self-cure at a low rate, returning to the ‘latent slow’ compartment. It is assumed that the rate of self-cure is zero for individuals with advanced HIV infection.

Parameter format

A parameter ($p_{\text{self-cure}}$) is defined as the rate at which individuals with active TB self-cure, stratified by HIV status.
Data sources

The rate of self-cure is difficult to measure empirically. However, indirect information on rates of self-cure is available from studies describing case fatality and duration of disease in the absence of treatment, which have been synthesized by Tiemersma et al (10). Based on this evidence we assumed a self-cure rate of 0.20 [0.15, 0.35].

1.14. HIV incidence

HIV transmission was not modeled dynamically. Instead, HIV incidence rates were taken as exogenous parameter inputs, assumed to vary varying by age, homelessness status, and calendar year.

Parameter format

An HIV incidence time trend for the general population was described based on historical estimates. Incidence time trends for individual age groups were specified as multiples of this general time trend (defined by a rate ratio for each age group). Similarly, the incidence rate in the high-risk compartment was assumed to be a multiple of the rate experienced by the low-risk compartment.

Data sources

A time series of estimated new HIV infections has been reported by the CDC based on HIV surveillance activities (75, 76). This time series was divided by the total population in each year to obtain an incidence time trend for the general population. We incorporated uncertainty in this historical time trend as a single parameter with a diffuse prior (1.0 [0.5, 1.5]) scaling the entire time series of estimates. Over recent years this time trend has demonstrated a slow decline of 1-2% per year, with incidence more or less stable as the population has grown. We assumed a 2% annual reduction in incidence for extrapolating this time trend into the future (Panel A, Web Figure 5). For future projections we allowed uncertainty in the slope of the trend line and set narrow bounds on this parameter (0.98 [0.96, 1.00]), given the compounding effect of annual changes. We also allowed for temporal variation in future HIV incidence, implemented via a Gaussian process. Web Figure 3 shows the resulting range of trajectories for HIV incidence in future years.

Web Figure 5: HIV Incidence Inputs.

Rate ratios describing HIV incidence for each age group as compared to the general population were calculated from the age distribution of new infections in the most recent set of estimates (76) as compared to the population age distribution over the same period. These rate ratios were assumed constant over time, and are shown in Panel B of Web Figure 5. There are imperfect data on relative HIV incidence rates in high vs. low risk groups, though differences in HIV prevalence is known to be high in the high risk group relative to the general population (38).
We assumed an incidence rate ratio of 2.0 [1.0, 3.0] for the high-risk group, which was applied to each age group separately. In combination with higher rates of entry to the high risk group for HIV infected individuals (see Section 1.14), these assumptions will produce HIV prevalence in the high-risk group approximately 5-8 times higher than the low risk group.

1.15. HIV progression

Individuals transition from early HIV to advanced HIV compartments to reproduce observed disease progression rates and the time pattern of HIV-associated mortality in treated/untreated cases.

Parameter format

A parameter ($\lambda_{HIV\_progression}^{HIV\_progression}$) is defined as the rate at which individuals in the early HIV compartments progress to the advanced HIV compartments, stratified by ART status.

Data sources

The progression rate for individuals not receiving ART was based on a synthesis of observational data on HIV progression in high-income countries prior to the widespread availability of ART (77), which reported data consistent with a mean duration from seroconversion to CD4<350 of 3.3 years, producing a transition rate of 0.30 [0.14, 0.53]. We used this rate for individuals <25 years of age and allowed for higher progression rates for older age groups (Web Figure 6), based on progressively shorter survival estimates described for older age groups (15). Progression rates for individuals on ART was set at 0.023 [0.0004, 0.088] for individuals <25 years of age in order to produce appropriate mortality rates, and assumed to increase for older age groups.

Web Figure 6: Rate of Progression from Early- to Late-Stage HIV, by Age and Receipt of ART.

References


Web Appendix 2. Routine interventions

Major interventions in the model include (i) screening of prospective migrants for active TB disease prior to arrival, (ii) testing for and treatment of latent TB infection, (iii) diagnosis and treatment of active TB, and (iv) provision of antiretroviral therapy for HIV. The first of these (migrant screening) is described in Section 1.4. While ART is not primarily provided for its impact on TB, it has an important role in preventing active TB among individuals with advanced HIV.

2.1. LTBI testing and treatment initiation

The model allows for rates of LTBI diagnosis and uptake onto LTBI treatment for those with latent TB infection, with these rates changing over time and by individual characteristics. False-positive treatment, i.e. treatment of those in the uninfected compartments, is described by calculations outside of the model structure, and not assumed to have any epidemiologic impact.

Parameter format

The model allows a smooth increase in LTBI treatment access between 1985 and 2015 (operationalized as a logistic curve), to reach a final rate of 0.025 [0.010, 0.100] for a non-US-born individual with LTBI. Differential rates of testing for other risk groups were specified via risk ratios, with risk ratios of 3.0 [1.5, 6.0] assumed for the HIV positive and homeless compartments, and 0.2 [0.1, 0.4] for individuals with no other risk factors. These parameters were calibrated to estimates of the total volume of LTBI treatment reported by Sterling et al. (1). Test characteristics were based on IGRA, with sensitivity of 0.80 [0.75, 0.84] assumed for LTBI diagnosis among HIV negative individuals, and 0.68 [0.60, 0.75] for individuals with HIV. Specificity was assumed to be 0.98 [0.97, 0.99] (2-6). For the base case projection, parameters describing LTBI treatment uptake by risk group and LTBI diagnostic performance were held constant at current values.

2.2. LTBI treatment completion, discontinuation and default

Individuals exit the LTBI treatment compartment to reflect treatment completion as well as default/discontinuation, with these mechanisms modeled separately. Individuals who are cured transition to the ‘Uninfected, partially-immune’ compartment, which individuals failing LTBI treatment return to the ‘Latent, slow’ compartment.

Parameter format

The fraction of individuals exiting LTBI treatment successfully cured was assumed to be a function of whether the full treatment course was completed, and the drug susceptibility of the treated individual’s TB strain. Cure rates for individuals defaulting from treatment were related to the cure rates for those completing treatment by a fixed odds ratio. Individuals treated with a regimen to which their strain is resistant were assumed to fail treatment.

Data sources

For individuals completing treatment, the exit rate was calculated as the inverse of the regimen duration, assuming a 9 month INH regimen (9H). Rates of default from LTBI treatment are known to be high (7-9), and programmatic data provided by the CDC indicate only 60-70% of all individuals starting the regimen complete LTBI treatment (89). For this study it was assumed that the fraction defaulting/discontinuing their regimen was 0.33 [0.29, 0.37], and the default rate was specified to match this fraction. The probability of cure for individuals completing treatment was specified as 0.90 [0.81, 0.96] (10, 11). The probability of cure for individuals defaulting from treatment was assumed to be zero, based on the observation that LTBI treatment default typically occurs shortly after treatment initiation (7, 12), and will likely have little therapeutic benefit. For individuals infected with an INH resistant strain, the probability of cure was also assumed to be zero (13). For the base case projection, parameters describing LTBI treatment completion and cure rates were held constant at current values.

2.3. TB diagnosis and treatment initiation

Individuals with active TB face rates of diagnosis and treatment initiation that vary over time and between model compartments, with slower diagnosis for the high-risk group. Newly diagnosed individuals with
pansensitive TB are assumed to be initiated the standard 1st-line regimen. Newly diagnosed individuals with drug-resistant TB are assumed to be initiated on either the standard 1st-line regimen or a regimen appropriate for their drug resistance profile, depending on the quality of diagnosis.

Parameter format

The rate at which individuals with active TB present for treatment was modeled with a penalized B-spline, operationalized with a quadratic basis function and a second order difference penalty. This produced a flexible function allowing smooth changes over time without imposing a restrictive structural form. The delay between initial attendance and treatment initiation was considered separately, and individuals with active TB were assumed to exhibit ongoing transmission risks during this period. This ‘provider delay’ was assumed to differ between smear-negative and smear-positive patients, and both the attendance rate and provider delay were assumed to differ by age and by homelessness status. Algorithm sensitivity was assumed to differ by smear-status. The probability of receiving DST was assumed to rise smoothly between 1985 and 2010 (operationalized as a logistic curve).

Data sources

There is little information on the average time delay between incidence of active TB and treatment initiation. Empirical studies to estimate this quantity collect data from individuals presenting as suspected TB cases, and are affected by recall bias (particularly since the start of clinical infection is difficult to discern). In addition, some TB cases will self-cure before reaching diagnosis, yet still contribute to transmission. For this reason, the rate of treatment initiation for individuals with active TB was identified through calibration to the case report time series, with a diffuse prior created for the attendance rate with a value of 0.25 in 1950 rising to 4.0 in 2015 (14, 15). Individuals with smear-negative disease were assumed to have a provider delay of 1.5 [1.0, 2.1] months, while a delay of 0.6 [0.4, 0.9] months was used for individuals with smear-positive disease (14, 15). The rate of attending healthcare with TB symptoms is known to be lower for homeless populations, as is the provider delay (due to inadequate retention between visits). The overall delay between developing TB symptoms and initiating TB treatment for the high-risk group was related to the same value for the general population by a ratio of 2.0 [1.5, 2.5] (14, 15). The sensitivity of TB diagnosis was assumed to be high – 0.95 [0.9, 0.99] for smear-negative TB and 0.995 [0.99, 0.999] for smear-positive TB. The probability of receiving drug-sensitivity testing (DST) was assumed to rise from zero in 1985 to 0.8 [0.7, 0.9] in 2000, based on routine reporting. In more recent years the incomplete coverage of DST results is due primarily to the inability to undertake culture (and therefor DST) for some individuals diagnosed with TB, rather than lack of access to DST itself. For the base case projection, parameters describing delays to diagnosis and treatment of active TB, and the performance of diagnostic services, were held constant at current values.

2.4. TB treatment completion, discontinuation and default

Individuals exit the TB treatment compartments to reflect treatment completion as well as default/discontinuation and death. These mechanisms were modeled separately.

Parameter format

The outcomes related to exit from TB treatment are shown in Web Figure 7. Mortality for individuals on treatment was assumed to be a function of background mortality, HIV-specific mortality, and TB-specific mortality. Background and HIV-specific mortality rates were assumed to be unaffected by TB treatment, while TB-specific mortality was calculated as the mortality rate associated with active TB multiplied by the fraction receiving ineffective treatment (1.0 - TxEffit, a parameter representing treatment effectiveness). A tuning parameter (Adj/TxMort) was also included in this equation to calibrate model predictions to observed reporting data (Web Figure 7). Treatment effectiveness (TxEffit) was assumed to be the product of regimen efficacy (specific to the strain/regimen combination) and a time-varying parameter describing overall treatment quality (TxQualt). Treatment completion rates were calculated to produce the average duration for each regimen. A single time-varying rate of default/discontinuation (r_defaultt) was adopted for the general population, and a higher rate assumed for the high-risk group.
Treatment completion and default/discontinuation were assumed to lead to one of several outcomes: (i) cure, with the individual returning to the slow latent compartment, (ii) failure/relapse, with the individual returning to the active disease compartment, or (iii) retreatment (in the case of treatment failure), with individuals reinitiated on a regimen appropriate for their drug resistance profile. Individuals reinitiated on treatment were assumed to represent a fixed fraction of all those with treatment failure after competing a treatment regimen. Cure probabilities for individuals completing treatment were determined by the parameter $TxEff_{it}$. We assumed the probability of cure for individuals defaulting from treatment represented a fixed multiple of the cure probability for individuals completing treatment. For simplicity, the acquisition of TB drug resistance was modeled as occurring as individuals exit treatment, with individuals with treatment failure assigned to a new drug resistance category based on the rates of acquired resistance associated with their strain:regimen combination.

**Web Figure 7: Approach for Modeling TB Treatment Outcomes.**

**Data sources**

**Regimen duration**

Regimen duration was assumed to be 9 months for a RIF-sensitive strain (1st line), and 21 months for a RIF-resistant strain. Individuals not receiving drug sensitivity testing were assumed to go onto the 1st line regimen.
Regimen efficacy

Regimen efficacy was assumed to be a consequence of both TB strain and regimen. For pansensitive TB treated with a 1st line regimen, efficacy was estimated at 0.96 [0.94, 0.98] (16), including both immediate failure and relapse. Efficacy for other strain:regimen combinations were indexed to the value for the 1st line regime and pansensitive TB via risk ratios. Efficacy of a first-line regimen with strains resistant to INH or RIF (but not both) was estimated with a risk ratio of 0.90 [0.82, 0.96] (16, 17), and a risk ratio of 0.50 [0.35, 0.65] was used for MDR and XDR-TB treated on the first-line regimen (17). Risk ratios for individuals treated on a 2nd line regimen were 0.90 [0.80, 0.95] for sensitive strains (18), and 0.50 [0.35, 0.65] for non-sensitive strains (the same as MDR-TB with a first-line regimen).

Treatment quality

A holistic measure of treatment quality is difficult to define, but for the purposes of this analyses it was assumed to be proportional the fraction of individuals successfully completing treatment within 12 months (among those eligible to do so), as reported by the national surveillance system (19) for the years 1993-2011. Little information is available for earlier years, and for simplicity we assumed stable treatment quality over the years 1950-1980 (with a prior of 0.70 [0.55,0.83]), and a smooth transition to the subsequent time trend. For years after 2011 we assumed treatment quality held stable at recent values. The implications of these decisions are shown in Panel A of Web Figure 8. By allowing to treatment quality in the early years to be either better or worse than in 1993, the model allows for the possibility of a transient degradation in the effectiveness of the treatment program in the late eighties and early nineties, associated with the spike in case notifications observed during that period (20).

Identification of treatment failure

It was assumed that some fraction of individuals who complete a regimen yet fail to achieve cure would be identified and reinitiated on a regimen tailored to their drug resistance profile. It was assumed that this fraction rose from zero in 1985 to 0.5 [0.1, 0.9] by 2000, operationalized as a logistic curve.

Treatment default/discontinuation

Rates of default/discontinuation were based on national treatment outcomes data (21). Under the assumption of constant mortality and default rates, and assuming an average regimen duration of 9 months, default rates were estimated from the treatment outcomes data, which describe the fraction defaulting/discontinuing for various reasons, out of the entire cohort initiating treatment. These estimates cover the period 1993-2011. A similar approach to that used for treatment quality was used for estimating values for early years, with default rates over the years 1950-1980 assumed to be stable with a prior of 0.18 [0.08,0.31]). Future values were obtained by assuming continuation of the most recent values. The implications these decisions are shown in Panel B of Web Figure 8. Default rates for the high-risk group were assumed to be 2.0 [1.2, 3.0] times higher than the general population rate. There are few studies that report outcomes for patients who have defaults or discontinued, but failure rates are thought to be high (22), and for this study it was assumed the cure probability for this group represented (0.25 [0.04, 0.64]) of the cure probability for individuals successfully completing treatment. For the base case projection, parameters describing treatment outcomes for active TB (cure, failure, default, death, acquired drug resistance) were held constant at current values.
Acquired TB drug resistance

A fraction of all individuals failing treatment were assumed to acquire drug resistance. Diffuse priors for the probability of developing drug resistance were based on the work of Dye, Espinal and Williams (23, 24) and a meta-analysis by Lew et al (25). For individuals with initially pansensitive TB failing a 1st line regimen this fraction ($p_{1}^{AR}$) was assumed to be 0.13 [0.03, 0.31], 0.08 [0.01, 0.22] and 0.07 [0.01, 0.21] for developing INH resistance, RIF resistance, and MDR-TB respectively. For individuals with mono-RIF or mono-INH resistance the probability of developing MDR-TB was assumed to be 0.49 [0.27, 0.77] if treated with the 1st line regimen and 0.13 [0.03, 0.31], and 0.08 [0.01, 0.22] respectively (i.e. the same as observed for pansensitive TB with the first line regimen), if treated with a regimen appropriate for their existing resistance pattern. For individuals with MDR-TB the probability of developing XDR-TB was assumed to be 0.20 [0.06, 0.32].

It is unlikely the probability of acquired resistance is a constant fraction of those failing a regimen—for poor quality treatment, failure to achieve cure will also be due to inadequate adherence to a regimen that would otherwise be effective. For this reason, the probabilities of acquired resistance were multiplied by a scaling function $1 - e^{-(s-TxEff_{1})} f^{AR}_{1}$, where $f^{AR}_{1}$ is a tuning parameter (with prior 2.0 [1.0, 4.0]) that allows for the probability of acquired resistance to be an increasing but concave function of the failure rate. For low values of $f^{AR}_{1}$ the rate of acquired resistance decline almost linearly as failure rates decline. For high values of $f^{AR}_{1}$ the rate of acquired resistance decline sub-linearly as failure rates decline.

2.5. ART initiation

ART initiation was modeled to track historical changes in ART eligibility criteria and changes in ART coverage.

Parameter format

Expansion of ART access was modeled as a smooth increase in the rate of ART initiation for HIV-infected individuals (operationalized as a logistic curve), stratified by patient characteristics.
For patients with CD4<350 and/or diagnosed TB, the rate of ART initiation was assumed to increase from zero in 1994 to a rate of 0.25 [0.05, 0.75] in 1997, consistent with the substantial reduction in HIV mortality over this period. ART access for patients with CD4>350 and no TB disease was assumed to be delayed, rising from zero in 2005 to a fraction (0.1 [0.05, 0.2]) of the rate for the CD4<350 category by 2015, following the progressive expansion in CD4-based ART eligibility expressed by the treatment guidelines at the time (26-28).

The prior distributions for these rates were deliberately broad, and information on reported ART coverage was included in the calibration (29). For individuals in the high-risk group, the rate of ART initiation was assumed to be a fraction (0.2 [0.1, 0.5]) of the rate on the corresponding non-high-risk compartment. For the base case projection, parameters describing ART initiation by risk group were held constant at current values.

2.6. ART default

ART patients were assumed to experience a small ongoing risk of default.

Parameter format

The risk of default was assumed to be constant in time and stratified by risk group, with an increased rate of default for the high-risk group.

Data sources

Retention in HIV care is known to be sub-optimal (30) and a default rate of 0.15 [0.05, 0.4] was used, with this rate assumed to be double for the high risk group (31).

References

**Web Appendix 3. Operationalization**

We programmed the model in C++ using the Rcpp package, version 0.12.3, and conducted analyses using R version 3.2.4. (1, 2). The distribution of the population across model compartments was updated with a monthly timestep, and transition probabilities calculated with a first-order Taylor series approximation.

The model was initiated in 1950. In order to create an appropriate population distribution at the start of 1950, the model was initiated with a small fraction of individuals in the active TB compartment and burned in by cycling the model at 1950 values until it reached equilibrium, rebalancing overall population size and distribution across risk and age groups in each cycle. The model was then used to simulate outcomes from 1950 to 2016, to reproduce time trends in TB epidemiology and the introduction of control interventions.

Partial rank correlation coefficients were calculated using the sensitivity package, version 1.12.1.

**References**

Web Appendix 4. Calibration approach

A Bayesian approach was used to calibrate the model, implemented with Incremental Mixture Importance Sampling (IMIS) (1, 2). Under this approach estimates for quantities of interest are obtained by combining information from two data sources: prior distributions for model parameters, and a likelihood function for evidence on modeled outcomes. This produces a large sample of parameter sets representing draws from the posterior parameter distribution. We used the distribution of model results to calculate mean estimates, and report uncertainty via equal-tailed 95% posterior intervals and graphical displays of the distribution of results.

We constructed prior distributions for model parameters using Beta distributions (for parameters bounded between 0 and 1), Gamma distributions (for non-negative parameters) and Normal distributions (for unbounded parameters). The mean and implied 95% confidence interval for each parameter is given in the relevant text of Sections 1 and 2.

Calibration data were compiled to describe a wide range of features of historical and current demography, TB and HIV epidemiology, and disease control efforts. The model was calibrated to the following evidence:

Demography
- Total non-US-born population in each decade 1960-2010 (3).
- Age distribution of the total population, by nativity, in 2014 (4).
- Membership of the high-risk compartment in 2010: calibration target created to produce prevalence of approx. 1% of US population.

TB Epidemiology
- IGRA positivity by age group and nativity, in 2011 (5). This was assumed to be related to latent TB infection prevalence by the values given for test sensitivity and specificity in Section 2.1.
- Total deaths with TB (ICD-10 codes A16-19, multiple cause of death reports) 1999-2014 (6).
- Distribution of TB deaths by age-group, 1999-2014 (6).
- Time pattern of active disease risk as a function of time since initial infection (7-9)\(^2\).
- Smear-positivity in the absence of treatment (10).

TB Diagnosis and Treatment
- The time series of total TB cases reported for 1953-2015 (11, 12).
- The fraction of all reported TB cases represented by non-US-born individuals, annually from 1993 to 2015 (11, 12).
- The fraction of all reported non-US-born TB cases present in the US for <2 years, annually from 1993 to 2014\(^4\) (11).
- The age distribution of reported TB cases stratified by nativity, annually from 1993 to 2014 (11).
- The fraction of all reported TB cases represented by previously treated individuals, annually from 1993 to 2014 (11).
- The fraction of all reported TB cases represented by HIV infected individuals, annually from 1993 to 2014\(^4\) (11).
- The fraction of all reported TB cases reporting homelessness in past year, annually from 1993 to 2014 (11).

\(^2\) Data from the U.S. Public Health Service INH trials (Ferebee and Mount 1962, Ferebee 1970) include co-incident TB cases among household members discovered at initial examination.

\(^3\) Final year of time series differs depending on availability of reported data for each outcome.

\(^4\) A substantial fraction of TB diagnoses in the early years of the time series have no HIV result available. To create the calibration target we calculated the fraction of TB cases HIV positive under two approaches: (i) using only the data with HIV status available (assuming availability of an HIV test result was unrelated to true HIV status), and (ii) using data on all TB cases and assuming those without reported HIV status were HIV negative. We created the calibration target for the annual fraction of TB cases with HIV as the average of these two approaches.
- The categorization of TB treatment outcomes into completed treatment, discontinued, and died, annually from 1993 to 2012 (13).
- Estimated LTBI treatment initiations and distribution across risk groups (non-US-born, HIV, homelessness, and other) in 2002 (14).

**HIV Epidemiology and Treatment**
- HIV prevalence 2006-2012 (15).
- Total HIV deaths, 2010-2013 (16).
- HIV survival in the absence of treatment, by age group (17).
- Estimated ART volume in 2010 (18).

**References**

Web Appendix 5. Modeled scenarios

We examined several hypothetical scenarios in addition to the base case. This section describes the operationalization of these scenarios.

Scenario 1: TLTBI for new migrants
This scenario modified the base case scenario to include provision of LTBI screening and treatment for all new legal migrants and refugees entering the United States. To operationalize this scenario we adjusted the estimated fraction of new migrants with LTBI. We assumed the intervention would be implemented progressively from 2016, to reach full coverage by 2021. We assumed the intervention would not apply to undocumented migrants or short-term visitors such as persons on work visas or student visas. For individuals with LTBI receiving the intervention, we estimated the fraction cured to be the product of LTBI diagnostic sensitivity, the fraction completing treatment, and the cure rate for individuals completing treatment. Parameters describing the performance of diagnosis and treatment were assumed to be the same as for existing LTBI treatment services (Section 2.1-2).

Scenario 2: Improved TLTBI in US
This scenario modified the base case scenario to allow increased uptake of LTBI screening and treatment among high-risk populations. There were two components to this intervention: greater access to LTBI treatment, and adoption of a shorter treatment regimen. To model greater access to LTBI treatment, we assumed that the rate at which individuals would be identified to receive LTBI treatment would be increased by 100% (see Section 2.1 for base case assumptions). As current treatment initiation rates differ by risk group (homelessness, HIV, non-US-born, other) we assumed that rates would be increased proportionally within each risk group. We assumed the same diagnostic approach would be used as in the current program (Section 2.1). To model adoption of a shorter isoniazid-rifapentine LTBI treatment regimen with higher completion rates, we reduced the duration of LTBI treatment from 9 months to 3 months, producing a reduction in default from 33% to 12%, and assumed treatment efficacy was the same as 9 month INH among those completing the regimen (1) (in all other scenarios, 9H is assumed to be the TLTBI regimen). Both elements of this scenario (greater access to LTBI treatment, adoption of a isoniazid-rifapentine treatment regimen) were assumed to be implemented progressively from 2016, to reach full coverage by 2021.

Scenario 3: Better case detection
This scenario modified the base case scenario to allow improved detection of active TB cases, such that the duration of untreated active disease is shortened to half the current estimated value. To model this scenario we assumed that the delay from development of TB symptoms to attendance at a health care provider (patient delay) would be reduced by 50%, which we implemented by increasing the rate for individuals with active TB by 100%. As this rate differs by risk group (Section 2.3) we assumed that rates would be increased proportionally within each risk group. Similarly, we assumed that the delay from attendance at a health care provider to treatment initiation (diagnostic delay) would also be reduced by 50% within each risk group. We assumed these changes would be implemented progressively from 2016, to reach full coverage by 2021.

Scenario 4: Better TB treatment
This scenario modified the base case scenario to allow improvements in the quality of diagnosis and treatment for individuals with active TB. This was operationalized through changes to three aspects of care (Section 2.4): (i) the probability that an individual with TB drug resistance would receive a regimen appropriate to their drug resistance profile, (ii) treatment completion, and (iii) the probability of cure for individuals completing their TB regimen (Section 2.4). For the probability that an individual would receive an appropriate regimen, we reduced the fraction of individuals without DST results by 50% from its current value, and assumed that all individuals with DST results indicative of TB drug resistance would receive an appropriate regimen. For treatment completion, we reduced the rate of treatment discontinuation or default by 50% compared to its current value. As this rate differed by risk group (homeless, other) we assumed the rate would change proportionally within each risk group. For the probability of cure, we assumed that the risk of treatment failure (for individuals completing their regimen) would be reduced by 50% within each regimen and drug resistance
category. We assumed these changes would be implemented progressively from 2016, to reach full coverage by 2021.

**Scenario 5: All improvements**

This scenario combined each of the 4 scenarios described above.

**Sensitivity analysis 1: No transmission after 2016**

For this sensitivity analysis we assumed that from 2016 onwards no individuals would acquire *M. tb* infection from transmission within the United States. In this scenario the only source of new LTBI cases was from existing infection in new migrants.

**Sensitivity analysis 2: No TB or LTBI in new migrants 2016**

For this sensitivity analysis we assumed that from 2016 onwards all individuals migrating to the United States would be free of TB and LTBI, while maintaining the same total volume of immigration (includes both legal and undocumented migrants).

**Sensitivity analysis 3: Reduced TB prevention efforts**

For this sensitivity analysis we assumed that the annual rate of screening and treatment of LTBI among target groups would be halved between 2016 and 2021, representing a weakening of US TB prevention efforts.

**Sensitivity analysis 4: Rising MDR-TB in new migrants**

For this sensitivity analysis we assumed that MDR-TB prevalence among new migrants would increase from current values to plateau at 19% and 55% among treatment naïve and previously treated individuals respectively by 2050. These values represent the current highest MDR-TB rates for any individual geographic region (2).

**Sensitivity analysis 5: Future in-migration at 2016 levels**

For this sensitivity analysis we assumed that a fixed cap would be placed on future immigration volume, restricting it to 2016 levels. While the actual pattern of a restricted immigration intervention would likely be non-linear, and may only partially restrict undocumented immigration, this scenario provides a simplified projection of the impact of restricting immigration.

**Sensitivity analysis 6: Improving global TB control**

For this sensitivity analysis we assumed that TB control in countries from which new migrants originate would see faster rates of reduction in TB burden, with rates dropping at 3% per year, double the current rate.

**Sensitivity analysis 7: Rising reactivation risks**

For this sensitivity analysis we assumed that population average reactivation risks for individuals with LTBI would rise progressively until 2050, resulting from rising diabetes prevalence, with reactivation rates increasing by a rate ratio of 1.4 by 2050 (3, 4).

**Sensitivity analysis 8: Recent transmission lower in non-US-born**

In a recent analysis of TB genotyping data, Yuen et al estimate values for the fraction of new TB cases attributable to recent infection (5). For the non-US-born this value was 7.5%, substantially lower than the mean estimate of 18% from our analysis. There are several differences in the approach adopted by Yuen and the values reported from our analysis:

- The Yuen analysis excluded non-US-born individuals resident in the US for less than 6 months, while our analysis excludes all those in the US for less than 2 year.
- Any TB resulting from recent infection acquired abroad, or from transmission from short-term non-immigrant visitors, would be coded as ‘unclustered’ in the Yuen analysis and therefore attributed to reactivation disease. In our analysis these individuals would be attributed to recent transmission.
In the Yuen analysis, the operational definition of recent transmission is based on the length of time between diagnosis date for source and secondary cases (with the recent transmission requiring a delay of between -0.25 and 2 years between source case and secondary cases). In our analysis the definition of recent transmission is based on date of incidence of active TB (as opposed to date of TB diagnosis). This value is calculated as the fraction of all incident cases of TB resulting from infection within the last 2 years, on expectation, averaged across model compartments.

Despite these definitional differences, the recent transmission values from these two analyses are sufficiently comparable to warrant further investigation. For this sensitivity analysis we modified the model to be consistent with the Yuen estimates for the fraction of diagnosed TB cases attributable to recent infection among non-US-born individuals. To do so we included an additional calibration target for the fraction of TB among non-US-born individuals due to recent transmission, centered at the Yuen estimate of 7.5%. With this new calibration target we fit the model and estimated outcomes as described for the main analysis.

**Sensitivity analysis 9: Lower LTBI prevalence.**

Currently available tests for LTBI have imperfect sensitivity and specificity. For the main analysis we explicitly model these factors to estimate IGRA positivity, which can be compared directly to empirical estimates produced by analysis of NHANES data. In this sensitivity analysis we adopted alternative assumptions whereby ‘dual-positive’ results (positivity to both IGRA and TST tests) were assumed to be equivalent to LTBI infection. The result of these alternative assumptions was lower LTBI prevalence estimates in the recalibrated model.

**Sensitivity analysis 10: LTBI reactivation risk plateaus after initial decline**

In the model used for the main analysis, LTBI progression risks are allowed to decline monotonically as a function of years since infection, such (for example) that the progression risk 30 years after infection would be substantially lower than the risk faces 10 years after infection. This approach contrasts with some earlier TB models, which assumed that temporal declines in LTBI reactivation risk following *M. tb* infection eventually plateau 

(6, 7). For this sensitivity analysis we recalibrated the model to assess outcomes under this alternate modeling approach.

**References**

Web Figure 9: Comparison of Calibration Data and Results from Fitted Model.
Web Figure 10: Comparison of Calibration Data and Results from Fitted Model.
Web Figure 11: Comparison of Calibration Data and Results from Fitted Model.
Web Figure 12: Comparison of Calibration Data and Results from Fitted Model.
Web Figure 13: Comparison of Calibration Targets and Results from Fitted Model.
Web Appendix 7. Additional results

Web Figure 14: Epidemiology of Recent and Current TB Transmission: Modeled Estimates for Fraction of Incident TB Cases due to Recent Infection, and Annual Rate of TB Infection*.

* Bold lines represent best estimate projection (posterior mean for a given population group and projection year). Faint lines represent 100 individual calibrated parameter sets, representing uncertainty in modeled results.
Web Figure 15: Projected Trends in Tuberculosis Outcomes Under Modeled Scenarios, 2015-2050, Stratified by Nativity.
### Web Table 1: Measures of TB Epidemiology Under Various Hypothetical Scenarios Projected to 2025, 2050, and 2100, Total US Population.

* Values for base-case scenario shown in Table 1. Point estimate represents posterior mean. Values in parentheses represent equal-tailed 95% posterior intervals.
<table>
<thead>
<tr>
<th>Scenario</th>
<th>TLTBI for new Migrants</th>
<th>Improved TLTBI in US</th>
<th>Better Case Detection</th>
<th>Better TB Treatment</th>
<th>All Improvements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcomes at 2025</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New infections per million</td>
<td>38 (31, 47)</td>
<td>37 (30, 45)</td>
<td>18 (15, 22)</td>
<td>41 (33, 50)</td>
<td>12 (10, 15)</td>
</tr>
<tr>
<td>Percent of base-case*</td>
<td>88 (86, 91)</td>
<td>85 (83, 88)</td>
<td>42 (38, 44)</td>
<td>95 (94, 95)</td>
<td>29 (27, 31)</td>
</tr>
<tr>
<td>LTBI prevalence (%)</td>
<td>1.2 (0.99, 1.5)</td>
<td>1.2 (0.94, 1.4)</td>
<td>1.2 (0.98, 1.5)</td>
<td>1.2 (0.99, 1.5)</td>
<td>1.1 (0.92, 1.4)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>100 (100, 100)</td>
<td>95 (93, 96)</td>
<td>99 (98, 99)</td>
<td>100 (100, 100)</td>
<td>94 (92, 95)</td>
</tr>
<tr>
<td>TB cases per million</td>
<td>7.4 (6.2, 8.8)</td>
<td>6.4 (5.3, 7.8)</td>
<td>5.6 (4.7, 6.6)</td>
<td>7.2 (6.1, 8.6)</td>
<td>4.2 (3.5, 5.2)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>95 (94, 97)</td>
<td>83 (80, 86)</td>
<td>72 (68, 76)</td>
<td>94 (93, 95)</td>
<td>55 (51, 59)</td>
</tr>
<tr>
<td>MDR-TB in incid TB (%)</td>
<td>0.94 (0.67, 1.3)</td>
<td>1.1 (0.75, 1.5)</td>
<td>1.1 (0.79, 1.6)</td>
<td>0.79 (0.54, 1.1)</td>
<td>1.1 (0.76, 1.7)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>100 (100, 101)</td>
<td>114 (111, 118)</td>
<td>119 (115, 124)</td>
<td>84 (80, 87)</td>
<td>120 (111, 130)</td>
</tr>
<tr>
<td>TB deaths per million</td>
<td>0.83 (0.64, 1.0)</td>
<td>0.70 (0.53, 0.90)</td>
<td>0.57 (0.42, 0.73)</td>
<td>0.77 (0.59, 0.99)</td>
<td>0.43 (0.32, 0.57)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>98 (97, 98)</td>
<td>84 (81, 87)</td>
<td>68 (64, 71)</td>
<td>93 (92, 94)</td>
<td>52 (48, 56)</td>
</tr>
<tr>
<td><strong>Outcomes at 2050</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New infections per million</td>
<td>18 (13, 26)</td>
<td>18 (11, 26)</td>
<td>9.9 (6.7, 14)</td>
<td>23 (15, 32)</td>
<td>5.2 (3.4, 7.6)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>76 (73, 81)</td>
<td>72 (66, 79)</td>
<td>41 (35, 45)</td>
<td>93 (91, 95)</td>
<td>21 (18, 24)</td>
</tr>
<tr>
<td>LTBI prevalence (%)</td>
<td>0.42 (0.35, 0.51)</td>
<td>0.39 (0.32, 0.47)</td>
<td>0.40 (0.33, 0.48)</td>
<td>0.43 (0.35, 0.52)</td>
<td>0.36 (0.29, 0.44)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>98 (97, 98)</td>
<td>89 (87, 92)</td>
<td>92 (90, 94)</td>
<td>99 (99, 99)</td>
<td>83 (79, 87)</td>
</tr>
<tr>
<td>TB cases per million</td>
<td>2.6 (2.0, 3.4)</td>
<td>2.1 (1.5, 2.9)</td>
<td>1.7 (1.4, 2.2)</td>
<td>2.9 (2.2, 3.7)</td>
<td>0.94 (0.72, 1.2)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>84 (80, 88)</td>
<td>66 (61, 71)</td>
<td>56 (49, 61)</td>
<td>91 (89, 93)</td>
<td>30 (26, 33)</td>
</tr>
<tr>
<td>MDR-TB in incid TB (%)</td>
<td>1.1 (0.73, 1.5)</td>
<td>1.3 (0.90, 2.0)</td>
<td>1.3 (0.91, 1.9)</td>
<td>0.84 (0.56, 1.2)</td>
<td>1.7 (1.1, 2.6)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>106 (102, 109)</td>
<td>133 (125, 143)</td>
<td>130 (120, 143)</td>
<td>83 (80, 87)</td>
<td>172 (149, 200)</td>
</tr>
<tr>
<td>TB deaths per million</td>
<td>0.29 (0.22, 0.39)</td>
<td>0.21 (0.15, 0.29)</td>
<td>0.18 (0.13, 0.24)</td>
<td>0.30 (0.22, 0.40)</td>
<td>0.10 (0.071, 0.13)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>89 (86, 93)</td>
<td>64 (61, 68)</td>
<td>54 (48, 60)</td>
<td>90 (88, 92)</td>
<td>30 (26, 34)</td>
</tr>
<tr>
<td><strong>Outcomes at 2100</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New infections per million</td>
<td>8.2 (2.9, 16)</td>
<td>8.3 (2.5, 18)</td>
<td>5.0 (1.5, 10)</td>
<td>11 (3.8, 22)</td>
<td>2.4 (0.72, 5.4)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>70 (66, 79)</td>
<td>69 (54, 84)</td>
<td>42 (31, 47)</td>
<td>93 (87, 96)</td>
<td>20 (14, 26)</td>
</tr>
<tr>
<td>LTBI prevalence (%)</td>
<td>0.065 (0.049, 0.089)</td>
<td>0.060 (0.044, 0.084)</td>
<td>0.051 (0.039, 0.067)</td>
<td>0.075 (0.053, 0.087)</td>
<td>0.039 (0.031, 0.050)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>84 (79, 90)</td>
<td>77 (73, 82)</td>
<td>66 (59, 74)</td>
<td>97 (95, 98)</td>
<td>51 (42, 62)</td>
</tr>
<tr>
<td>TB cases per million</td>
<td>0.86 (0.37, 1.6)</td>
<td>0.76 (0.26, 1.6)</td>
<td>0.52 (0.20, 1.0)</td>
<td>1.07 (0.43, 2.1)</td>
<td>0.23 (0.090, 0.48)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>72 (67, 82)</td>
<td>62 (48, 75)</td>
<td>44 (32, 49)</td>
<td>89 (83, 93)</td>
<td>19 (14, 24)</td>
</tr>
<tr>
<td>MDR-TB in incid TB (%)</td>
<td>1.0 (0.49, 1.9)</td>
<td>1.2 (0.55, 2.2)</td>
<td>1.1 (0.53, 2.0)</td>
<td>0.72 (0.30, 1.5)</td>
<td>1.4 (0.64, 2.5)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>110 (103, 116)</td>
<td>131 (117, 154)</td>
<td>122 (110, 152)</td>
<td>78 (69, 85)</td>
<td>153 (122, 213)</td>
</tr>
<tr>
<td>TB deaths per million</td>
<td>0.084 (0.040, 0.15)</td>
<td>0.068 (0.027, 0.14)</td>
<td>0.041 (0.019, 0.077)</td>
<td>0.098 (0.044, 0.18)</td>
<td>0.018 (0.009, 0.035)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>75 (70, 85)</td>
<td>59 (47, 71)</td>
<td>37 (26, 43)</td>
<td>87 (81, 91)</td>
<td>16 (12, 20)</td>
</tr>
</tbody>
</table>

**Web Table 2: Measures of TB Epidemiology Under Various Hypothetical Scenarios Projected to 2025, 2050, and 2100, US-Born Population.**

* Values for base-case scenario shown in Table 1. Point estimate represents posterior mean. Values in parentheses represent equal-tailed 95% posterior intervals.
### Table 3: Measures of TB Epidemiology Under Various Hypothetical Scenarios Projected to 2025, 2050, and 2100, Non-US-Born Population.

* Values for base-case scenario shown in Table 1. Point estimate represents posterior mean. Values in parentheses represent equal-tailed 95% posterior intervals.

Web Figure 16: Correlations Between Determinants of Long-Term TB Outcomes and New Infections Per Million, TB Cases Per Million, and TB Deaths Per Million in 2050*.

* Outcomes estimated under the base case scenario. Bars and values show partial rank correlation coefficients. Negative associations between the factor and the outcome are marked in purple. Coefficients of greater magnitude indicate a greater sensitivity of the outcome to the factor, controlling for changes in other factors. For factors relating to immigrants, this group represents all individuals entering the US, including both legal and undocumented immigrants.
Web Figure 17: Projected Trends in Measures of TB Epidemiology 2015-2050, Assuming no Future TB Transmission Within the US or No Future TB or LTBI Among Immigrants Entering the US (Sensitivity Analysis).
### Web Table 4: Measures of TB Epidemiology Projected to 2025, 2050, and 2100, Assuming No Future TB Transmission Within the US or No Future TB or LTBI Among Immigrants Entering the US (Sensitivity Analysis).

* Values for base-case scenario shown in Table 1. Point estimate represents posterior mean. Values in parentheses represent equal-tailed 95% posterior intervals.
Web Figure 18: Projected Trends in Measures of TB Epidemiology 2015-2050, Under Alternative Assumptions About Future Trends in Key Epidemiologic Determinants (Sensitivity Analysis).
**Web Table 5: Measures of TB Epidemiology Projected to 2025, 2050, and 2100, Under Alternative Assumptions About Future Trends in Key Epidemiologic Determinants (Sensitivity Analysis).**

* Values for base-case scenario shown in Table 1. Point estimate represents posterior mean. Values in parentheses represent equal-tailed 95% posterior intervals.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Reduced TB Prevention Efforts</th>
<th>Rising MDR-TB in Migrants</th>
<th>Future Immigration at 2016 Levels</th>
<th>Improving Global TB Control</th>
<th>Rising Reactivation Risks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcomes at 2025</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New infections per million</td>
<td>57 (48, 68)</td>
<td>53 (44, 64)</td>
<td>52 (44, 63)</td>
<td>52 (43, 62)</td>
<td>57 (48, 69)</td>
</tr>
<tr>
<td>Percent of base-case*</td>
<td>106 (104, 108)</td>
<td>99 (98, 100)</td>
<td>98 (96, 99)</td>
<td>97 (96, 98)</td>
<td>107 (106, 108)</td>
</tr>
<tr>
<td>LTBI prevalence (%)</td>
<td>3.2 (2.8, 3.6)</td>
<td>3.1 (2.8, 3.5)</td>
<td>3.1 (2.7, 3.5)</td>
<td>3.1 (2.7, 3.5)</td>
<td>3.1 (2.8, 3.5)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>102 (102, 103)</td>
<td>100 (100, 100)</td>
<td>99 (99, 100)</td>
<td>99 (98, 99)</td>
<td>100 (100, 100)</td>
</tr>
<tr>
<td>TB cases per million</td>
<td>25 (22, 28)</td>
<td>24 (21, 27)</td>
<td>23 (20, 26)</td>
<td>23 (20, 26)</td>
<td>25 (22, 29)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>106 (104, 107)</td>
<td>100 (99, 101)</td>
<td>97 (95, 99)</td>
<td>95 (94, 96)</td>
<td>107 (106, 109)</td>
</tr>
<tr>
<td>MDR-TB in incid TB (%)</td>
<td>1.6 (1.4, 1.9)</td>
<td>3.4 (2.9, 4.1)</td>
<td>1.6 (1.4, 1.9)</td>
<td>1.6 (1.4, 1.9)</td>
<td>1.6 (1.4, 1.9)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>97 (95, 98)</td>
<td>207 (194, 220)</td>
<td>99 (97, 100)</td>
<td>98 (97, 100)</td>
<td>96 (95, 98)</td>
</tr>
<tr>
<td>TB deaths per million</td>
<td>2.2 (1.7, 2.8)</td>
<td>2.1 (1.6, 2.6)</td>
<td>2.0 (1.6, 2.6)</td>
<td>2.0 (1.6, 2.6)</td>
<td>2.3 (1.7, 2.9)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>106 (105, 108)</td>
<td>101 (100, 101)</td>
<td>99 (97, 100)</td>
<td>98 (97, 98)</td>
<td>109 (108, 109)</td>
</tr>
<tr>
<td><strong>Outcomes at 2050</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New infections per million</td>
<td>38 (25, 52)</td>
<td>32 (20, 44)</td>
<td>30 (19, 41)</td>
<td>27 (17, 38)</td>
<td>42 (28, 58)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>115 (101, 131)</td>
<td>96 (85, 108)</td>
<td>90 (78, 102)</td>
<td>82 (72, 93)</td>
<td>126 (110, 144)</td>
</tr>
<tr>
<td>LTBI prevalence (%)</td>
<td>2.0 (1.5, 2.5)</td>
<td>1.9 (1.4, 2.3)</td>
<td>1.7 (1.3, 2.2)</td>
<td>1.6 (1.3, 2.0)</td>
<td>1.9 (1.4, 2.4)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>107 (100, 115)</td>
<td>100 (93, 107)</td>
<td>94 (86, 101)</td>
<td>86 (79, 92)</td>
<td>101 (94, 108)</td>
</tr>
<tr>
<td>TB cases per million</td>
<td>16 (11, 22)</td>
<td>14 (10, 20)</td>
<td>13 (8.6, 17)</td>
<td>11 (7.4, 15)</td>
<td>18 (12, 25)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>114 (100, 130)</td>
<td>101 (89, 115)</td>
<td>88 (75, 102)</td>
<td>75 (65, 85)</td>
<td>126 (109, 144)</td>
</tr>
<tr>
<td>MDR-TB in incid TB (%)</td>
<td>1.7 (1.1, 2.5)</td>
<td>11 (5.9, 17)</td>
<td>1.9 (1.2, 2.7)</td>
<td>2.0 (1.2, 2.8)</td>
<td>1.7 (1.0, 2.4)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>91 (79, 104)</td>
<td>564 (456, 676)</td>
<td>99 (87, 113)</td>
<td>103 (90, 118)</td>
<td>87 (76, 100)</td>
</tr>
<tr>
<td>TB deaths per million</td>
<td>1.5 (1.1, 2.0)</td>
<td>1.3 (0.91, 1.8)</td>
<td>1.2 (0.82, 1.6)</td>
<td>1.0 (0.74, 1.4)</td>
<td>1.7 (1.2, 2.3)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>120 (109, 130)</td>
<td>104 (96, 114)</td>
<td>92 (83, 101)</td>
<td>82 (75, 89)</td>
<td>132 (121, 144)</td>
</tr>
<tr>
<td><strong>Outcomes at 2100</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New infections per million</td>
<td>20 (7.2, 41)</td>
<td>16 (5.4, 35)</td>
<td>14 (4.7, 29)</td>
<td>10 (3.3, 23)</td>
<td>22 (8.0, 44)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>119 (97, 143)</td>
<td>96 (80, 111)</td>
<td>83 (62, 111)</td>
<td>58 (45, 76)</td>
<td>131 (102, 163)</td>
</tr>
<tr>
<td>LTBI prevalence (%)</td>
<td>1.0 (0.32, 2.0)</td>
<td>0.92 (0.29, 1.9)</td>
<td>0.76 (0.26, 1.5)</td>
<td>0.45 (0.19, 0.83)</td>
<td>0.93 (0.30, 1.9)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>110 (104, 115)</td>
<td>100 (95, 105)</td>
<td>84 (68, 105)</td>
<td>51 (43, 67)</td>
<td>102 (97, 107)</td>
</tr>
<tr>
<td>TB cases per million</td>
<td>8.5 (2.9, 17)</td>
<td>7.3 (2.3, 15)</td>
<td>5.8 (2.0, 11)</td>
<td>3.0 (1.1, 6.0)</td>
<td>9.2 (3.1, 19)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>118 (96, 140)</td>
<td>101 (84, 120)</td>
<td>82 (60, 113)</td>
<td>42 (34, 51)</td>
<td>129 (101, 153)</td>
</tr>
<tr>
<td>MDR-TB in incid TB (%)</td>
<td>1.9 (0.48, 4.1)</td>
<td>12 (2.4, 28)</td>
<td>2.1 (0.56, 4.5)</td>
<td>2.9 (0.76, 6.4)</td>
<td>1.8 (0.48, 4.1)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>89 (74, 107)</td>
<td>573 (400, 722)</td>
<td>101 (85, 119)</td>
<td>141 (116, 171)</td>
<td>87 (71, 104)</td>
</tr>
<tr>
<td>TB deaths per million</td>
<td>0.86 (0.26, 1.6)</td>
<td>0.72 (0.26, 1.4)</td>
<td>0.56 (0.21, 1.0)</td>
<td>0.33 (0.14, 0.61)</td>
<td>0.92 (0.34, 1.8)</td>
</tr>
<tr>
<td>Percent of base-case</td>
<td>127 (111, 142)</td>
<td>105 (96, 120)</td>
<td>83 (65, 108)</td>
<td>49 (43, 60)</td>
<td>135 (116, 150)</td>
</tr>
</tbody>
</table>
Web Figure 19: Percentage Reduction in TB incidence Produced by Each Intervention Scenario in 2050 with Different Assumptions About Current Epidemiology (Sensitivity Analysis)*.

* ‘Original model’ represents results from main analysis. ‘Recent transmission lower in non-US-born’ represents model calibrated to Yuen et al estimate of 7.5% of all transmission among non-US-born individuals due to recent transmission. ‘Lower LTBI prevalence’ represents model calibrated to lower LTBI prevalence estimates from NHANES 2011, based on a more specific positivity criterion (positive to both TST and IGRA). ‘LTBI reactivation risk plateaus after initial decline’ represents model structure revised to assume that temporal declines in LTBI reactivation risk following M. tb infection eventually plateau, an assumption made by earlier TB models.