HHS Public Access

Author manuscript

*Epidemiol Infect*. Author manuscript; available in PMC 2019 March 01.

Published in final edited form as:

*Epidemiol Infect*. 2018 September ; 146(12): 1478–1494. doi:10.1017/S0950268818001760.

# Quantifying TB transmission: a systematic review of reproduction number and serial interval estimates for tuberculosis

**Y. Ma**[*,1], **C. R. Horsburgh Jr.**[2], **L. F. White**[1,#], and **H. E. Jenkins**[1,#]

[1]Department of Biostatistics, Boston University School of Public Health, Boston, MA, 02118

[2]Department of Epidemiology, Boston University School of Public Health and Department of Medicine, Boston University School of Medicine, Boston, MA, 02118

## Summary

Tuberculosis (TB) is the leading global infectious cause of death. Understanding TB transmission is critical to creating policies and monitoring the disease with the end goal of TB elimination. To our knowledge, there has been no systematic review of key transmission parameters for TB. We carried out a systematic review of the published literature to identify studies estimating either of the two key TB transmission parameters: the serial interval and the reproductive number. We identified five publications that estimated the serial interval and 56 publications that estimated the reproductive number. The serial interval estimates from four studies were: 0.57, 1.42, 1.44 and 1.65 years; the fifth paper presented age-specific estimates ranging from 20-30 years (for infants <1 year old) to less than five years (for adults). The reproductive number estimates ranged from 0.24 in the Netherlands (during 1933-2007) to 4.3 in China in 2012. We found a limited number of publications and many high TB burden settings were not represented. Certain features of TB dynamics, such as slow transmission, complicated parameter estimation, require novel methods. Additional efforts to estimate these parameters for TB are needed so that we can monitor and evaluate interventions designed to achieve TB elimination.

## INTRODUCTION

Tuberculosis (TB), an airborne bacterial infection caused by the organism *Mycobacterium tuberculosis* (*Mtb*), has surpassed HIV/AIDS as the leading cause of death due to a single infectious organism worldwide [1]. It primarily attacks the lungs but can also infect other areas of the body [2,3]. Those exposed to *Mtb* often develop latent TB infection (LTBI) and have a 5-10% lifetime risk of progressing to active TB [4,5]. Worldwide, 2-3 billion people

[*]Corresponding author: Y. Ma, Department of Biostatistics, 801 Massachusetts Avenue, Boston, MA 02118, USA. ym48@bu.edu. Telephone: +1-646-298-8151.
[#]LFW and HEJ contributed equally to the work

**Conflict of interest:**
None.

**Ethical standards:**
The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

are infected with TB; an estimated 10.4 million people developed active TB disease in 2015 [4]. Major innovations in strategies and tools to monitor the success of new strategies are needed to achieve the World Health Organization (WHO)'s ENDTB goals of reducing TB deaths by 95% and new cases by 90% by 2035 [4].

The reproductive number and serial interval (SI) are two key quantities in describing transmission of an infectious disease. The reproductive number is defined as the average number of secondary cases a primary infectious case will produce. In a totally susceptible population, it is referred to as the basic reproductive number ($R_0$); it is referred to as the effective reproductive number ($R_e$) if the population includes both susceptible and non-susceptible persons [6]. An $R_e$ greater than 1 indicates that the disease will continue to spread while an $R_e$ less than 1 indicates that the disease will eventually die out. Although the reproductive number is usually defined as the average number of secondary cases, it is occasionally defined as the average number of secondary infections [7–10], a distinction that is important for a disease with a long incubation period (the time between infection and developing symptomatic disease) and/or only a fraction of infections progressing to disease. Depending on the setting, the reproductive number can be expressed as a function of parameters such as infection rate, contact rate, recovery rate, making it useful in determining whether or not a disease can spread through a population.

The serial interval (SI), defined as the time between disease symptom onset of a case and that of its infector [11], is a surrogate for the generation interval— an unobservable quantity defined as the time between the infection of a case and the time of infection of its infector [12]. The SI is an important quantity in the interpretation of infectious disease surveillance data, in the identification of outbreaks, and in the optimization of quarantine and contact tracing.

These two quantities have been used to inform control policies during outbreaks [13] by quantifying the transmission of infectious diseases such as influenza A (H1N1) [11,12,14], Severe Acute Respiratory Syndrome (SARS) [12,15], and Ebola [16,17], where progression to disease upon transmission occurs quickly. For example, Wallinga and Teunis [18] in 2004 demonstrated the impact of the first global alert against SARS on the change of the effective reproductive number.

TB has a slower transmission rate due to its much longer incubation period. Of the 5-10% of infections that develop into active (symptomatic and infectious) TB disease, it is thought that the majority occur within the first two years after infection [2,5,19], although active TB disease can develop decades after initial infection [20]. This is much longer than the aforementioned infectious diseases where cases show symptoms within days of infection. Although there is an increasing consensus that some transmission events may occur before the infector shows symptoms, many likely occur after the infector is symptomatic, therefore, the longer the incubation period is, the longer the SI (Figure 1).

Development of TB disease can be caused by de novo infection, reactivation of the same bacterial strain as a previous infection [5,21] or by infection with a bacterial strain different from the original infection (reinfection TB). This complicates estimation of the serial

interval, unless molecular techniques are used to distinguish reinfection and reactivation [21]. To our knowledge, there has been no systematic review of methods to estimate the serial interval and reproductive number for TB. Therefore, in this paper we systematically review the literature to examine the methods applied to the estimation of TB transmission parameters and the estimates obtained from these methods. This compilation informs the gaps in our understanding of TB and identifies areas where further research is needed to develop methods to better understand TB transmission.

## METHODS

We conducted two searches in PubMed for publications in English—one for tuberculosis and serial interval; one for tuberculosis and reproductive number:

### 1. Tuberculosis and serial interval

("Tuberculosis"[MeSH] OR "Mycobacterium tuberculosis"[MeSH] OR "tuberculosis"[TI]) and ("serial interval"[tiab] or "generation interval"[tiab] or "serial distribution" [tiab] or "secondary infections" [tiab] or "secondary cases" [tiab])

### 2. Tuberculosis and reproductive number

("Tuberculosis"[MeSH] OR "Mycobacterium tuberculosis"[MeSH] OR "tuberculosis"[TI] OR "pulmonary, tuberculosis [MeSH]") and ("reproductive number"[tiab] or "reproduction number"[tiab] or "reproductive rate"[tiab] or "reproduction rate"[tiab] or "reproduction ratio" [tiab] or "reproductive ratio"[tiab] or "reproduction value"[tiab] or "reproductive value"[tiab] or "R0" [tiab] or "secondary infections" [tiab] or "secondary cases" [tiab])

Titles and abstracts of the publications referenced in the articles we found were reviewed for inclusion for either parameter. For the SI, as limited number of publications met our inclusion criteria, we also reviewed the titles and abstracts of publications that cited the serial interval articles that we included in a full-text review.

Two reviewers (two of YM, HEJ, LFW) independently screened all titles and abstracts, resolving discrepancies by consensus. Each publication was then independently reviewed by two reviewers (two of YM, HEJ, LFW) for inclusion. From the included articles, the same pairs of reviewers extracted the following details for all parameter estimates (if available): point estimates, confidence intervals, ranges, sample size and location/setting. We summarized the methods for analysis and aggregated those with similar estimation approaches.

## RESULTS

### Serial interval

The serial interval query returned 171 articles (Figure 2), of which 163 were excluded as they did not present any estimates.. Leung et al. [22] reported the serial interval as the time from identification of primary case to secondary case as median 1.4 years (range: 0.4-5.2 years). This study used household transmission data from Hong Kong and focused on MDR- and XDR-TB. Vynnycky and Fine [23] analyzed a population of white males in England and

Wales in the 20th century using a mathematical compartmental model to estimate the SI as dependent on the age when infection occurred, distinguishing reinfection and reactivation in the model. In this model, the risk of developing disease was calibrated on incidence data. The estimates were presented as a frequency distribution. The most frequent time to develop disease was estimated at: between 20 to 30 years due to reinfection for those infected in the first year of life; between 10 and 14 years due to reinfection for those infected at age 10; less than 5 years due to recent infection for those infected at age 20 and those infected at age 40. ten Asbroek et al. [24] analyzed genetic data for a Dutch sample from 1993-1996 to link infectors and infected people using DNA fingerprinting based on restriction fragment length polymorphism (RFLP) and estimated the serial interval at a geometric mean of 0.57 years (95% Confidence Interval (CI): 0.44-0.73). In this 4-year study, the probability of observing both the infector and the infected person depended on the time interval between isolates—the shorter this time interval was, the more likely that this couple was observed. Therefore, the observed serial intervals were weighted by the inverse of the difference between the length of the study period and the time between isolates of the infector and the infected person, allowing a rough correction for underrepresentation of longer SIs.

Two articles that cited the articles that met our inclusion criteria in the PubMed search reported estimates of the SI and were included for full-text review. Borgdorff in [25] used the same method on genetic data as [24] to estimate the median SI as 1.44 years (95% CI 1.29-1.63 years) for a Dutch sample from 1993-2007. Brooks-Pollock [26] in 2011 analyzed cross-sectional household data for a sample in Lima, Peru from 1996-2002 and reported the time between the diagnosis of the infector and the infected person as an estimate for the SI with mean at 3.5 years and the median at 1.65 years.

## Reproductive number

Two hundred and thirty-seven articles were identified for the reproductive number of TB. Additionally, 6 articles were included based on reviewing titles and abstracts of the articles that were referenced in the 237 articles, making the total number of articles 243. Fifty-six articles met our inclusion criteria and are described below. Three articles used either approximate Bayesian or exact likelihood methods, 24 articles used either a mathematical model fit with empirical data or a descriptive/regression approach on empirical data, and 29 articles used a simulation based mathematical model (Figure 3). Explicit estimates were extracted and summarized in Figure 4. The estimates range from as low as 0.26 for the Netherlands in 1993-2007 to as high as 4.3 in China in 2012.

Three articles (Table 2) used the same genetic RFLP data from TB diseased individuals during an outbreak in San Francisco in 1991-1992 [27]. They all estimated the effective reproductive number in a Bayesian framework. Tanaka et al. [28] used an approximated computation method to obtain an estimate of 3.4 (95% CI: 1.4-79.7). Stadler [29] in 2013 used an exact likelihood method to obtain an estimate of 1.02 (95% CI: 1.01-1.04) and claimed that the difference from the estimate in [28] was due to the lack of precision in the approximation of the posterior distribution in [28]. Aandahl et al. [30] in 2014 reconciled the two methods by specifying an informative prior for two parameters in [28] and improving the convergence performance of the Markov chain Monte Carlo (MCMC) sampler in [29].

The reconciled estimates were: 2.1 (95% CI: 1.54-2.66) for the approximate method in [28] and 2.05 (95% CI: 1.55-2.63) for the exact method in [29]. These papers used the same model but differed in the methods used to obtain the estimates. The assumptions of the model are listed in Table 2.

Twenty-four articles analyzed the reproductive number with empirical data (Table 3). Seventeen articles reported explicit estimates, with five estimating the effective reproductive number and twelve estimating the basic reproductive number. The majority of these articles used mathematical compartmental models with different variations in structure and parameterization to address issues such as seasonality [31], the effect of age [32,33], and HIV-TB co-epidemics [9].

Two articles [7,34] used the Wells-Riley model or a modified version of the model. In these models, the reproductive number was expressed as a function of infection risk, which was further expressed as proportionate to environmental factors such as the number of infectious people in a given space, per-person breathing rate and inversely proportionate to germ-free ventilation rate. One article derived the reproductive number as a function of the transmission index— defined as the ratio of the number of secondary cases to the sum of the number of source cases (infectors) and non-clustered cases where clusters are defined as groups of patients that had isolates with identical fingerprints [35]. The largest reproductive number (effective) was estimated in [36] using the Chinese Centre for Disease Control and Prevention (CDC) data from 2005-2012, where the annual reproductive number ranged between 3.33 and 4.30 for years 2005-2012 in China. The lowest (effective) reproductive number was estimated at 0.24 (95% CI:0.17-0.31) using RFLP data in San Francisco, USA from 1991 to 1996 [37]. Vynnycky and Fine [33] in 1998 used an age-structured mathematical model and estimated the effective reproductive number to be around 1 from 1900-1950 in England and Wales; the basic reproductive number was estimated to have declined from about 3 in 1900 to 2 by 1950, and first fell below 1 in about 1960. The assumptions of these models are listed in Table 3.

One article defined the reproductive number as the number of secondary infections caused by an infectious case [7]. As only a fraction of the infected people develop active disease, the estimated reproductive number was larger than those in the other papers. The median of the reproductive number in this article ranged from 14 to 45 as exposure time increased from 1 month to 5 months.

Twenty-nine articles analyzed the reproductive number through simulation based on a mathematical modeling framework (Table 4). These articles all used mathematical compartmental models with different variations to address issues such as reinfection [38], the interaction between HIV and TB [39], and drug-resistant and drug-sensitive TB [40]. The majority of them focused on studying the effect of these issues on TB transmission dynamics through simulations that were not based on a specific population. In this case, parameters for the model were based on estimates from studies performed in diverse settings or sampled over a range of feasible values. The analytical expression of the basic reproductive number was derived to study the disease-free equilibrium and endemic-persistent state of TB in these papers. Five articles [10, 39, 40, 79, 82] included drug-

resistant TB cases as a compartment and four articles [38, 67, 78, 81] included HIV+ TB cases as a compartment.

## Discussion

We found very few publications that reported estimates for the serial interval of TB. Estimates of the reproductive number were limited to seven countries, with the majority of the publications using mathematical compartmental models that did not base estimates on actual data. This indicates a need for a better understanding of these crucial parameters of TB transmission, which can help inform public health decisions in order to reach the WHO's End TB goals [4] of reducing TB deaths by 95% and incident cases by 90% by 2035.

### Serial interval

We found only five articles that discussed the estimation of the SI for TB and presented explicit estimates. ten Asbroek [24] estimated the serial interval over four years as a geometric mean of 0.57 years (95% CI:0.44-0.73). Using the same method over a longer study period (15 years compared to four years in [24]), the estimated median was 1.44 years, which is comparable to the median serial interval of 1.65 years in [26] with a six-year study period. This indicates that the study period could potentially bias the serial interval estimates, even though the method in [24] corrected for the underrepresentation of longer serial intervals. In contrast with other infectious diseases that progress much faster and have SIs measured in days, the SI of TB can be weeks, years and even decades [23]. This unique feature of TB makes it difficult to obtain an unbiased estimate of the SI as lengthy follow-up is required to observe the long period between presence of symptoms of the infector and the infected person. Additionally, uncertainty regarding the presence and impact of multiple infection events further complicates the observation of this interval. Currently, the most common way of monitoring TB is by looking at annual incidence rates in studies that are often no longer than five years [41,42]. This creates two issues: right censoring as symptoms of the infected people can develop long after the end of studies, and interval censoring as the symptom onset time can fall during long intervals between two observed time points. Another issue is patients' and doctors' delay. Patients may not seek medical assistance immediately after symptoms develop and diagnosis may require lab processing time which causes delay in establishing the diagnosis [24], creating a left censoring issue. Survival analysis techniques can be considered to address these issues but may need substantial modification. Further ambiguity exists due to the inconsistent availability of genetic typing of strains to link cases, and the further uncertainty about how to best link strains when genetic information is available, as such information may not account for mutation rate, or infection with multiple bacterial strains.

### Reproductive number

The majority of the articles used mathematical compartmental models (a brief introduction can be found in the appendix) to describe the transmission dynamics of TB. These models have been widely used to understand the dynamics of infectious diseases including SARS, influenza and TB, and they either use empirical data to estimate the parameters in the model or are based on simulation.

The compartmental models using empirical data are distinguished from simulation-based models in two key ways. First, empirical models use data to estimate some of the model parameters, while others are taken directly from the literature or assumed. Simulation-based models do not use empirical data to parameterize the models. For example, in [43] where empirical data was used, the mortality rate due to drug susceptible TB was estimated from Taiwanese Centre of Disease Control data and the effective contact rate for TB was estimated based on the literature; in [44] where simulation was used, the recruitment rate was taken from the literature and awareness rate of TB was estimated from data.

A second distinction between models based on empirical data and simulation-based models is that the former often report explicit estimates of the reproductive number for a specific region, while the latter usually focus on studying the impact of a certain feature on TB transmission dynamics. For example, in [45] where empirical data was used, the reproductive number was reported for India overall and by regions; in [40] where a simulation-based approach was used, the impact of drug-sensitive and drug-susceptible strains mixed together on TB transmission dynamics was studied.

In developed countries, the reproductive number was sometimes estimated to be well below 1: for example, 0.55 in the USA from 1930 to 1995 [46] and 0.26 in the Netherlands from 1993 to 1995 [35]. In developing countries, the reproductive number was as high as 4.3 in China in 2012 [36] and 3.55 in Southern India from 2004 to 2006. In the Netherlands, the reproductive number has been consistently estimated at well below one, ranging from 0.24 [37] to 0.48 [47].

The same dataset in San Francisco, USA in 1991-1992 (published in 1994) was used to estimate the effective reproductive number in two separate studies [28,29] that yielded disparate results. The estimates from these two papers were reconciled in [30] to an estimated effective reproductive number of approximately 2.1 by specifying an informative prior for two parameters in [28] and improving the convergence performance of the MCMC sampler in [29]. One can contrast this estimate with other estimates for the USA to see the range of values obtained. A study of the entire USA in [46] estimated the reproductive number to be 0.55 using case rates of active TB in USA from 1955 to 1994. As shown in [48], TB incidence in San Francisco peaked between 1991 and 1993, due to the TB/HIV co-epidemic, which is consistent with the higher estimated reproductive number (around 2.1) in [28–30]. When using TB case rates in the entire USA from 1955 to 1994 as in [46], the potential geographical and temporal heterogeneity in the estimates is not well-represented, resulting in an estimated reproductive number of 0.55. We would expect a lower reproductive number, and in particular, a reproductive number below one, when using data from 1955-1994 because by 1955, effective antibiotics were in use and BCG had also been developed, both leading to a reduction in TB incidence across the USA. In addition, Borgdorff [37] reported an effective reproductive number of 0.24 using RFLP data in San Francisco from 1991-1996. In this paper, the ratio of secondary cases and source cases was used to estimate the reproductive number, which may be an over-simplified estimator of the reproductive number. Issues such as linking the secondary cases and the sources cases have not been addressed. These divergent results indicate the need for the use of whole genome sequencing (WGS), which can be used to effectively link source and secondary cases.

Similar to the more statistical analysis of the San Francisco and the entire USA data, we observe that mathematical models lead to inconsistent results, at least partially attributable to the varying assumptions they make in their structure and parameterization. For example, even though both [36] and [43] used mathematical compartmental models with different variations for similar regions (China and Taiwan), they have quite different estimates: between 3.3 and 4.3 in China from 2005 to 2012 as compared to 0.9 for drug-sensitive TB, around 0.38 for multidrug-resistant TB (defined as a TB strain resistant to at least isoniazid and rifampicin) in Taiwan from 2005 to 2010. Both articles used incidence data from Chinese and Taiwanese CDC but formulated the compartments in the models differently. In [36], compartments “exposed”, “infectious and hospitalized” and “infectious but not hospitalized” were included; in [43], compartments “latent”, “infected” were used for two sub-populations: drug-sensitive and multidrug-resistant. The model parameters were also differently specified: in [36], some parameters were assumed while others were estimated using minimum sum of square; in [43], some parameters were given a probabilistic distribution and estimated with a root mean squared error method while others were assumed. The difference between the estimated reproductive numbers produced from these two modeling exercises is striking, as the two regions and populations are quite comparable in terms of demographics, economic status and access to health care. One could similarly contrast the modeling approaches and estimates obtained in [49] and [31], two other studies from China and Taiwan from similar time periods that produced different estimates. The differing model structures, as well as the parameter estimates, including the recruitment rate, incidence rate, and mortality rate, likely drive these observed differences. It is difficult to say which model might be a more accurate reflection of reality.

The example above illustrates the challenges of interpreting and using mathematical models for estimation of the reproductive number. However, most estimates to date make use of this approach. One shortcoming of these models is that they require assumptions about parameter values that may be difficult to estimate, such as the transmission rate, the treatment rate, and the recovery rate, which are often unobservable and not reliably estimated. As a result, most of the articles assume values for the parameters in the model based on evidence in the published literature, where it exists, sometimes without measures of uncertainty (e.g. standard errors). Model structure also varies substantially from study to study, with no generally agreed upon approach to model TB and estimate parameters. For example, in [50], compartments of different vaccine strategies were included in the model and in [36], a compartment of hospitalization was included in the model. These models also often require assumptions about the parameters used to run the models, which are likely to differ by country and time period. Sometimes sufficient data are unavailable to parameterize a model and generalizations need to be made that may not always be appropriate. The majority of the existing publications use mathematical compartmental models, which are not often ideal for statistical inference and estimation due to strong assumptions for the model structure and parameters used to run the models. While these models have the flexibility of using different compartments to evaluate the impact of policies, they are not ideal for real-time analysis where the appropriate model structure and parameter values required to fit the model may not be clear. The complexity of the natural history of TB and important factors such as HIV and drug resistance complicate these models and require additional parameters

for which the data are sometimes not available. We believe that it is important to develop, as a complementary approach to compartmental models, likelihood-based data driven analytic tools. Ideally, these estimators can be used with datasets using minimal assumptions. In addition, as WGS data become more ubiquitous [51], it will be important to develop methods that use these data to estimate the reproductive number.

This review found that the reproductive number estimates for TB are very divergent - in reality, we would expect different results in different parts of the world, reflecting diversity in TB epidemics geographically. Therefore, it is important to have estimates from a wide range of settings. An ultimate goal of methods to estimate the reproductive number should be to use routinely collected data (including potentially WGS data) to be able to monitor the reproductive number in “real-time” and evaluate interventions through this process.

Our review is subject to a number of limitations. It is possible that some useful papers could have been excluded due to our selection of search terms and our inclusion of reports in only English. These limitations are difficult to avoid in systematic reviews, in which the potential for increased yield from a wider search must be weighed against the increased feasibility of a tighter search. Additionally, our query was limited to searching in abstracts and titles, making it possible that we excluded articles where the keywords only appear in the text [25].

In conclusion, a limited number of studies have yielded explicit estimates for the serial interval and reproductive number of TB. When estimating the serial interval, it is difficult to observe the symptom onset of the infector and infected person with precision. Estimates of the reproductive number were limited geographically (Figure 6) with estimates only available for seven countries. Settings with high TB burdens, especially high drug-resistant TB burdens such as the former Soviet Union [52] are not included in these papers. In addition, there was only one estimate from a high TB and high HIV burden country [53]. The lack of estimates could be because incidence and mortality rates are currently used to monitor TB control. These rates are not suitable for monitoring transmission; reductions in mortality could be attributed to improvements in treatment outcomes rather than any change in transmission and, due to the long incubation period of TB, changes in transmission could take years to impact incidence rates. In contrast, the reproductive number can provide a direct estimate of TB transmission itself. Most studies used mathematical models with various assumed model structures and parameters, making it difficult to compare the estimates and draw useful conclusions about the TB transmission dynamics by evaluating the reproductive number.

The WHO End TB goals [4] include reducing TB deaths by 95% and incident cases by 90% by 2035. To achieve these goals, it is necessary to obtain improved estimates of the reproductive number and the SI as they can be used for monitoring and evaluating the effect of interventions on TB transmission. For example, the serial interval of TB can be used to determine how long one must monitor contacts of an infectious TB case to see if they will develop symptoms [54]. The effective reproductive number can be used to monitor the efficacy of interventions in reducing transmission. As interventions decrease transmission, estimates of the reproductive number should correspondingly decrease[44]; in particular, if

the reproductive number can be maintained below one, the disease can potentially be eliminated.

The limited number of articles that we found, and the lack of geographic representation, demonstrate a substantial gap in our understanding of these crucial parameters of TB transmission in diverse settings.

## Acknowledgments

This work was supported by the National Institute of Allergy And Infectious Diseases of the National Institutes of Health under Award Number U19AI111276; and by the Providence/Boston Center for AIDS Research (P30AI042853). HEJ was supported by the U.S. National Institutes of Health (US NIH K01AI102944 award). LFW was supported by U.S. National Institutes of Health (R01GM122876). The content is solely the responsibility of the authors and does not necessarily represent the views of the U.S. National Institute of Allergy and Infectious Diseases or the U.S. National Institutes of Health.

## Appendix

- Mathematical compartmental models

One of the simplest forms of these models divides the population into three groups (compartments): susceptible to the disease (S), infected by the disease (I) and recovered/removed from the diseased (R) (Figure 6(a)). The number of people in each compartment fluctuates as the disease dynamic changes. People move from one compartment to another at different rates: the transmission rate is β—the rate of transitioning from the susceptible compartment to the infected compartment; the recovery rate is $\gamma$. This SIR system can be expressed by a set of ordinary differential equations:

$$\frac{dS}{dt} = -\frac{\beta IS}{N};$$
$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I;$$
$$\frac{dR}{dt} = \gamma I$$

In this system, S, I, R represent the number of people in each compartment and is a function of time t. N represents the total number of people in the population and is equal to the sum of S, I and R. The basic reproductive number is defined as $R_0=\beta/\gamma$. A typical model for TB includes compartments “susceptible (S)”, “latently infected (L)”, “infectious (I)” and “recovered (R)” (Figure 6(b)). The additional compartment would require more parameters in the model for the rates between compartments, hence more differential equations are needed.

## Glossary

**Mycobacterium tuberculosis (Mtb)**
the bacterium that causes tuberculosis.

**Latent tuberculosis infection (LTBI)**
a state of persistent immune response to stimulation by Mycobacterium tuberculosis antigens without evidence of clinically manifested active TB.

**Generation interval/time**
the time between the infection of a case and the time of infection of its infector; usually unobservable.

**Serial interval/distribution**
the time between disease symptom onset of a case and that of its infector; usually observable; can serve as a surrogate for the generation interval under certain conditions.

**Incubation period/time**
the time between infection and developing active disease.

**Reproductive/reproduction number/rate/ratio/value**
the average number of secondary cases a primary infectious case will produce.

**Net reproductive number**
an interchangeable term for the reproductive number.

**Basic reproductive number**
reproductive number in a totally susceptible population where everyone is at risk of infection.

**Effective reproductive number**
reproductive number in a population with both susceptible and non-susceptible persons.

**De novo infection**
an event of a new, recent infection.

**Reactivation**
the event of latent bacteria becoming active to cause disease; can occur when the immune system becomes weakened.

**Reinfection**
the event of infection with a bacterial strain different from the original strain.

**Restriction fragment length polymorphism (RFLP)**
a technique that exploits variations in homologous DNA sequences.

**Multi-drug-resistant tuberculosis (MDR-TB)**
a form of tuberculosis (TB) infection caused by bacteria resistant to treatment with at least two of the most powerful medications, isoniazid and rifampin.

**Contact tracing**
the identification and diagnosis of people who may have come into contact with an infected person.

## References

1. World Health Organization (WHO). Global tuberculosis report. 2015

2. Center for Disease Control and Prevention (CDC). Basic TB facts, risk factors [Internet]. 2016. Available from: http://www.cdc.gov/tb/topic/basics

3. American Lung Association. Learn about tuberculosis [Internet]. 2016. Available from: http://www.lung.org/lung-health-and-diseases/lung-disease-lookup/tuberculosis/learn-about-tuberculosis.html

4. World Health Organization (WHO). Global tuberculosis report. 2016

5. Horsburgh CR, et al. Priorities for the Treatment of Latent Tuberculosis Infection in the United States. The New England Journal of Medicine. 2004; 350:2060–7. [PubMed: 15141044]

6. Biggerstaff M, Cauchemez S, Reed C, Gambhir M, Finelli L. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. BMC Infectious Diseases. 2014; 14(480)

7. Furuya H, Nagamine M, Watanabe T. Use of a mathematical model to estimate tuberculosis transmission risk in an Internet café. Environmental Health and Preventive Medicine. 2009; 14(2): 96–102. [PubMed: 19568853]

8. Bhunu CP, Garira W, Mukandavire Z, Magombedze G. Modelling the effects of pre-exposure and post-exposure vaccines in tuberculosis control. Journal of Theoretical Biology. 2008; 254(3):633–49. [PubMed: 18644386]

9. Long EF, Brandeau ML. Controlling Co-Epidemics: Analysis of HIV and Tuberculosis Infection Dynamics. Operations Research. 2009; 56(6):1366–81.

10. Gumel Abba B, Song B. Existence of multiple-stable equilibria for a multi-drug-resistant model of Mycobacterium tuberculosis.pdf. Mathematical Biosciences & Engineering. 2008; 5(3):437–55. [PubMed: 18616351]

11. Boëlle PY, Ansart S, Cori A, Valleron AJ. Transmission parameters of the A/H1N1 (2009) influenza virus pandemic: A review. Influenza and Other Respiratory Viruses. 2011; 5(5):306–16. [PubMed: 21668690]

12. Moser CB, Gupta M, Archer BN, White LF. The impact of prior information on estimates of disease transmissibility using bayesian tools. PLoS One. 2015; 10(3):1–16.

13. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic Potential of a Strain of Influenza A(H1N1):Early Findings. Science (80-). 2009; 324(5934):1557–62.

14. White LF, Wallinga J, Finelli L, Reed C, Riley S, Lipsitch M, et al. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA. Influenza and Other Respiratory Viruses. 2009; 3(6):267–76. [PubMed: 19903209]

15. Riley S, Fraser C, Donnelly CA, Ghani AC, Laith J, Hedley AJ, et al. Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions. Science (80-). 2003; 300(5627):1961–6.

16. White LF, Pagano M. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. Statistics in Medicine. 2008; 27(16):2999–3016. [PubMed: 18058829]

17. Chowell G, Hengartner NW, Castillo-Chavez C, Fenimore PW, Hyman JM. The basic reproductive number of Ebola and the effects of public health measures: The cases of Congo and Uganda. Journal of Theoretical Biology. 2004; 229(1):119–26. [PubMed: 15178190]

18. Wallinga J, Teunis P. Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal. American Journal of Epidemioly. 2004; 160(6):509–16.

19. Hartman-Adams H, Clark K, Juckett G. Update on latent tuberculosis infection. American Family Physician Journal. 2014; 89(11):889–96.

20. Lillebaek T, Dirksen A, Baess I, Strunge B, Thomsen VØ, Andersen AB. Molecular evidence of endogenous reactivation of Mycobacterium tuberculosis after 33 years of latent infection. The Journal of Infectious Diseases. 2002; 185(3):401–4. [PubMed: 11807725]

21. Lambert ML, Hasker E, Van Deun A, Roberfroid D, Boelaert M, Van der Stuyft P. Recurrence in tuberculosis: Relapse or reinfection? The Lancet Infectious Diseases. 2003; 3(5):282–7. [PubMed: 12726976]

22. Leung ECC, Leung CC, Kam KM, Yew WW, Chang KC, Leung WM, et al. Transmission of multidrug-resistant and extensively drug-resistant tuberculosis in a metropolitan city. European Respiratory Society. 2013; 41(4):901–8.
23. Vynnycky E, Fine PEM. Lifetime risks, incubation period, and serial interval of tuberculosis. American Journal of Epidemiology. 2000; 152(3):247–63. [PubMed: 10933272]
24. ten Asbroek AH, Borgdorff MW, Nagelkerke NJ, Sebek MM, Deville W, van Embden JD, et al. Estimation of serial interval and incubation period of tuberculosis using DNA fingerprinting. International Journal of Tuberculosis and Lung Disease. 1999 May; 3(5):414–20. [PubMed: 10331731]
25. Borgdorff MW, Sebek M, Geskus RB, Kremer K, Kalisvaart N, van Soolingen D. The incubation period distribution of tuberculosis estimated with a molecular epidemiological approach. International Journal of Epidemiology. 2011; 40(4):964–70. [PubMed: 21441552]
26. Brooks-pollock E, Becerra MC, Goldstein E, Cohen T, Murray MB. Epidemiologic Inference From the Distribution of Tuberculosis Cases in Households in Lima, Peru. The Journal of Infectious Diseases. 2011; 203(11):1582–9. [PubMed: 21592987]
27. Small PM, Hopewell PC, Singh SP, Paz A. The Epidemiology of Tuberculosis in San Francisco A Population-Based Study Using Conventional and Molecular Methods. The New England Journal of Medicine. 1994; 330(24):1703–9. [PubMed: 7910661]
28. Tanaka MM, Francis AR, Luciani F, Sisson Sa. Using approximate bayesian computation to estimate tuberculosis transmission parameters from genotype data. Genetics. 2006; 173(3):1511–20. [PubMed: 16624908]
29. Stadler T. Inferring epidemiological parameters on the basis of allele frequencies. Genetics. 2011; 188(3):663–72. [PubMed: 21546541]
30. Aandahl RZ, Stadler T, Sisson SA, Tanaka MM. Exact vs. approximate computation: Reconciling different estimates of Mycobacterium tuberculosis epidemiological parameters. Genetics. 2014; 196(4):1227–30. [PubMed: 24496011]
31. Liao C-M, Cheng Y-H, Lin Y-J, Hsieh N-H, Huang T-L, Chio C-P, et al. A probabilistic transmission and population dynamic model to assess tuberculosis infection risk. Risk Analysis. 2012; 32(8):1420–32. [PubMed: 22211354]
32. Brooks-Pollock E, Cohen T, Murray M. The impact of realistic age structure in simple models of tuberculosis transmission. PLoS One. 2010; 5(1):3–8.
33. Vynnycky E, Fine PE. The long-term dynamics of tuberculosis and other diseases with long serial intervals: implications of and for changing reproduction numbers. Epidemiology and Infection. 1998; 121(2):309–24. [PubMed: 9825782]
34. Andrews JR, Morrow C, Wood R. Modeling the role of public transportation in sustaining tuberculosis transmission in South Africa. American Journal of Epidemiology. 2013; 177(6):556–61. [PubMed: 23423215]
35. Borgdorff MW, Nagelkerke N, van Soolingen D, de Haas PE, Veen J, van Embden JD. Analysis of tuberculosis transmission between nationalities in the Netherlands in the period 1993-1995 using DNA fingerprinting. American Journal of Epidemiology. 1998; 147(2):187–95. [PubMed: 9457010]
36. Zhang J, Li Y, Zhang X. Mathematical modeling of tuberculosis data of China. Journal of Theoretical Biology. 2015; 365:159–63. [PubMed: 25451959]
37. Borgdorff MW, Behr MA, Nagelkerke NJD, Hopewell PC, Small PM. Transmission of tuberculosis in San Francisco and its association with immigration and ethnicity. International Journal of Tuberculosis and Lung Disease. 2000; 4(4):287–94. [PubMed: 10777075]
38. Feng Z, Castillo-chavez C, Capurro AF. A Model for Tuberculosis with Exogenous Reinfection. Theoretical Population Biology. 2000; 57(3):235–47. [PubMed: 10828216]
39. Sharomi O, Podder C. Mathematical analysis of the transmission dynamics of HIV/TB coinfection in the presence of treatment. Mathematical Biosciences and Engineering. 2008; 5(1):145–74. [PubMed: 18193936]
40. Sergeev R, Colijn C, Cohen T. Models to understand the population-level impact of mixed strain M. tuberculosis infections. Journal of Theoretical Biology. 2011; 280(1):88–100. [PubMed: 21514304]

41. Morrison J, Pai M, Hopewell PC. Tuberculosis and latent tuberculosis infection in close contacts of people with pulmonary tuberculosis in low-income and middle-income countries: a systematic review and meta-analysis. The Lancet Infectious Diseases. 2008; 8(6):359–68. [PubMed: 18450516]
42. Fox GJ, Barry SE, Britton WJ, Marks GB. Contact investigation for tuberculosis: A systematic review and meta-analysis. European Respiratory Society. 2013; 41(1):140–56.
43. Liao CM, Lin YJ. Assessing the transmission risk of multidrug-resistant Mycobacterium tuberculosis epidemics in regions of Taiwan. International Journal of Infectious Diseases. 2012; 16(10):e739–47. [PubMed: 22818291]
44. Okuonghae D, Omosigho SE. Analysis of a mathematical model for tuberculosis: What could be done to increase case detection. Journal of Theoretical Biology. 2011; 269(1):31–45. [PubMed: 20937288]
45. Narula P, Azad S, Lio P. Bayesian Melding Approach to Estimate the Reproduction Number for Tuberculosis Transmission in Indian States and Union Territories. Asia Pacific Journal of Public Health. 2015; 27(7):723–32. [PubMed: 26182939]
46. Salpeter EE, Salpeter SR. Mathematical model for the epidemiology of tuberculosis, with estimates of the reproductive number and infection-delay function. American Journal of Epidemiology. 1998; 142(4):398–406.
47. Ypma RJF, Altes HK, Van Soolingen D, Wallinga J, Van Ballegooijen WM. A sign of superspreading in tuberculosis: Highly skewed distribution of genotypic cluster sizes. Epidemiology. 2013; 24(3):395–400. [PubMed: 23446314]
48. Higashi J. Report on Tuberculosis in San Francisco: 2012 [Internet]. 2013. Available from: https://www.sfdph.org/dph/hc/HCCommPublHlth/Agendas/2013/2013/March/TB2012healthcommission.pdf
49. Zhao Y, Li M, Yuan S. Analysis of transmission and control of tuberculosis in Mainland China, 2005-2016 based on the age-structure mathematical model. International Journal of Environmental Research and Public Health. 2017; 14(10)
50. Liu S, Li Y, Bi Y, Huang Q. Mixed vaccination strategy for the control of tuberculosis: A case study in China. Mathematical Biosciences and Engineering. 2017; 14(3):695–708. [PubMed: 28092959]
51. Wyllie D, Jennifer D, Walker T, Rathod P, Crook D, Peto T, et al. A quantitative evaluation of MIRU-VNTR typing against whole-genome sequencing for identifying Mycobacterium tuberculosis transmission: A prospective observational cohort study. bioRxiv. 2018
52. Olson S, English R, Claiborne A. The New Profile of Drug-Resistant Tuberculosis in Russia [Internet]. 2011. Available from: https://www.ncbi.nlm.nih.gov/books/NBK62461/
53. Basu S, Friedland GH, Medlock J, Andrews JR, Shah NS, Gandhi NR, et al. Averting epidemics of extensively drug-resistant tuberculosis. Proceedings of the National Academy of Sciences. 2009; 106(18):7672–7.
54. Vink MA, Bootsma MCJ, Wallinga J. Serial intervals of respiratory infectious diseases: A systematic review and analysis. American Journal of Epidemiology. 2014; 180(9):865–75. [PubMed: 25294601]
55. Yang Y, Guo C, Liu L, Zhang T, Liu W. Seasonality Impact on the Transmission Dynamics of Tuberculosis. Computational and Mathematical Methods in Medicine. 2016; 2016:1–12.
56. Nebenzahl-Guimaraes H, Verhagen LM, Borgdorff MW, Van Soolingena D. Transmission and progression to disease of mycobacterium tuberculosis phylogenetic lineages in the Netherlands. Journal of Clinical Microbiology. 2015; 53(10):3264–71. [PubMed: 26224845]
57. Liu L, Wu J, Zhao X-Q. The impact of migrant workers on the tuberculosis transmission: general models and a case study for China. Mathematical Biosciences and Engineering. 2012; 9(4):785–807. [PubMed: 23311422]
58. Borgdorff MW, Van Den Hof S, Kremer K, Verhagen L, Kalisvaart N, Erkens C, et al. Progress towards tuberculosis elimination: Secular trend, immigration and transmission. European Respiratory Society. 2010; 36(2):339–47.
59. Liu L, Zhao XQ, Zhou Y. A tuberculosis model with seasonality. Bulletin of Mathematical Biology. 2010; 72(4):931–52. [PubMed: 20063125]

60. Borgdorff MW, Van Der Werf MJ, De Haas PEW, Kremer K, Van Soolingen D. Tuberculosis elimination in the Netherlands. Emerging Infectious Diseases journal. 2005; 11(4):597–602.
61. Davidow AL, Aicabes P, Marmora M. The Contribution of Recently Acquired Mycobacterium tuberculosis Infection to the New York City Tuberculosis Epidemic, 1989-1993. Epidemiology. 2000; 11(4):394–401. [PubMed: 10874545]
62. Ren S. Global stability in a tuberculosis model of imperfect treatment with age-dependent latency and relapse. Mathematical Biosciences and Engineering. 2017; 14(5/6):1337–60. [PubMed: 29161864]
63. Kheiri H, Song B, Nazari F, Castillo-Chavez C, Jabbari A. A two-strain TB model with multiple latent stages. Mathematical Biosciences and Engineering. 2016; 13(4):741–85. [PubMed: 27775384]
64. Okuonghae D, Ikhimwin BO. Dynamics of a mathematical model for tuberculosis with variability in susceptibility and disease progressions due to difference in awareness level. Frontiers in Microbiology. 2016; 6:1–23.
65. Liu L, Wang Y. A mathematical study of a TB model with treatment interruptions and two latent periods. Computational and Mathematical Methods in Medicine. 2014
66. Silva CJ, Torres DFM. Optimal control for a tuberculosis model with reinfection and post-exposure interventions. Mathematical Biosciences. 2013; 244(2):154–64. [PubMed: 23707607]
67. Hu X. Threshold dynamics for a tuberculosis model with seasonality. Mathematical Biosciences and Engineering. 2012; 9(1):111–22. [PubMed: 22229399]
68. Emvudu Y, Demasse R, Djeudeu D. Optimal control of the lost to follow up in a tuberculosis model. Computational and Mathematical Methods in Medicine. 2011
69. Roeger L-IW, Feng Z, Castillo-Chavez C. Modeling TB and HIV co-infections. Mathematical Biosciences and Engineering. 2009; 6(4):815–37. [PubMed: 19835430]
70. Gerberry DJ. Trade-off between BCG vaccination and the ability to detect and treat latent tuberculosis. Journal of Theoretical Biology. 2009; 261(4):548–60. [PubMed: 19733577]
71. Bhunu CP, Garira W, Mukandavire Z. Modeling HIV/AIDS and tuberculosis coinfection. Bulletin of Mathematical Biologyl. 2009; 71(7):1745–80.
72. McCluskey CC. Lyapunov functions for tuberculosis models with fast and slow progression. Mathematical Biosciences and Engineering. 2006; 3(4):603–14. [PubMed: 20361835]
73. Martcheva M, Thieme HR. Progression age enhanced backward bifurcation in an epidemic model with super-infection. Journal of Mathematical Biology. 2003; 46(5):385–424. [PubMed: 12750833]
74. Aparicio JP, Capurro AF, Castillo-Chavez C. Transmission and dynamics of tuberculosis on generalized households. Journal of Theoretical Biology. 2000; 206(3):327–41. [PubMed: 10988019]
75. Beatriz M, Leite F, Bassanezi RC, Yang HMO, De Campinas UE, Depart I, et al. The basic reproduction ratio for a model of directly transmitted infections considering the virus charge and the immunological response. Mathematical Medicine and Biology: A Journal of the IMA. 2000; 17(1):15–31.
76. Castillo-Chavez C, Feng Z. Global stability of an age-structure model for TB and its applications to optimal vaccination strategies. Mathematical Biosciences. 1998; 151(2):135–54. [PubMed: 9711046]
77. Lietman T, Porco T, Blower S. Leprosy and tuberculosis: The epidemiological consequences of cross-immunity. American Journal of Public Health. 1997; 87(12):1923–7. [PubMed: 9431277]
78. Sanchez MA, Blower SM. Uncertainty and Sensitivity Analysis of the Basic Reproductive Rate. American Journal of Epidemiology. 1997; 145(12)
79. Singer BH, Kirschner DE. Influence of backward bifurcation on interpretation of r(0) in a model of epidemic tuberculosis with reinfection. Mathematical Biosciences and Engineering. 2004; 1(1): 81–93. [PubMed: 20369961]
80. Trauer JM, Denholm JT, McBryde ES. Construction of a mathematical model for tuberculosis transmission in highly endemic regions of the Asia-pacific. Journal of Theoretical Biology. 2014; 358:74–84. [PubMed: 24878110]

81. Dye C, Williams BG. Criteria for the control of drug-resistant tuberculosis. Proceedings of the National Academy of Sciences. 2000; 97(14):8180–5.

82. Blower SM, Chou T. Modeling the emergence of the “hot zones”: tuberculosis and the amplification dynamics of drug resistance. Nature Medicine. 2004; 10(10):1111–6.

83. Blower S, McLean A, Porco T, Small P. The intrinsic transmission dynamics of tuberculosis epidemics. Nature Medicine. 1995; 1(8):815–21.

84. Blower SM, Gerberding JL. Understanding, predicting and controlling the emergence of drug-resistant tuberculosis: a theoretical framework. Journal of Molecular Medicine. 1998; 76(9):624–36. [PubMed: 9725765]


**Figure 1.**

Important infectious disease intervals. The time between a and c is the serial interval; the time between b and c is the incubation period.

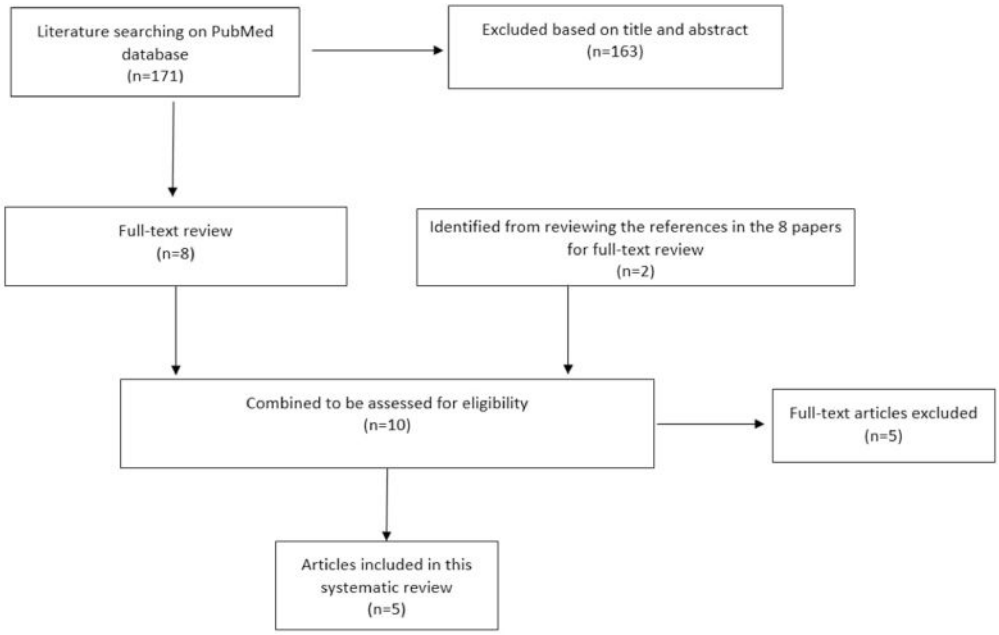


**Figure 2.**
Flow diagram of articles included in the search of estimates of the serial interval.

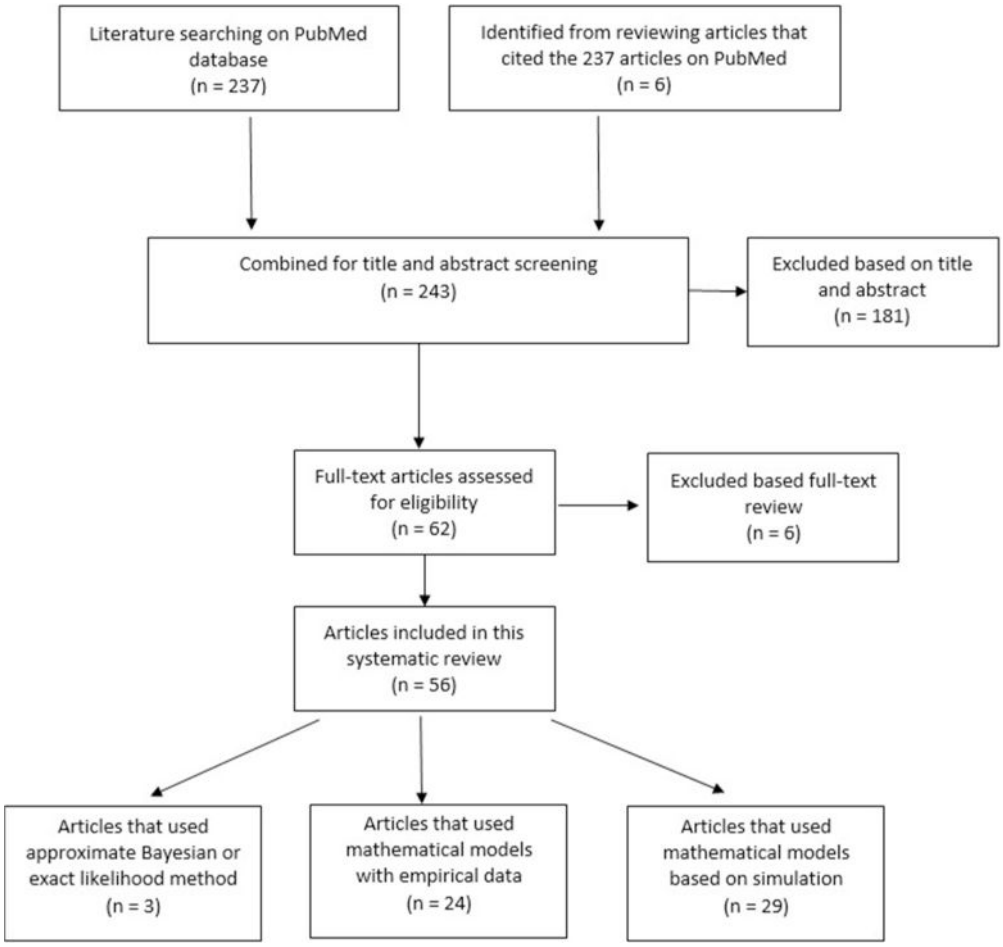


**Figure 3.**
Flow diagram of articles included in the search of estimates of the reproductive number.

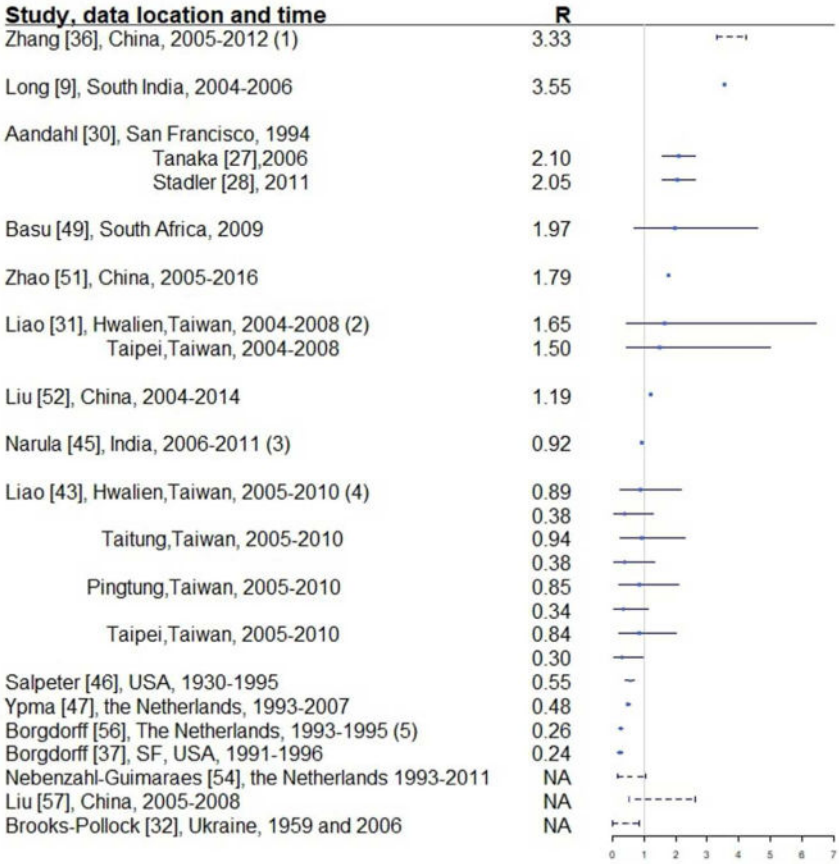


**Figure 4.**

Reproductive number from studies with explicit R estimate from empirical data.

Notes:

(1) The range is for years 2005-2012, with the reproductive number estimated at 3.33, 3.72, 3.38, 3.97, 4.29, 3.32, 3.92 and 4.30, respectively.

(2) For each location, the first R corresponds to drug-sensitive population and the second correspond to drug-resistant population.

(3) R estimated for 35 states and union territories of India with estimates ranging from 0.72 to 0.98; 0.92 is the overall estimate for India.

(4) For each location, the first R corresponds to drug-sensitive population and the second correspond to drug-resistant population.

(5) Bordgorff in [35, 56, 58] estimated the reproductive number for the Nethelands from 1993 to 2007 at around 0.26 with lower bound of the 95% CI around 0.20 and upper bound around 0.32.

(6) Broken lines indicate range; solid lines indicate 95% confidence interval.

(7) Vynnycky [23] in 1998 estimated the basic reproductive number to decline from about 3 in 1900 to 2 in 1950 and to below 1 in about 1960 for England and Wales, which is not included in this graph.

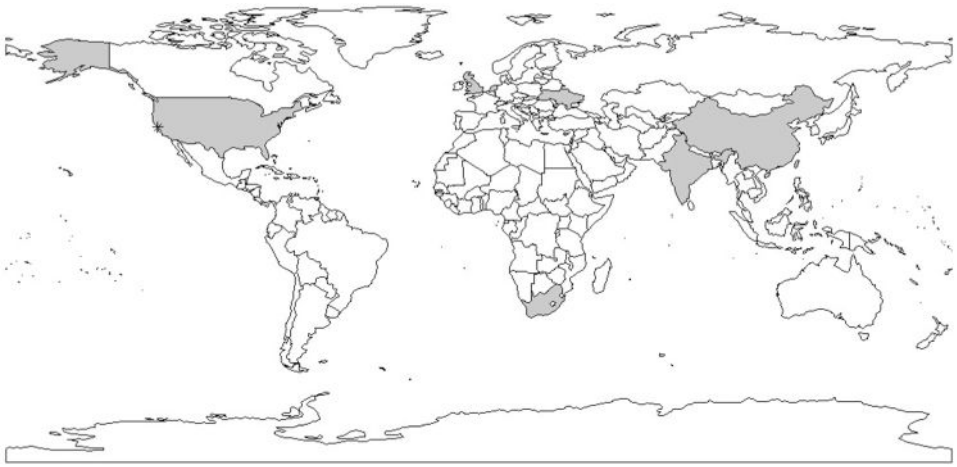

**Figure 5.**

Shaded areas and stars indicate countries and cities with reproductive number estimates. Multiple estimates: China, Taiwan, USA, India; one estimate: Ukraine, the Netherlands, South Africa, the UK. *indicates San Francisco corresponding to data used in [28, 29, 30].

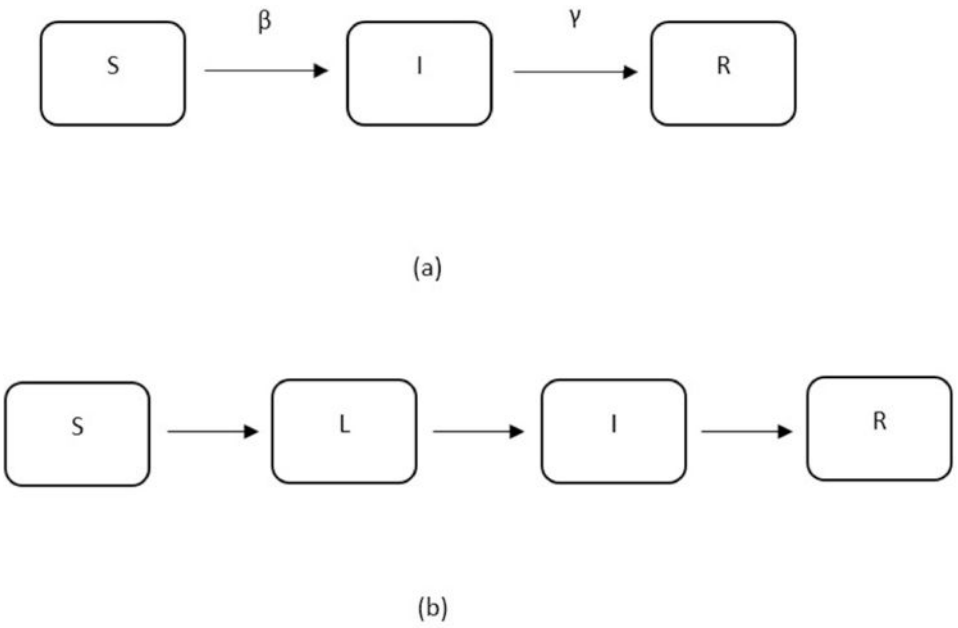


**Figure 6.**
Examples of mathematical compartmental models.

**Table 1**

Estimates of the Serial Interval

| Name of 1st author, publication year | Objective | Method | Assumptions | Estimated serial interval |
|---|---|---|---|---|
| Leung [22], 2013 | Study household transmission of MDR-TB | Data on all MDR-TB in Hong Kong 1997-2006. Did contact investigations and DNA fingerprinting and linked index to secondary | No censoring in this estimate, not clear how long people were followed up for cases | 1.4 (0.4-5.2) years (median) |
| Vynnycky [23], 2000 | Demonstrate how the lifetime risk of disease, the incubation period and the serial interval changed | An age-dependent compartmental model | Assumed values for model input parameters such as the annual risk of infection | Estimated as dependent on age of infection and summarized as frequency distributions |
| ten Asbroek [24], 1999 | To determine the serial interval and incubation period of tuberculosis within 4 years of ransmission | Descriptive approach on RFLP data (used to link infectors and infected people) | One source of infection for each infected cluster | 0.57 years (95% CI: 0.44-0.73) |
| Borgdorff [25], 2011 | Same as [24] | Same as [24] | NA | Median: 1.44 years (95% CI: 1.29-1.63) |
| Brooks-Pollock [26], 2011 | Estimate the relative contributions of household and community transmission, the serial interval, and the immunity afforded by a previous TB infection | Descriptive approach for the serial interval | All members of the study cohort have been exposed to TB by living with someone with active disease | Mean serial interval: 3.5 years; median serial interval 1.65 years |

**Table 2**

Estimates of the Reproduction Number Using Approximate Bayesian Computation and Exact Likelihood Methods (All methods used data from San Francisco on cases reported in 1994)

| Name of 1st author, publication year | Objective | Method | Assumptions | Estimated reproductive number (95% credible interval) |
|---|---|---|---|---|
| Tanaka [28], 2006 | Estimate TB transmission parameters: net transmission rate, doubling time and reproductive number | Approximate Bayesian computation | Constant supply of susceptible people; all genotypes are selectively neutral; mutation and transmission are independent; infinite alleles; epidemic spreads until N individuals are infected | 3.4 (1.4-79.7) |
| Stadler [29], 2011 | Estimate TB transmission parameters: net transmission rate and reproductive number | Exact likelihood | Constant birth-death rate; infinite alleles; epidemic started at a random time in the past; an isolate is sampled from an individual with probability p | 1.02 (1.01-1.04) |
| Aandahl [30], 2014 | Reconcile the different estimates in [28] and [29] | Improved the method in [28] by specifying informative priors; improved convergence performance of the MCMC sampler in [29] | Fixed mutation rate; used Gaussian prior for the death/recovery rate | 1) 2.1 (1.54-2.66) in [28]<br>2) 2.05 (1.55-2.63) in [29] |

**Table 3**

Estimates of the Reproductive Number from Mathematical Models with Empirical Data

| Name of 1st author, publication year | Location, time of data | Type of data | Objective | Methods | Assumptions | Reproductive number type | Estimated reproductive number |
|---|---|---|---|---|---|---|---|
| Zhao [49], 2017 | China, 2005-2016 | CDC data | To investigate the impact of age on TB transmission | SEIR model with age structure; use least squares to get parameters that align with TB data in China; use latin hypercube sampling to get CI | Although the susceptible compartment was stratified by age, the other compartments were not age-stratified thus assuming no difference in age for those compartments | Basic | 1.786 (95% CI: 1.775 - 1.796) |
| Liu [50], 2017 | China, 2004-2014 | Annual TB case data | To use modeling to investigate the impact of different vaccination strategies (constant or pulse BCG) on TB transmission | Compartmental models with vaccination compartments | Assumptions made for all parameter values | Basic | 1.19 |
| Yang [55], 2016 | Shaanxi, China, 2004-2012 | Notifiable active TB cases by month | Study the seasonality impact on TB transmission dynamics | A seasonality TB compartmental model: subjects either entered latent or diseased compartment; contact rate, reactivation rate and disease-induced death rate are periodic continuous functions | Parameter values for recruitment rate, natural death rate, recovery rate | Basic | Dependent on parameter values |
| Nebenzahl-G uimaraes [56], 2015 | The Netherlands, 1993-2011 | Surveillance and RFLP data | Determine if mycobacterial lineages affect infection risk, clustering, and disease progression among Mycobacterium tuberculosis cases | Descriptive and regression approach; DNA fingerprinting to link cases | All secondary cases captured in surveillance data; genetic matching accurately reflects transmission patterns | Effective | Range: 0.17-1.04 |
| Narula [45], 2015 | India, 2006-2011 | Quarterly reported data from Central TB Division | Estimate basic R0 for TB | Compartmental model with | Some parameter | Basic | 0.92, averaged for India overall with range 0.72-0.98 |

| Name of 1st author, publication year | Location, time of data | Type of data | Objective | Methods | Assumptions | Reproductive number type | Estimated reproductive number |
|---|---|---|---|---|---|---|---|
| | | | | Bayesian melding technique to estimate parameters; Susceptible, latent, infected compartments instead of SIR | values assumed with reference in the differential equations | | |
| Zhang [36], 2015 | China, 2005-2012 | Monthly case reporting data from CDC | Estimate effective R0 of TB by year | Compartmental model adding hospitalized compartment; Chi-Square test for optimal parameters | An upper bound for number of initially susceptible people, natural death rate, initial number of latent individuals | Effective | Range from 3.318-4.302 from year 2005 to 2012 |
| Ypma [47], 2013 | The Netherlands, 1993-2007 | RFLP data | Explore the high heterogeneity in the number of secondary cases caused per infectious individual for TB | Model “superspreading” parameter as a negative binomial distribution | Immigrants who have been in the country for less than six months at diagnosis are index cases themselves | Fingerprint reproductive number as a function of the effective reproductive number and the probability that the fingerprint of the infected person is different than its infector | 0.48 (95% CI 0.44-0.59) |
| Andrews [34], 2012 | Cape Town, South Africa, 2011 | Carbon dioxide data, public transit usage data from national survey | Estimate risk of TB transmission on 3 modes of public transit | Modified Wells-Riley model for airborne disease transmission | Duration of infectiousness of 1 year; used TB and HIV parameters from studies in the same area; natural history parameters from literature | Basic | Dependent on duration of infectiousness and frequency of transit usage |
| Okuonghae [44], 2011 | Benin city, Nigeria, 2008 | Survey data | Assess how control strategies on addressing TB transmission parameters can minimize incidence | Compartmental model adding compartments of disease awareness level, identified infectiousness | Model parameter values such as recruitment rate, recovery rate from literature | Basic, under treatment | Dependent on parameter values |
| Liao [43], 2012 | Taiwan, 2005-2010 | Monthly data from CDC | Estimate MDR-TB infection risk | Mathematical probabilistic two-strain model with compartments for drug-sensitive and drug-resistant subjects; dose-response model | Some model parameter values from data, some from literature; assumed 0.99 of people | Basic | **Hwalien County**: 0.89 (95% CI: 0.23-2.17) for drug sensitive; 0.38 (95% CI: 0.05-1.30) for multi-drug resistant; **Taitung County**: 0.94 (95% CI: 0.24-2.28) for drug sensitive; 0.38 (95% CI: 0.05-1.33) for multi-drug resistant; |

| Name of 1st author, publication year | Location, time of data | Type of data | Objective | Methods | Assumptions | Reproductive number type | Estimated reproductive number |
|---|---|---|---|---|---|---|---|
| | | | | for relationship between R0 and total proportion of infected population | latently infected were drug sensitive and 0.01 were drug resistant | | **Pingtung County**: 0.85 (95% CI: 0.21-2.08) for drug sensitive; 0.34 (95% CI: 0.04-1.13) for multi-drug resistant; **Taipei City**: 0.84 (95% CI: 0.21-2.00) for drug sensitive; 0.30 (95% CI: 0.04-0.97) for multi-drug resistant; |
| Liao [31], 2012 | Taiwan, 2004-2008; selected 3 areas with the highest incidence, 1 with the lowest incidence | Monthly disease burden TB data from Taiwan CDC | Examine TB population dynamics and assess potential infection risk | Compartmental model with susceptible, latently infected, infectious, non-infectious and recovered compartments; incorporated reactivation, relapse and reinfection | Some parameter values taken from literature, some estimated from data | Basic, estimated as sum of fast, slow and relapse | Highest R0 total in Hwalien: 1.65 with 95th percentile range 0.45-6.45; Taipei lowest at 1.5 (0.45-4.98); Taitung: 1.72; Pingtung: 1.65 |
| Liu [57], 2012 | China, 2000-2008 | Data from National Bureau of Statistics | Incorporate migration to study TB transmission | SEIR compartments for rural residents, migrant workers and urban population | Model parameters calculated from website data; migration rates | Basic | No explicit estimate |
| Borgdorff [58], 2010 | The Netherlands, 1993-2007 | RFLP data | Determine to what extent tuberculosis trends in the Netherlands depend on secular trend, immigration and recent transmission | DNA fingerprinting to link cases | All secondary cases captured in surveillance data; genetic matching accurately reflects transmission patterns | Basic | 0.24 (95% CI 0.21-0.26) |
| Liu [59], 2010 | China, Jan, 2005-Dec, 2008 | Monthly notification data from Ministry of Health | Develop a model incorporating seasonality and define basic reproduction ratio | Used periodic infection rate and reactivation rate to incorporate seasonality in the compartmental model; considered fast and slow progression | Parameters such as recruitment rate, natural death rate were assumed to be constants; some parameter values assumed and some taken from literature | Basic | Dependent on parameter values with range 0.4-2.6 |
| Brooks-Pollock [32], 2010 | Ukraine, 1959 and 2006 | Mortality data | Explore the effect of age structure on TB infection and disease | Basic SEIR mathematical model with assumptions | A survivorship function which could be described | Basic | Dependent on progression rate with range 0-0.85 |

| Name of 1st author, publication year | Location, time of data | Type of data | Objective | Methods | Assumptions | Reproductive number type | Estimated reproductive number |
|---|---|---|---|---|---|---|---|
| | | | prevalence, basic reproductive number and impact of intervention | about survivorship | in terms of age and life expectancy | | |
| Basu [53], 2009 | KwaZulu-Natal, South Africa | Extensively drug-resistant TB data (XDR-TB) | Model XDR-TB transmission dynamics | Model XDR-TB incorporating the existing XDR detection rate and treatment system | Even mixing of air; range of key parameters in the model | Effective | 1.97, range 0.7-4.6; 1.23, range 0.4-3.1 when combining screening and therapy; 1.38, range 0.6-3.3 with South African strategic plan alone. |
| Furuya [7], 2008 | Japan, 2000-2005 | Exposure data | Quantify the risk of TB infection in an internet café where people without homes stayed overnight | Wells-Riley model to estimate the reproductive number | Patients stayed in a confined space for 150 days; some values in the Wells-Riley equation assumed, others from literature | Estimated as a function of exposure period | Dependent on exposure period |
| Long [9], 2008 | Southern India, 2004-2006 | HIV-TB co-epidemics data | Model HIV-TB co-epidemics and explore hypothetical treatment effect | First model: susceptibility to either or both diseases compartments; second model: $SII^{*}SEI$ | A linear relationship between treatment levels and the associated parameters; model parameters from literature | Basic | R=3.55 when no active treatment for TB |
| Borgdorff [60], 2005 | The Netherlands, 1995-2002 | RFLP data | Assess progress towards TB elimination | DNA fingerprinting to link cases; survival analysis | All secondary cases captured in surveillance data; genetic matching accurately reflects transmission patterns | Basic | Dutch index cases: 0.23, non-Dutch index cases: 0.25 |
| Borgdorff [37], 2000 | San Francisco, USA, 1991-1996 | RFLP data | Determine tuberculosis transmission dynamics in San Francisco and its association with country of birth and ethnicity | Define effective reproductive number as a function of transmission index, which is a function of number of secondary cases and potential source cases in a given subgroup | Each cluster originates from a single source case in the database; either the first case of a cluster was its source case, or that the probability of being a source case declined exponentially | Effective, recent transmission | 0.24 (95% CI: 0.17-0.31) |

| Name of 1st author, publication year | Location, time of data | Type of data | Objective | Methods | Assumptions | Reproductive number type | Estimated reproductive number |
|---|---|---|---|---|---|---|---|
| | | | | | over time by 0.77% per day | | |
| Davidow [61], 2000 | New York City, 1989-1993 | TB and AIDS surveillance data | Evaluate the importance of recent M. tuberculosis transmission | Estimated # of TB infectious cases 1 year ago and computed short-term R0; R0= the average # of new infections caused by each case per year of infectiousness*the average duration of infectiousness*the probability of progressing to active TB within 1 year after infection | Some clinical assumptions; parameter values in equation taken from literature or calculated from neighborhood-specific data | Short-term | No explicit estimates; focused on percentage of TB cases due to infection 1 year ago |
| Vynnycky [33], 1998 | England and Wales, 1900 | Surveillance data; age and time-specific mortality rates | Describe transmission dynamics of all forms of pulmonary TB | Age-structured mathematical model with compartments for endogenous and exogenous diseases | General relationship between: 1st primary episode and age at infection, risk of exogenous disease and age at reinfection, endogenous disease and current age; risk of reinfection and 1st infection are identical; parameter values from literature | Basic and net which is the same as effective | Net R at about 1 from 1900-1950; basic R0 declined from about 3 in 1900, reached 2 by 1950, and first fell below 1 in about 1960 |
| Salpeter [46], 1998 | USA, 1930-1995 | Case rates, correction for rates before 1975 | Estimate time delay from infection to disease and R | Estimate R as a function of case rate and the shape of the delay function | R and case rate constant with calendar time t; incidence rate of latent infection is independent of the age | Effective | 0.55, range 0.4-0.7 in sensitivity analysis |
| Borgdorff [35], 1998 | Netherlands, 1993-1995 | RFLP data | Quantifying transmission of TB between and within nationalities | Effective R0 estimated as a function of transmission index | Probability of a patient being the source of a cluster was proportional | Effective | 0.26, 95% CI: (0.20-0.32); also estimated for different nationalities |

| Name of 1st author, publication year | Location, time of data | Type of data | Objective | Methods | Assumptions | Reproductive number type | Estimated reproductive number |
|---|---|---|---|---|---|---|---|
| | | | | | to the incidence rate of potential sources times the probability that a potential source would give rise to a cluster | | |

**Table 4**

Estimates of the Reproductive Number from Mathematical Models based on Simulation

| Name of 1st author, publication year | Objective | Methods, setting | Assumptions | Basic or effective R0 | Estimated R0 |
|---|---|---|---|---|---|
| Ren [62], 2017 | Develop SEIR model for imperfect treatment with age-dependent latency and relapse | SEIR model | TB infectious in latent period; age-dependence | Basic | Dependent on parameters |
| Jabbari [63], 2016 | To set up a model that can examine two TB strains (DS and DR) with multiple latent stages | Mathematical compartmental model with compartments for latency stages | The drug-sensitive strain will not play a role in the process of exogenous reinfection for the drug-resistant strain | Basic | Dependent on parameters |
| Okuonghae [64], 2016 | Study the effects of additional heterogeneities from the level of TB awareness on TB transmission dynamics and case detection rate | Expanding [49] by dividing both susceptible and latently infected compartment by level of TB awareness | Reasonable values and bounds for parameters such as transmission rate, recovery rate from literature | Effective | Dependent on parameters such as active case finding rate and treatment rate |
| Liu [65], 2014 | Evaluate effect of treatment for TB | Compartmental model with treatment and two latent periods incorporated | Once the treatment of active TB cases is interrupted, there is no more treatment; specified model parameter values and their relationship with one another | Basic | Dependent on transmission coefficients |
| Silva [66], 2013 | Study optimal strategies for the controlling active TB infectious and persistent latent individuals | Compartmental model considering reinfection and post-exposure interventions with the addition of early latent and persistent latent compartments | Parameter values taken from literature | Basic | Dependent on transmission coefficient |
| Hu [67], 2012 | Study the threshold dynamics of TB | Compartmental model with periodic functions for reactivation rate and infection rate; include additional compartment for treated people that do not return to the hospital for examination | NA | Basic | Dependent on transmission coefficient |
| Emvudu [68], 2011 | Address the problem of optimal control for TB transmission dynamics | Compartmental model with an additional compartment for loss to follow-up | Half of the parameter values were assumed; others taken from Cameroon literature | Basic | Dependent on parameters such as transmission rate |
| Sergeev [40], 2011 | How drug-sensitive and drug-resistance strains mixed together can impacts long-term TB dynamics | Compartmental with the 3 compartments for both latent and infected: drug-resistant, drug sensitive and mixed strains | Reasonable values for many parameters; few data exist to inform model parameters | Basic; estimated for drug-resistant, drug-sensitive and mixed strains | Dependent on model parameters |
| Roeger [69], 2009 | Model TB and HIV co-infection | Compartmental model for joint dynamics of TB and HIV and compute independent reproductive numbers for the two diseases | Probability of infection is the same for those treated with TB and those susceptible; assumed | Overall R0 as the max of R0 for TB and HIV | Dependent on model parameters |

| Name of 1st author, publication year | Objective | Methods, setting | Assumptions | Basic or effective R0 | Estimated R0 |
|---|---|---|---|---|---|
| | | | relationship among model parameters | | |
| Gerberry [70], 2009 | Study the trade-off between BCG and detection, treatment of TB | Compartmental model with additional compartments for latently infected and unvaccinated, latently infected and vaccinated; establish thresholds for basic R0 | Throughout the duration of the vaccine's efficacy, latent TB completely undetectable | Basic | Dependent on model parameters |
| Bhunu [71], 2009 | Model HIV/AIDS and TB coinfection | Compartmental model for TB, HIV separately without intervention; full model with intervention | Parameter values from Central Statistics Office of Zimbabwe and literature; relationship amongst parameters in the model | Basic | Dependent on model parameters |
| Bhunu [8], 2008 | Model the effect of pre-exposure and post-exposure vaccines | Compartmental model with additional compartments for susceptible (vaccinated or not) and latent (history of vaccine or not) | Homogeneous mixing; recovered people would not develop disease from reinfection, but could be re-infected; parameter values taken from Central Statistics Office and literature | Basic | Dependent on model parameters |
| Sharomi [39], 2008 | Address the interaction between HIV and TB | TB-only, HIV-only and full model analyzed with both susceptible and latent compartments divided according to TB and HIV status | Dually infected people could not transmit both diseases; some parameters taken from literature, others assumed | Basic | Dependent on model parameters |
| McCluskey [72], 2006 | Address global stability of high dimensional TB model | Use Lyapunov function to demonstrate the stability of the endemic equilibria in mathematical models for TB: SEIR, SEIS and SIR; fast and slow progression incorporated | | Basic | No explicit estimate |
| Martcheva [73], 2003 | Address the issue of an infected person being subject to further contacts with infectious individuals—"su per infection" | Subdivide the latent stage into one where the disease progresses and one where the disease development is on hold | Relationship among model parameters | Basic | No explicit estimate |
| Aparicio [74], 2000 | Express basic R0 as a function of cluster size | Divide individuals into either active clusters or otherwise | Homogeneous mixing | Basic | No explicit estimate; expressed as a function of household size |
| Feng [38], 2000 | Examine how exogenous reinfection changes the TB transmission dynamics | Include additional parameters in the mathematical model to model exogenous reinfection | Constant per capita removal rate to focus on the role of reinfection | Basic | No explicit estimate; analytical expression |
| Beatriz [75], 2000 | Assess the effects of heterogeneous infectivity | Divide infective period into k stages | Homogenous mixing; bilinear incidence rate | Basic | No explicit estimate; analytical expression |
| Castillo-Ch avez [76], 1998 | Use an age-structure model to study the dynamics of TB | Use age-specific parameters in the compartmental model; transmission dynamics studied for with and without vaccine | Mixing between individuals is proportional to their age-dependent activity level; | Net and basic | No explicit estimate; analytical expression |

| Name of 1st author, publication year | Objective | Methods, setting | Assumptions | Basic or effective R0 | Estimated R0 |
|---|---|---|---|---|---|
| | | | disease-induced death rate neglected | | |
| Lietman [77], 1997 | Test the hypothesis that exposure to TB leads to disappearance of leprosy | Add in leprosy compartment in the mathematical model | Cross-immunity is symmetric: same immunity for TB and leprosy | Basic | Dependent on R0 of leprosy and cross-protection rate |
| Sanchez [78], 1997 | Evaluate the effects of parameter estimation uncertainty on the value of Ro | Latin hypercube sampling used on parameters in the compartmental model in Blower [78] to evaluate uncertainty of R0 | Range for parameters in the compartmental model | Sum of R0 for fast, slow and relapse | Dependent on parameters in the model |
| Gumel [10], 2008 | Study the transmission dynamics of TB with multiple strains, in the presence of exogenous reinfection | Included drug-sensitive and resistant strains in the compartmental model; exogenous reinfection incorporated | Homogenous mixing | Effective R0 for the two strains | Dependent on parameters in the model |
| Singer [79], 2004 | Study the impact of different reinfection levels of latently infected individuals on TB transmission dynamics | Compartmental model for heterogeneous population: one group more susceptible to infection than the other; | Parameter range uniformly distributed according to previous papers | Basic | No explicit estimate |
| Trauer [80], 2014 | Model TB transmission for highly endemic regions of the Asia-Pacific where HIV-coinfection is low | Compartmental models with compartments for immunization, latency, reinfection, drug-resistance, etc. | Parameters fixed values according to papers and WHO | Basic | Dependent on parameters; computed as 8.34 for drug-susceptible and 5.84 for drug-resistant at baseline |
| Dye [81], 2000 | To establish criteria for MDR-TB control | Compartmental models with compartments for drug-susceptible, drug-resistant, treatment failure, etc. | Parameters calculated from different populations | Basic | Dependent on parameters; best estimated of the model parameters yielded $R_0$=1.6 (95% CI: 1.02-2.67) |
| Blower [82], 2004 | Track the emergence and evolution of multiple strains of drug-resistant TB | Non-compartmental mathematical model. | NA | Basic | Dependent on drug-susceptibility of TB |
| Blower [83], 1995 | Model the transmission dynamics of TB | Compartmental models with latently infected, infectious, non-infectious, recovered compartments. | Some model parameters assumed; some taken from references | Basic; defined as the sum of slow progression, recent transmission and relapse | Median of 4.47, range: 0.74-18.58 |
| Blower [84], 1998 | Understand, predict and control TB | Compartmental models with drug-sensitive and drug-resistant compartments | NA | Basic | Dependent on model parameters |
| Aparicio [74],2009 | Evaluate homogeneous mixing and heterogeneous mixing models for TB | Three types of compartmental models: a standard incidence homogenous mixing mode; a heterogeneous mixing model; an age-structured model. | Assumptions on model parameters | Basic | Dependent on model parameters |