Epidemics xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Optimally capturing latency dynamics in models of tuberculosis transmission

Romain Ragonnet[a,b,*], James M. Trauer[a,c,d], Nick Scott[b,c], Michael T. Meehan[e], Justin T. Denholm[a,d,f], Emma S. McBryde[a,e]

[a] Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Australia
[b] Burnet Institute, Australia
[c] School of Population Health and Preventive Medicine, Monash University, Australia
[d] Victorian Tuberculosis Program, Melbourne, Australia
[e] Australian Institute of Tropical Health and Medicine, James Cook University, Australia
[f] Royal Melbourne Hospital, Melbourne, Australia

ARTICLE INFO

Article history:
Received 13 February 2017
Received in revised form 30 May 2017
Accepted 14 June 2017
Available online xxx

Keywords:
Tuberculosis latency
Mathematical modelling
Risk of disease activation
Parameter estimation

ABSTRACT

Although different structures are used in modern tuberculosis (TB) models to simulate TB latency, it remains unclear whether they are all capable of reproducing the particular activation dynamics empirically observed. We aimed to determine which of these structures replicate the dynamics of progression accurately. We reviewed 88 TB-modelling articles and classified them according to the latency structure employed. We then fitted these different models to the activation dynamics observed from 1352 infected contacts diagnosed in Victoria (Australia) and Amsterdam (Netherlands) to obtain parameter estimates. Six different model structures were identified, of which only those incorporating two latency compartments were capable of reproducing the activation dynamics empirically observed. We found important differences in parameter estimates by age. We also observed marked differences between our estimates and the parameter values used in many previous models. In particular, when two successive latency phases are considered, the first period should have a duration that is much shorter than that used in previous studies. In conclusion, structures incorporating two latency compartments and age-stratification should be employed to accurately replicate the dynamics of TB latency. We provide a catalogue of parameter values and an approach to parameter estimation from empiric data for calibration of future TB-models.

1. Introduction

Tuberculosis (TB) is a major health issue with 10.4 million active cases and 1.8 million deaths worldwide in 2015 (WHO, 2016). Furthermore, around one quarter of the world's population is estimated to be infected with TB (Houben and Dodd, 2016), representing a huge reservoir of potential disease. Accordingly, fully understanding latent TB infection is crucial for assessing the future epidemic trajectory and designing effective TB control policies. Despite this, much reinfection occurs in high incidence cohorts, hampering accurate estimation of latency dynamics. Therefore insights into the activation dynamics following a single infection episode of *Mycobacterium tuberculosis* provided by recent studies in very low transmission settings are particularly valuable (Trauer et al., 2016a; Sloot et al., 2014). These works provide detailed information on patterns of activation, highlighting that most active cases occur within the first few months of infection.

Mathematical modelling has informed TB control programs by simulating interventions, or by explaining the mechanisms underlying observed epidemiological trends (Vynnycky and Fine, 1997; Gomes et al., 2004; Castillo-Chavez and Feng, 1997; Abu-Raddad et al., 2009; Cohen et al., 2008; Dye, 2012; Menzies et al., 2012; Trauer et al., 2016b), yet little is known about whether such modelling has been able to capture latency dynamics accurately. In the past, TB models have been constructed to capture the lifelong probability of disease and, although some models allowed for marked differences between the early and late dynamics of infection (Dowdy et al., 2013; Lin et al., 2011; Aparicio and Castillo-Chavez, 2009), estimates for the associated parameters have not been fit closely to longitudinal data. Despite this, it has been shown

* Corresponding author at: 85 Commercial Road, Burnet Institute, Melbourne, VIC 3003, Australia.
E-mail address: romain.ragonnet@burnet.edu.au (R. Ragonnet).

http://dx.doi.org/10.1016/j.epidem.2017.06.002

that when modelling infectious diseases, it is critical to employ appropriate distributions of latent periods (Wearing et al., 2005). Focusing on emerging infectious diseases, Wallinga and Lipsitch further demonstrated that capturing the mean of the generation times is not sufficient to characterise transmission accurately, as the shape of the distribution of the generation intervals also plays a critical role in infection dynamics (Wallinga and Lipsitch, 2007). Although TB is an ancient disease, its epidemiology is continuously evolving. In particular, changes in TB epidemiology in response to emerging phenomena, such as introduction of drug-resistant forms of TB or stronger control programs, are likely to affect the shape of the generation time distribution. Therefore, the recent detailed characterisation of TB activation dynamics represent a valuable opportunity to review and improve modelling practices for the simulation of TB latency.

Compartmental dynamic transmission models – the most common type of TB mathematical model – simulate TB latency with various levels of complexity. While some modellers employ a single latency compartment that precedes the active disease compartment (Colijn et al., 2008; Blower and Chou, 2004), others incorporate a second latency compartment in order to capture two different rates of progression from latent infection to active disease (Hill et al., 2012; Cohen et al., 2006; Trauer et al., 2014). When two latency compartments are incorporated, they can either be positioned in series or in parallel, involving different underlying assumptions regarding the progression pathways to active disease. First, the serial structure implies that newly infected individuals remain at high risk of disease during the initial phase and then, if TB activation has not occurred, they transition to another compartment where their risk of developing active TB is reduced. By contrast, with a parallel compartmental structure, the underlying assumption is that a proportion of infected individuals belong to a high risk category, while the remainder are at lower risk of TB disease. While TB modelling has been used extensively for over 40 years, it remains unclear which of these structures are best adapted to the natural history of TB.

In this study, we aim to determine the most appropriate model structures to simulate TB latency and provide estimates for the parameters associated with these structures across different age categories. We use the distribution of the estimated times from infection to TB activation in 1352 infected contacts of individuals with active pulmonary TB from Victoria (Australia) and Amsterdam (Netherlands) to calibrate the latency structures of different candidate models to the dynamics observed in the data.

2. Methods

2.1. Literature review

Our search was based on the literature review of mathematical and economic TB modelling articles provided by the TB Modelling and Analysis Consortium, available online at http://tb-mac.org/Resources/Resource/4 (see Appendix in Supplementary file for more details). From this database we identified all 88 publications reporting the use of a deterministic compartmental transmission dynamic model. All selected papers were reviewed independently by two authors (RR, JMT) who classified the manuscripts according to the structure used to model TB latency. These two independent investigations led to the same classification which is presented in the Appendix in Supplementary file (Table S1).

2.2. Analytical solution

For each latency structure found in the literature, we associated a basic dynamic model comprised of the latency structure in combination with compartments representing susceptibility to infection and active disease. We then found analytical solutions for the TB activation dynamics corresponding to each model. Namely, considering that individuals were infected at time $t = 0$, we determined the proportion $I(t)$ of infected individuals that had developed active TB after each time t ($t \geq 0$). Analytical expressions are also presented for the total proportion of infected individuals progressing to active disease, obtained by calculating the limit of $I(t)$ as t approaches plus infinity. The detailed method used to obtain the analytic solutions is described in the Appendix in Supplementary file.

2.3. Data used to calibrate the models

The models described above were calibrated to individual data on close contacts of individuals with active pulmonary TB notified in the Australian state of Victoria from January 2005 to December 2013. These data are derived from a very low endemic setting and were described in detail by Trauer et al. They consist of 613 infected contacts of whom 67 (10.9%) developed active TB during the study period (Trauer et al., 2016a). To enhance our dataset, we also used the published data on close contacts of pulmonary TB patients from Amsterdam (Netherlands) notified between 2002 and 2011, as reported by Sloot et al. (Sloot et al., 2014). These data include 739 infected individuals, of whom 71 (9.6%) developed active TB. The detailed approaches used to determine both dates of infection and activation in individuals in the two studies are presented in the respective manuscripts. The activation times measured in these data were used to calibrate the different models. In order to validate our approach involving merging of two datasets, we present a comparison of the estimates obtained from the separate fittings to the two datasets (Appendix in Supplementary file Section 8.2). The approach used to extract data from Sloot and colleagues' article is described in detail in the Appendix in Supplementary file, along with a validation analysis of the extraction method while the distribution of the times to activation measured in the two datasets (Victoria and Amsterdam) is presented in Fig. S7 (Appendix in Supplementary file).

Trauer et al. also proposed an imputation method which takes into account the censorship for migration, death, and preventive treatment (Trauer et al., 2016a). We used this approach, which is associated with higher estimates concerning the risk of TB activation, in a supplementary analysis.

2.4. Model fitting

Model fitting to data was made by building the survival likelihood defined as follows. For a given model associated with a given set of parameters, θ, we obtain an analytical survival function $S_\theta(t)$ which represents the probability that activation has not occurred yet at time t given that infection occurred at $t = 0$. This function is associated with a hazard function $\lambda_\theta(t)$ defined by $\lambda_\theta(t) = -S'_\theta(t)/S_\theta(t)$, characterising the chance that progression to active TB occurs at precisely time t, given survival up to that time. Then, for each infected case i of our dataset, for whom t_i designates the time of either TB activation or end of follow-up, we define an individual likelihood component by $\mathcal{L}_{\theta,i} = S_\theta(t_i)$ if the case was not known to develop active TB; and $\mathcal{L}_{\theta,i} = \lambda_\theta(t_i) \times S_\theta(t_i)$ if the case effectively activated TB at time t_i. Finally, we aim to maximise the multi-dimensional likelihood obtained by multiplying all the individual likelihood components together: $L_\theta = \prod_i \mathcal{L}_{\theta,i}$. This problem is equivalent to maximising the following log-likelihood that we define as the fitting score: $FS_\theta = \sum_i \log(\mathcal{L}_{\theta,i})$.

Another fitting method was used for validation and when the data did not allow for the survival likelihood to be utilised. Specifically, a least squares optimisation was performed to minimise the

Please cite this article in press as: Ragonnet, R., et al., Optimally capturing latency dynamics in models of tuberculosis transmission. Epidemics (2017), http://dx.doi.org/10.1016/j.epidem.2017.06.002