***Epidemiologic considerations in age-structured models of tuberculosis transmission***

Snow K, Graham S, Sawyer S, Denholm J

**Target journal:** Am J Epi (Practice of Epidemiology section)

**Word limit:** 4,000

Abstract (200 words, unstructured)

TB remains a disease of great public health significance, and mathematical modelling is increasingly used to estimate the potential impacts of control interventions. Some interventions, notably novel vaccines, would likely be targeted towards specific age groups, however most mathematical models of TB developed to date deal with age either crudely, indirectly or not at all.

This article reviews the age-related epidemiology of TB relevant to the construction of age-structured mathematical models, and demonstrates appropriate age-structured parameterisation of a simple model simulating a vaccination program targeting adolescents and young adults. The output of the age-structured model is compared to that of a traditional model that deals with age dichotomously (0-14 vs. ≥15 years).

[results]

**[Background / methods]**

*TB and TB modelling practice*

Tuberculosis (TB) is now the leading cause of death due to infectious disease worldwide, despite the availability of curative antimicrobial treatments since the 1940s. The epidemiology, clinical presentation, and outcomes of TB change markedly across the life course: probabilities of infection, progression to active disease, transmission of the organism to others, successful treatment, and TB-related death all vary substantially with age. Many public health interventions related to TB age likewise age-sensitive – BCG vaccination is highly efficacious only among infants; preventative therapy is most widely implemented for children under 5 years of age, and active case finding interventions are most often targeted at adults.

Furthermore, as TB epidemics develop and resolve over long time scales, the age-related epidemiology of TB changes together with the age-structure of the source population. High-intensity epidemics are most common in low and middle-income settings with young populations, and high transmission disproportionately impacts children and young adults. In resolving epidemics, however, the burden of disease shifts into progressively older age groups as reactivation disease becomes more prevalent than primary progressive disease. As a result, targeting a given age group for TB control activities will have markedly different efficacy in different countries. This fact differentiates TB from many other infectious diseases such as polio, measles, influenza, and HIV, where the age groups at highest risk are more consistent.

Despite the importance of age as a determinant of both TB risk and control measures, mathematical models of TB transmission rarely deal with age in a nuanced fashion. A majority of TB models developed to date deal with age in a dichotomous manner, with a single step in TB force of infection and/or risk of progression to disease from “childhood” (usually age 0-14 years) into late adolescence and adulthood (age 15 years and above). In reality both of these key parameters, and many others, vary markedly throughout each of these broad age categories.

Depending on the purpose of the specific TB model the dichotomous approach to age may not have material impacts. However in models attempting to estimate the impact of interventions that would in reality be targeted at specific age groups, these over-simplifications may introduce bias. Models are increasingly used to estimate the potential impact and cost effectiveness of novel TB vaccines, of which several are currently in stage 3 clinical trials. It is expected that such vaccines will be targeted at either adolescents or adults, rather than at infants, as young children with TB cause few secondary infections. Young adolescents are a key group of interest, as mass vaccination campaigns can more easily be delivered through schools than to the general adult population – an approach recently adopted in the UK for influenza vaccination. However, because the age-related epidemiology of TB varies markedly with both population age-structure and TB epidemic intensity, the efficiency of targeting a given age group for TB vaccination may vary markedly between settings.

If a novel TB vaccine reaches the market, modelling studies will be needed to identify the optimal age group for vaccination in different settings. The aim of this article is to review the key issues in TB epidemiology in early life that are relevant for TB modellers, in order to inform the development of appropriately parameterised models that may better predict the potential impact of age-sensitive interventions. A demonstration is provided using two simple models to estimate the impact of a novel vaccine targeted at different age groups in a variety of epidemic scenarios.

***Age-related changes in key model parameters***

*Force of infection and effective contact rates*

Force of infection is arguably the most fundamental parameter of a mathematical model of an infectious disease, and the hardest to estimate accurately in the specific case of TB. The force of infection experienced by people living in TB endemic settings is a function of their risk of contact with individuals with infectious TB, together with the intensity and duration of contact experienced. As such, it likely changes considerably with age throughout early life. A parent with infectious TB presents an exceptionally high risk of infection to a small child, with this risk declining as the child ages and contact with parents reduces in both duration and intensity. As children age they come into contact with a broader range of adults, and the nature of this contact changes, as children transition from care in the home to schooling, and to increasingly extensive socialisation outside the home. Adolescents have markedly higher numbers of total respiratory contacts than do either children or adults, and also the greatest number of same-age contacts (ref POLYMOD). Unlike young children, adolescents often develop infectious TB, creating a risk of infection for same age peers.

Directly observing the age-specific forces of TB infection to children and adolescents is challenging. Historical TST surveys have tended to target either a narrow age range (for example, 8-10 year olds), or to pool data from children and adolescents of varying ages to produce an annual risk of infection estimate which is effectively a mean. Uncertainty around age-specific risks of progression and reactivation complicates “backward engineering” of forces of infection from age-specific notification data on active disease in the same age group.

Perhaps the most feasible approach for modellers is to estimate force of infection to children and adolescents based on what is known or suspected about the prevalence of infectious TB in adolescents and adults, together with estimates of the contact rates between the susceptible age group and the infectious age groups (Dodd). Alternatively, where the force of infection in a model is varied as part of the calibration process, the relative frequencies of contact between different age groups should be preserved. In our model the age-specific force of infection is calculated from age-specific contact rates together with the age-specific prevalence of infectious TB.

At present, detailed age-specific contact data are only available from Europe, which may not perfectly capture contact patterns in high TB settings where high school attendance is lower, households are larger, the age difference between children and their parents may be smaller, and patterns of employment differ. However, the bias introduced by using European age-specific contact data to estimate contact rates in low and middle income countries may be less than the bias introduced by ignoring age entirely and using average values for contact rates in two broad age groups. Using average values for “children” and “adults” will likely over-estimate rates of contact between young children and between adults, and under-estimate rates between adolescents (Wallinga).

*Progression to disease*

Rates of progression to disease after infection by age are better characterised than are age-specific rates of infection, although uncertainty increases as age-groups become narrower. Some data suggest that rates of progression decline continually between infancy and adulthood, while others document an apparent transient increase during adolescence, followed by a continued decline. In contrast, many current models assume a single step down in progression that occurs at age fifteen (e.g. assuming a 15% risk of progression among children aged 0-14 and a 5% risk thereafter). Dichotomous approaches to risk of progression will dramatically under-estimate the risk of progression among young children, will over-estimate the risk among young adolescents, and may under-estimate the risk among older adolescents and young adults. Stochastic models should incorporate substantial uncertainty around any estimates of progression in narrow age groups.

In our model, we assumed a risk of primary progression of 30% among children under 5, with 5% absolute declines for every subsequent five year age group, with risk stabilising at 5% in 25-49 and 50 + year olds (based on data from Comstock, Trauer, etc).

*R0*

Young children only rarely develop cavitating pulmonary TB that presents a meaningful risk of transmission to others. Cavitation becomes increasingly common through late childhood and adolescence, and by age 15-19 adolescents appear as likely as adults to experience smear-positive disease (Cruz, me). This has implications for both the public health impact and the cost-effectiveness of age-specific interventions, as improved TB control among older adolescents and adults will prevent greater numbers of secondary cases than would interventions that prevent TB in childhood (although conversely, TB prevention among children will often avert more deaths). It should be noted that while older adolescents and adults may be equally likely to develop cavitating disease, the greater number of respiratory contacts made by adolescents implies that the age-specific R0 will be greater for older adolescents than for adults in the same epidemics.

*Risk of adverse TB outcomes*

Access to TB diagnosis varies with age. Children with TB are less likely to be diagnosed than adults, and even when diagnosed, may be less likely to be reported (Lestari, du Preez). In some settings, national prevalence surveys have documented differential access to diagnosis by age-group among adolescents and adults, with access to diagnosis improving with increasing age (ref Philippines). Mortality amongst children, adolescents or adults who remain undiagnosed will be very considerable, likely largely irrespective of age, although young children are at particularly high risk.

In the subset of patients who are successfully diagnosed, treatment outcomes are also known to vary with age. Elevated risk of premature treatment discontinuation has now been documented among adolescents in several settings, and difficulties in supporting medication adherence in adolescents with other chronic conditions have long been recognised. In contrast, TB-related mortality will usually be much rarer in adolescents on TB treatment than in young children or among the elderly (ref Philippines, du Preez, Mori). Age-specific treatment outcome data are increasingly available owing to the implementation of case-based electronic surveillance data, and could be used by modellers to incorporate age- and setting-specific probabilities of recovery and mortality on treatment in some settings.

Off-treatment mortality was assumed to be 75% per year among children aged 0-4 years and 50% per year in other age groups. For simplicity, on-treatment mortality was assumed to be equal to background mortality.

*HIV prevalence and incidence*

HIV status is a key determinant of age-specific TB risks, particularly risks of progression to disease and of adverse treatment outcomes. Age-stratified HIV prevalence estimates are sometimes available for children, adolescents, and adults in the same setting, and should be incorporated into models for high HIV prevalence settings where possible. It should be noted that age-specific prevalence in young children and to a lesser extent adolescents may change much more quickly than adult prevalence in the same setting, owing to the impact of prevention programs and the more rapid turnover of the population in narrower age groups.

*Aging populations*

TB epidemics develop resolve over much longer time scales than those of any other respiratory infectious disease. Rising wealth tends to cause both a shift in population age structures due to falling birth rates, lower child mortality, and increased life expectancy, and concurrent declines in TB incidence. Where high quality surveillance data are available for extended periods of time, it can be shown that birth cohort effects emerge as TB epidemics decline: TB risk falls more quickly among younger generations, with older generations continuing to experience high rates of reactivation disease long after actual TB transmission has fallen (Mori, Morabia). Consequently, population aging can arrest declines in per capita TB incidence, as the older generations at continued high risk constitute an increasing proportion of the total population. For models that run over periods of 20 years or more, these demographic shifts and differential declines in TB risk may be relevant, particularly if combined with substantial changes in HIV incidence over the same time period. As such, incorporating projected changes in the birth rate may have material impacts for models with long run times.

*Conceptual summary*

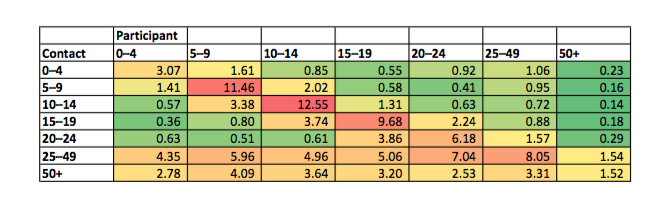
The public health impact and cost-effectiveness of age-sensitive TB interventions cannot be modelled accurately without accounting for age-specific changes in key parameters. Models that do not deal with age appropriately may either over- or under-estimate the impact of age-sensitive interventions, depending on the nature of the TB epidemic in the target setting. This will be particularly relevant for models attempting to estimate the potential impact and cost effectiveness of novel vaccines, which are likely to be targeted at young adolescents. To date, a large majority of models attempting to estimate the impact of novel vaccines have not handled age-related changes in force of infection or risk of progression in a nuanced fashion, but have rather treated the values of these and other key parameters as dichotomous.

***Demonstration Models***

Two Susceptible-Exposed-Infectious-Susceptible (SEIS) models were developed: 1) a dichotomous model stratifying the population into 0-14 years and ≥15 years, and 2) an age-stratified model considering the age groups 0-4, 5-9, 10-14, 15-19, 20-24, 25-49 and ≥50 years. Both models were set up with an initial population of 30 million people, with the age structure and background mortality rates based on data for the UK from the period 1900-1939. Births were set to equal deaths in order to keep the population approximately stable in size.

Age specific contact rates were based on casual contact data from the POLYMOD study, and adjusted downward by a constant factor across all age groups until TB incidence stabilised in the region of 800, 400, and 200 per 100,000py after a run-in period of 100 years.

Table X. Contact matrix values

  
Table X. Biologic parameter values in the age-structured model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **0-4** | **5-9** | **10-14** | **15-19** | **20-24** | **25-49** | **50+** |
| Primary progression rate | 0.30 | 0.25 | 0.20 | 0.15 | 0.10 | 0.05 | |
| Proportion infectious | 0.05 | 0.10 | 0.25 | 0.50 | | | |
| TB-mortality rate (untreated) | 0.75 | 0.50 | | | | | |

∂Sa = + birth rate + recovery rate\*(Ia) – lambda\*(Sa) – non-TB mortalitya\*(Sa) – aging (Sa)

∂La = + lambda\*(Sa) – progression rate\*(Ea) – non-TB mortalitya\*(Ea) – aging (La)

∂Da = + progression rate\*(Ea) – recovery\*(Ia) – non-TB mortalitya\*(Ia) – TB mortalitya\*(Ia) – aging (Ia)

Where

Lambdaa = ∑(Sa\*ßja\*(proportion infectiousj)\*(Ij))

Progression ratea = (progressiona)\*(Ea)

Recovery ratea = (diagnosis ratea \* treatment success ratea)\*(Ia)

TB mortality rate = ((treatment initiation ratea\*non-TB mortality ratea) +

((1-treatment initiation ratea)\*(TB mortality ratea+ non-TB mortality ratea)))\*(Ia)

The compartments were initially populated based on TB notification data for the UK in 1913 – Infectious compartments were populated according to the number of TB notifications in each age group that year, with Exposed compartments populated using the reciprocal of the age-specific primary progression rate to reflect recent infections, and the remainder of the population allocated to the Susceptible compartment.

The dichotomous model was set up in the same fashion, using weighted averages of all parameters for the narrow age bands to set the parameters for the <15 and ≥15 year age groups.

***Analysis***

Infections and cases attributable to people with active TB in each age group were calculated using next generation using the model’s output for year 101 of the epidemic. R0 for infections and new cases was calculated by diving the result of the next generation matrix of the number of infectious individuals in year 100. The contribution of each age group was compared to the 25-49 year age group by calculating the ratio of the comparison group’s R0 to that of the 25-49 year group.

**Results**

*Initial conditions*

The age structured model reached stable incidence around 350 per 100,000 when contact rates were adjusted by a factor of 0.25.

Figure X. Age specific incidence and prevalence of TB according to each models

Table X. Next generation matrices for age groups in each model

When the same conditions were imposed on the dichotomous model.

*Next generation matrices*

**Discussion**