Epidemiol. Infect. (1998), **121**, 309–324. Printed in the United Kingdom

The long-term dynamics of tuberculosis and other diseases with long serial intervals: implications of and for changing reproduction numbers

E. VYNNYCKY* AND P. E. M. FINE

Infectious Disease Epidemiology Unit, Department of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, England

(*Accepted 18 May 1998*)

SUMMARY

The net and basic reproduction numbers are among the most widely-applied concepts in infectious disease epidemiology. A net reproduction number (the average number of secondary infectious cases resulting from each case in a given population) of above 1 is conventionally associated with an increase in incidence; the basic reproduction number (defined analogously for a 'totally susceptible' population) provides a standard measure of the 'transmission potential' of an infection. Using a model of the epidemiology of tuberculosis in England and Wales since 1900, we demonstrate that these measures are difficult to apply if disease can follow reinfection, and that they lose their conventional interpretations if important epidemiological parameters, such as the rate of contact between individuals, change over the time interval between successive cases in a chain of transmission (the serial interval).

The net reproduction number for tuberculosis in England and Wales appears to have been approximately 1 from 1900 until 1950, despite concurrent declines in morbidity and mortality rates, and it declined rapidly in the second half of this century. The basic reproduction number declined from about 3 in 1900, reached 2 by 1950, and first fell below 1 in about 1960. Reductions in effective contact between individuals over this period, measured in terms of the average number of individuals to whom each case could transmit the infection, meant that the conventional basic reproduction number measure (which does not consider subsequent changes in epidemiological parameters) for a given year failed to reflect the 'actual transmission potential' of the infection. This latter property is better described by a variant of the conventional measure which takes secular trends in contact into account. These results are relevant for the interpretation of trends in any infectious disease for which epidemiological parameters change over time periods comparable to the infectious period, incubation period or serial interval.

INTRODUCTION

Tuberculosis has declined steadily in Western Europe and North America for at least a century, though this trend was arrested temporarily in several countries during the World Wars and during the late 1980s [1]. Although long-term trends in developing countries have not been well documented, many populations have experienced dramatic increases in tuberculosis since the 1980s, attributable largely to HIV [2]. The future implications of these recent increases, which led WHO to declare tuberculosis a global emergency in 1993, are not known.

* Author for correspondence.

Epidemiologists often reduce the problem of predicting disease trends to one of estimating reproduction numbers. The reproduction number concept is derived from demography. It was first applied to infectious diseases by Macdonald [3, 4] who defined a reproduction 'rate' as the average number of secondary infectious cases resulting from each (infectious) case in a given population. Recent workers have pointed out that the statistic is a simple number and not a rate *per se* [5, 6]. Increases or decreases in infection or morbidity incidence are assumed to reflect the current magnitude of the 'net reproduction number' (i.e. whether or not it exceeds one). For infections which lead to any immunity, the number of secondary cases to result from an infectious case will be maximum if she/he is introduced into a 'totally susceptible' population. This limiting value, often referred to as the 'basic reproduction number' [6–8] thus provides a measure of the 'transmission potential' of an infection under 'ideal' conditions.

Although the reproduction number concepts are simple to apply to acute infections, their application to tuberculosis is complicated by three particular features of the natural history of this disease: *first*, immunity after an exposure is not solid, and individuals can be reinfected [9–11]; *second*, the time period between infection and disease (the incubation period) and thus between successive cases in a chain of transmission (the serial interval [12]) can be very long, i.e. as long as a lifetime, and depends on age [13]; and *third*, the fact that important epidemiological factors, such as the nature and probability of contact between individuals have changed appreciably over periods comparable to the serial interval of the disease [14]. Hitherto, the only analyses of the basic and net reproduction numbers for tuberculosis have been based on a substantial simplification of its basic natural history, and have neglected to consider these features [15, 16].

This paper considers the implications of the complex natural history of tuberculosis for the derivation and interpretation of its reproduction numbers, and hence for the problem of assessing long-term trends of the disease. In so doing, it illustrates important and counter-intuitive properties of the conventional reproduction number measures. These are relevant for many other infections for which either the incubation period is long or reinfection can occur. We present the logic based on a dynamic transmission model of tuberculosis.

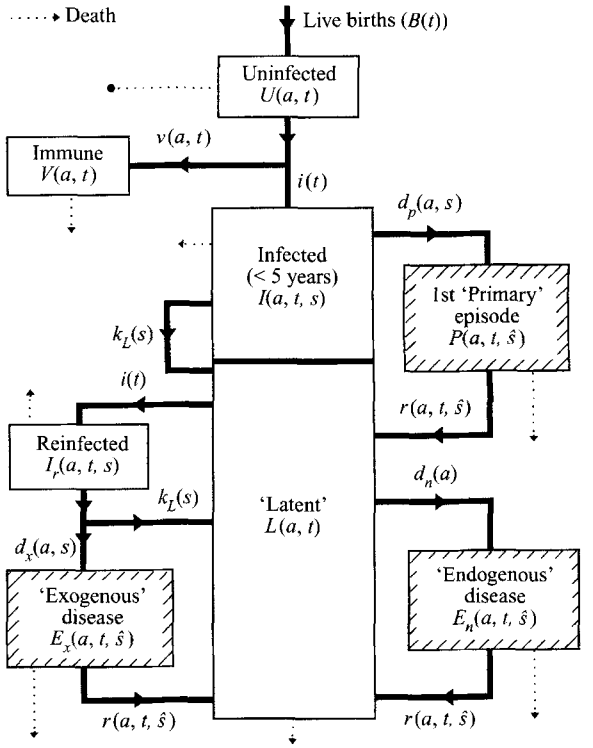

Fig. 1. Schematic diagram of the model. Primary disease is defined as disease within 5 years of initial infection [21]; exogenous disease is defined as the first disease episode within 5 years of the most recent reinfection event. Endogenous disease includes disease occurring more than 5 years after the most recent (re)infection event, and *second or subsequent* disease episodes occurring less than 5 years after the most recent (re)infection event.

METHODS

Structure of the model

The model explored here (which extends the classic work of Sutherland and colleagues [17] on adult tuberculosis in the Netherlands) describes the transmission dynamics of all forms of pulmonary tuberculosis in England and Wales since 1900 [11]. To avoid the complications of gender differences [18], and the contribution of immigrants and the HIV epidemic to the recent epidemiology of tuberculoses [19, 20], the analyses described here are restricted to the white male population in the absence of HIV. The model's structure is shown in Figure 1.

Individuals are born into the uninfected category, $U(a, t)$. Infected individuals are divided into those who have not yet developed primary disease (defined by convention as disease within 5 years of initial infection [21] ($I(a, t)$) and those at risk of endogenous reactivation and/or of reinfection followed by exogenous disease (see definitions in Fig. 1). The risks of

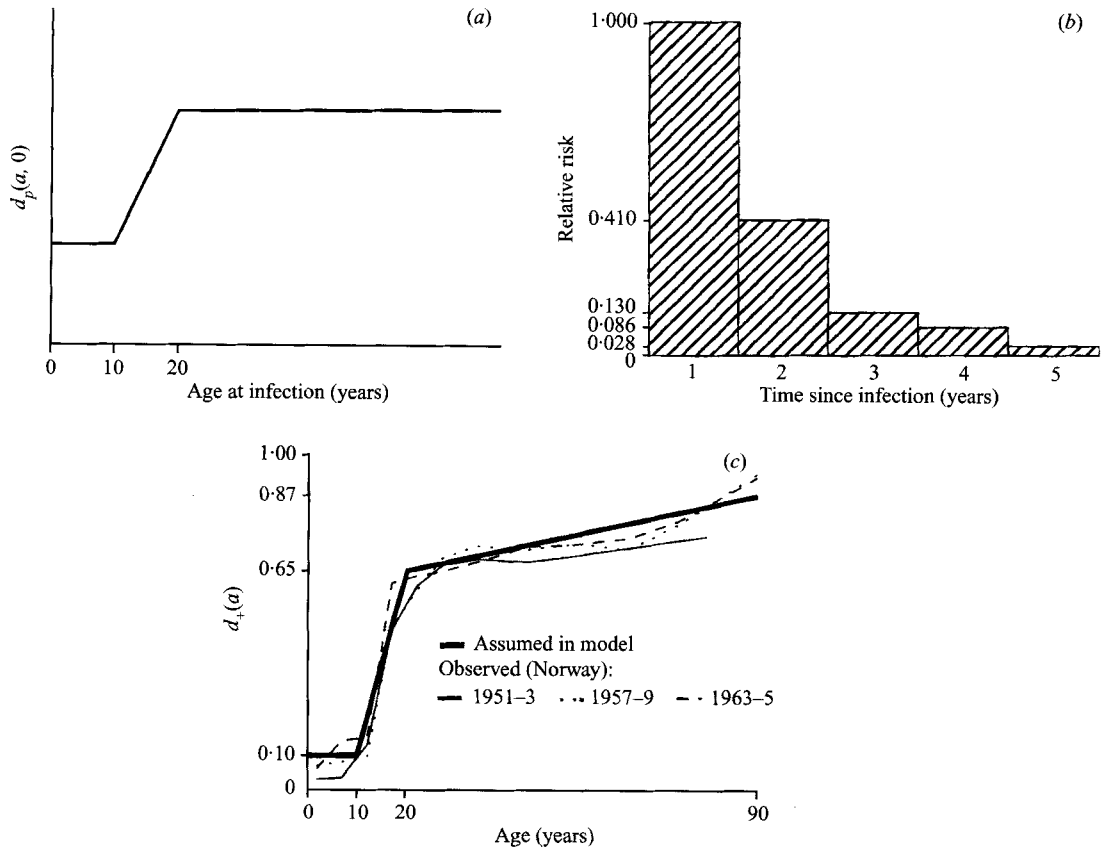

Fig. 2. Summary of main assumptions in the model relating to the risks of developing disease. (*a*) General relationship between the risk of developing the first primary episode (during the first year after infection) and age at infection. An identical relationship is assumed to hold between the risk of exogenous disease and the age at reinfection and between the risk of endogenous disease and the current age of individuals. See Table 1 for the magnitude of the disease risks. (*b*) Risk of developing the first primary episode (or exogenous disease) in each year following initial infection (or reinfection), relative to that experienced in the first year after infection (or reinfection). The relationship was derived using data from the UK MRC BCG trial during the 1950s [43]. (*c*) Proportion of respiratory disease incidence manifested as sputum-positive (i.e. infectious). (Data source: Dr K. Styblo (TSRU) and Dr K. Bjartveit (Norwegian National Health Screening Service).)

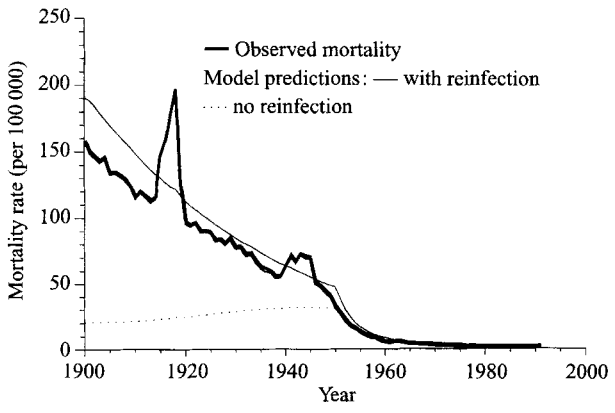

Fig. 3. Comparison between the observed average age-standardized mortality rates of respiratory tuberculosis in England and Wales in males since 1901, against those predicted by the model, assuming that reinfection did or did not occur, using the best-fitting disease risks for the corresponding assumption [14]. (The risks of disease based on the assumption that reinfection did not affect disease reinfection and of first infection ($i(t)$) are assumed to be identical, and to depend only on calendar year. However, reinfection is less likely to lead to disease than is initial infection, due to some immunity induced by prior infection [11]. We also assume that individuals cannot be reinfected whilst at risk of developing either the first primary episode or exogenous disease (i.e. within 5 years of an initial infection or reinfection event).

The age of an individual is assumed to determine the risks of developing disease (Fig. 2*a*) and the probability that the episode is infectious (sputum smear/culture-positive) (Fig. 2*c*) [11]. The risks of

risks were: 2·53% and 13·6% for individuals during the first year after infection at age 0–10 years and over 20 years respectively; annual risk of developing endogenous disease: $4{\cdot}72 \times 10^{-4}$% and 0·0416% for 0–10 year olds and those aged over 20 years respectively [11]).

Table 1. *Summary of parameter values used in the model*

Variable	Definition	Assumption
$i(t)$	Infection and reinfection rates at time t.	20% until 1880, declining by 2% pa until 1901, by 4% pa until 1950 and 13% pa thereafter [44].
$v(a, t)$	Proportion of uninfected individuals of age a immunized at time t.	Vaccination introduced in 1954 and restricted to 13 year olds. Vaccine efficacy assumed to be 77% and vaccine coverage increasing to approximately 80% since 1960 [11].
$d_p(a, s)$	Risk of developing the first primary episode at time s after infection at age a.	Declines with time since initial infection (Fig. 2*b*). Cumulative risks within 5 years of initial infection: 4·06%, 8·98% and 13·8% for 0–10 year olds, 15 year olds and individuals aged over 20 years respectively [11].
$d_x(a, s)$	Risk of developing exogenous disease at time s after reinfection at age a.	Declines with time since reinfection (Fig. 2*b*). Risks within 5 years of reinfection: 6·89%, 7·57% and 8·25% for 0–10 year olds, 15 year olds and individuals aged over 20 years respectively [11].
$d_n(a)$	Annual risk of developing endogenous disease at age a.	$9{\cdot}82 \times 10^{-8}$%, 0·0150%, and 0·0299% for 0–10 year olds, 15 year olds and individuals aged over 20 years respectively [11]
$d_+(a)$	Proportion of total disease incidence among cases aged a assumed to be infectious.	10% for 0–10 year olds, increasing linearly to 65% for 20 year olds and increasing linearly to 85% for 90 year olds (Fig. 2*c*).
$k_L(s)$	Rate at which individuals who have been infected or reinfected for time s without developing disease move into the 'latent' class.	Transition occurs exactly 5 years after infection/ reinfection. i.e. $k_L(s) = 0$ if $0 < s < 5$ and ∞ for $s = 5$ years.
$r(a, t, \hat{s})$	Recovery rate for cases of age a at time t at time $\hat{s}$ after disease onset.	Individuals are diseased for 2 years unless they die in the meantime (see below).
$m_+(t, \hat{s})$	Case-fatality of infectious pulmonary cases at time t and time $\hat{s}$ since disease onset.	Case fatality in second year after disease onset is 65% of that in first year. Overall case-fatality: 50% until 1950, declining to 30% and 25% by 1953 and 1956 respectively, and constant until 1976. Identical to mortality in general population thereafter [11].
$m_g(a, t)$	Mortality rate of non-infectious and non-diseased individuals in the general population of age a at time t.	Identical to all-cause mortality (after subtracting deaths among infectious cases, estimated in the model). Annual age-specific all-cause mortality rates obtained from Government's Actuary's Department since 1841. Data until 1841 obtained by back-extrapolation.

developing the first primary episode and exogenous disease depend also on the time since infection and reinfection respectively (Fig. 2*b*). These risks have been estimated by fitting model predictions to data on disease incidence in England and Wales [11]. The case-fatality rate in the prechemotherapy era, following data from well-known studies [22], is assumed to be 35% and 23% during the first and second year after disease onset, which corresponds to an overall case-fatality rate of 50%.

For reference, Table 1 summarizes the parameters and the assumptions used in the model. We have shown elsewhere [11] that the model's predictions closely mimic the pattern of tuberculosis in England and Wales during this century (see also Fig. 3).

Complications in the net and basic reproduction numbers for tuberculosis

The net reproduction number is defined formally as *the average number of secondary infectious cases resulting from each infectious case in a given population.* The formal definition of the basic reproduction number is analogous: *the average number of secondary infectious cases resulting from the introduction of a typical infectious case into a [totally] susceptible and demographically stable population* [6–8]. The natural history of tuberculosis introduces complications into the derivation and interpretation of these measures.

First, the fact that immunity after infection with the tubercle bacillus is not solid means that both uninfected and infected individuals are 'susceptible'

to infection, though to different degrees. Thus the net reproduction number depends on the number of individuals infected *or reinfected* by each case. As the risk of developing disease following *reinfection* appears to be lower than that after *initial infection* [11, 17], a 'totally susceptible population' in the definition of the basic reproduction number is best defined for tuberculosis as one containing *no* individuals who have ever been infected with *M. tuberculosis*.

Second, the basic and net reproduction numbers depend both on the *age structure* of the population under study, since the risks of developing disease are age-dependent (higher among adults than for children) [11, 23–26], and, given the long time interval between infection and disease onset [13], on *age-specific mortality rates* of the population, which determine the proportion of individuals infected by a given case who will survive long enough to develop disease.

The dramatic decline in the degree of contact between individuals, which resulted from many social changes and interventions over recent history [14], further complicates the basic reproduction number for tuberculosis. This decline is shown in Figure 4*c*, which presents the 'effective contact number', or the average number of individuals 'effectively' contacted by each infectious case (in each year), where an effective contact is defined following Frost [27], as one sufficient to lead to infection of a 'susceptible' individual. This decline implies that the basic reproduction number itself must have declined over time.

A second complication arises because the basic reproduction number is conventionally interpreted as the true 'transmission potential' of infection in an ideal ('totally susceptible') population, but is derived assuming that *no changes in epidemiological conditions occur over time* [7]. For tuberculosis, the fact that the effective contact number declined over a period comparable to the serial interval means that the basic reproduction number in England and Wales for a given year, e.g. 1900, fails to reflect the transmission potential of infection at that time. For example, the decline in the effective contact number affected both (*a*) the number of individuals infected by a given case,

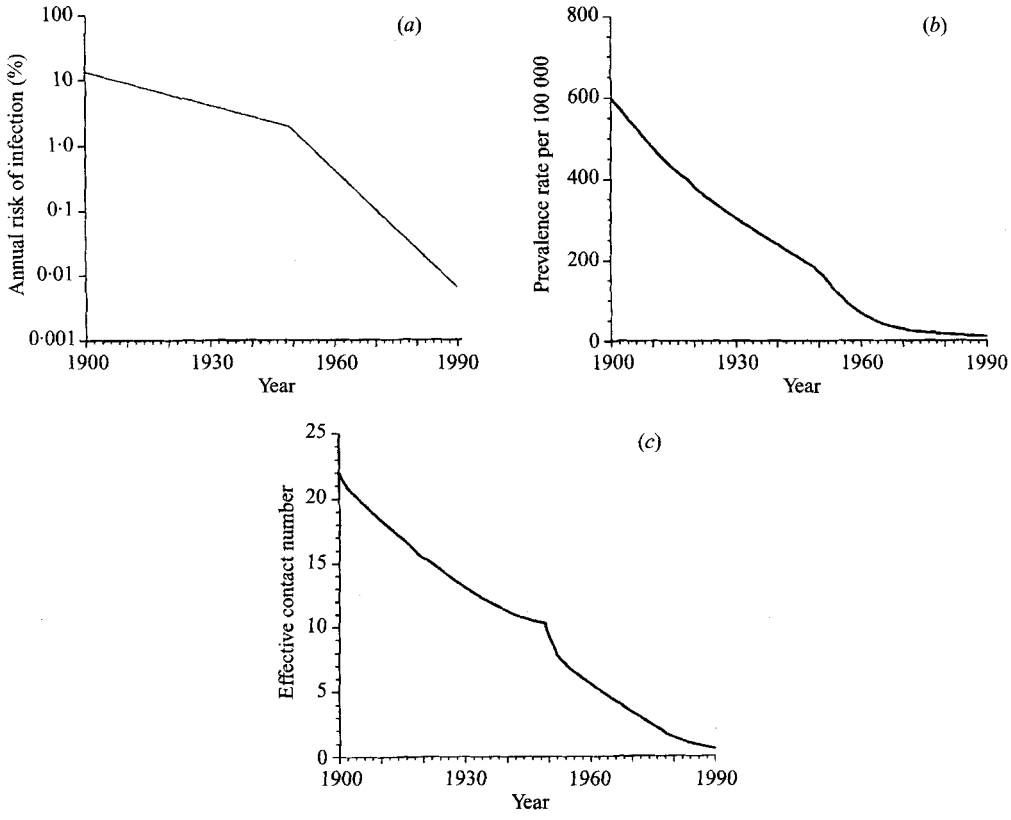

Fig. 4. Estimating the effective contact number in England and Wales since 1900. (*a*) The annual risk of infection (derived using tuberculous meningitis statistics and tuberculin sensitivity data [44]). (*b*) Prevalence of infectious pulmonary cases, derived using model predictions. (*c*) The estimated effective contact number, derived as the ratio between the annual risk of infection and the prevalence of infectious cases [14].

which would have changed even during the infectious period of a given case, and (*b*) whether those initially infected experienced further reinfections and hence, if subsequently diseased, could be accepted as 'secondary' to that case. We now examine this formally.

Definitions of the reproduction numbers

In these analyses, we define the *net* reproduction number as the **average number of secondary infectious cases resulting from either infection or reinfection by a *typical* infectious pulmonary case in a given population under any given conditions.**

Following the above discussion, we distinguish between the conventional ('simple') basic reproduction number and a variant of this measure, which we call the 'ultimate basic reproduction number'. Both are assumed to depend on the demography of the population considered. The 'simple basic reproduction number' is here defined formally as: **'The average number of secondary infectious cases resulting from a "typical" infectious pulmonary case following his/her introduction into a given population, in which no individual has hitherto been infected with the tubercle bacillus, and *assuming that no subsequent changes in epidemiological conditions occur.*'**

The ultimate basic reproduction number is defined analogously, except that it allows for changes in epidemiological parameters over time, and, is thus interpretable for any given year as the actual transmission potential of the infection in that population. Its formal definition is thus: **'The average number of secondary infectious cases resulting from a "typical" infectious pulmonary case following his/her introduction at a particular time into a population in which no individual has hitherto been infected with the tubercle bacillus, and *allowing for subsequent changes in epidemiological conditions.*'**

The derivation of these measures is described below.

Derivation of the net reproduction number for tuberculosis

Using the model described above, we first estimated how the age distribution of individuals infected and reinfected at a particular time, and the age-dependent lifetime risks of developing infectious pulmonary tuberculosis attributable to a given infection/reinfection event, have changed during this century in England and Wales. The net reproduction number in England and Wales for each year t since 1900 was then derived directly from the model, by taking the ratio between (*a*) the *total* number of individuals who developed disease attributable to infection or reinfection by cases with disease onset in year t, and (*b*) the total number of infectious pulmonary tuberculosis cases with onset in year t (see Appendix).

The fact that reinfection can occur complicates the definition of a secondary infectious case for tuberculosis, as it depends on whether bacilli causing a current disease episode stem from just one or from several (re)infection events (e.g. an individual would be the secondary case of several source cases if the latter holds). Although several studies have provided direct evidence for simultaneous infection with multiple strains of bacilli [28–30], it is not known how frequently this occurs. For simplicity, each disease episode is here attributed to the most recent (re)infection event, i.e. the bacilli causing the current disease episode are assumed to be those from the most recent (re)infection event. In order to assess how this assumption affects our conclusions about the relationship between the magnitude of the net reproduction number and trends in morbidity, we compared trends in morbidity and in the net reproduction number predicted by the model in two ways, including and excluding the assumption that reinfection can occur. Inclusion of reinfection-attributable disease greatly improves the fit between the model's predictions and the magnitude and trend in observed mortality rates [14] (Fig. 3) and the notification rates after 1950 (shown in detail in reference [11]).

Derivation of the basic reproduction numbers

Both the simple and ultimate basic reproduction numbers for tuberculosis in England and Wales since 1900 were estimated by direct simulation. For a given year t, we first simulated the introduction of a single infectious pulmonary 'founder' case into an uninfected population, which was identical to that of England and Wales at time t with respect to its age-structure, age and time-specific mortality rates, and effective contact number. We then summed the number of individuals infected by this case who developed disease *attributable to that infection* during their lifetime (truncated at 100 years), assuming either that no further changes in contact occurred (for the *simple* basic reproduction number) or that its decline matched that estimated (Fig. 4) (for the *ultimate* basic reproduction number).

The founder case is assumed to be infectious for 2 years. This reflects an average infectiousness duration, though it is recognized that some individuals may experience several disease episodes, and thus have a long and intermittent infectious period. Each disease episode is attributed to the most recent (re)infection event in these simulations (which allow reinfection-attributable disease), and no individuals are assumed to be vaccinated. For simplicity, when considering events after 1990, we have assumed that the effective contact number, the number of births and age and time-specific mortality rates remain at their 1990 levels.

The system of differential equations describing the disease dynamics following the introduction of a founder case into an uninfected population is identical to that presented elsewhere [11], except that the risk of infection in a given year t ($i(t)$) is here the product of the prevalence of infectious pulmonary cases and the effective contact number in that year. These equations were programmed in the 'C' programming language using forward Euler differencing with time steps of 1 year in both age and time [31].

RESULTS

Analyses of the net reproduction number

Figures 5*a*,*b* demonstrate how the age distribution of individuals initially infected or reinfected by each infectious case, and the proportion of these individuals who ultimately develop infectious pulmonary disease attributable to a given infection/reinfection event, have changed during this century in England and Wales. In 1900, most primary (i.e. initial) infections occurred in individuals aged under 15 years, most of the reinfections occurred among adults (Fig. 5*a*) and there were more reinfections than primary infections. By 1960, approximately half of all transmissions constituted new infections (mainly among 0–24 year olds). The estimated lifetime risks of developing infectious pulmonary tuberculosis following (re)infection, in absence of further reinfection, are age-dependent (lower for infants and young children than among adults – see Fig. 5*b*). They increased slightly with calendar year of (re)infection (Fig. 5*b*), as a consequence of lengthening life expectancy and reductions in (subsequent) reinfection risks.

Figure 6*a* compares the disease incidence predicted assuming that disease either could or could not arise following reinfection. Disease incidence derived using the 'full' model (assuming that disease could result from reinfection) follows a similar downward trend to the mortality rates (Fig. 3). In contrast, the incidence derived assuming that disease could not arise from reinfection would have *increased* at least until 1950. This is explained in the Discussion.

Figure 6*b* shows the implications of these assumptions for the net reproduction number. The net reproduction number derived using the full model is estimated to have been slightly above one until the 1930s (despite the decline in morbidity), with little change until 1950, after which it rapidly declined. In contrast, if it is assumed that disease could not result from reinfection, the net reproduction number would have declined steadily between 1900 and 1950, and the decline would have accelerated from 1950. Under these circumstances the net reproduction number would have been below one from the 1920s, even though an increasing disease incidence would have been observed until the 1950s.

Analyses of the basic reproduction numbers

Figure 7 portrays the trends estimated for the net reproduction number and for the simple (R_0) and ultimate (R_u) basic reproduction numbers. R_u is estimated to have exceeded the simple basic reproduction number slightly until about 1950. Both declined more rapidly than the net reproduction number until about 1950 (e.g. from about 3 in 1900 to about 2 by 1950). The decline in all these parameters is estimated to have accelerated appreciably from 1950.

Figures 8 and 9 illustrate the implications of the simple and ultimate basic reproduction numbers for the likelihood of continued transmission of infection following the introduction of an infectious pulmonary case into an uninfected population similar in structure to that in England and Wales during this century. As shown in Figure 8*a*, the incidence of infectious pulmonary disease would have peaked (at 1200 per 100000) about 25 years after introduction of the founder case in 1900, had there been no subsequent decline in the effective contact number. In contrast, the peak would have been lower (i.e. approximately 450 per 100000 individuals) and occurred later (about 30 years after his/her introduction) had the effective contact number subsequently declined (Fig. 8*b*). The disease incidence would have declined rapidly following the introduction of the founder case in 1960,

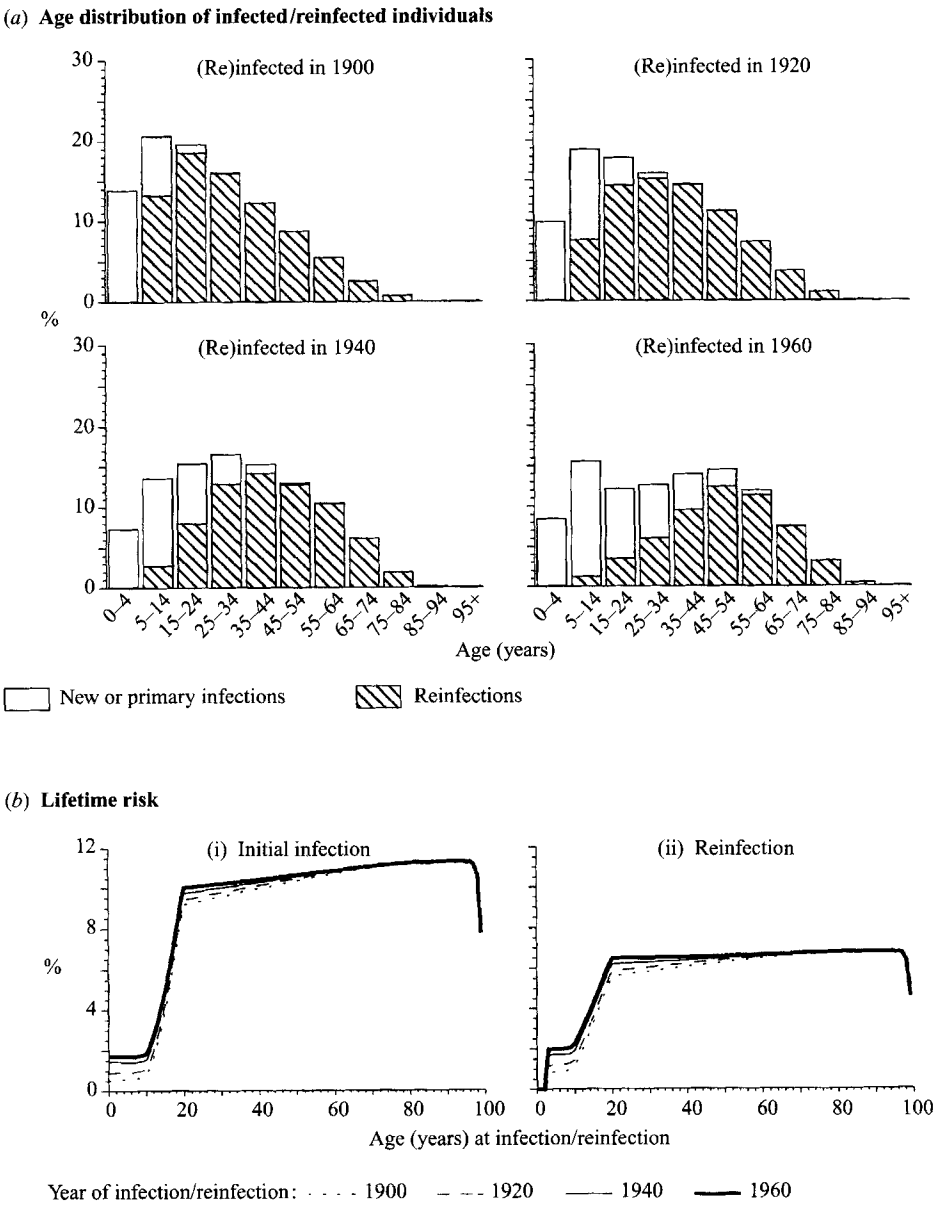

Fig. 5. (*a*) Predicted age-distribution of individuals infected or reinfected in 1900, 1920, 1940 and 1960. (*b*) Estimated age-specific lifetime risks of developing infectious pulmonary tuberculosis attributable to (i) *initial infection* and (ii) *reinfection* experienced in 1900, 1920, 1940 and 1960.

irrespective of whether there was any further decline in the effective contact number (Figs 8*a*, *b*).

The risk of infection would have been higher following the introduction of an infectious case into an uninfected population if the effective contact number had remained constant (at the level of the year of introduction) than if it subsequently declined (Fig. 9). It would have peaked at an unrealistically high value (i.e. 42%) about 30 years after the introduction of the founder case in 1900 had there been no decline in the effective contact number, and would have stabilized at about 20%. In contrast, the peak (at about 8%) would have occurred about 40 years after the introduction of the founder case under the conditions of a declining contact number (as in Fig. 4*c*).

DISCUSSION

Recent increases in tuberculosis in many countries have led the World Health Organization to declare the disease a 'global health emergency'. Given the

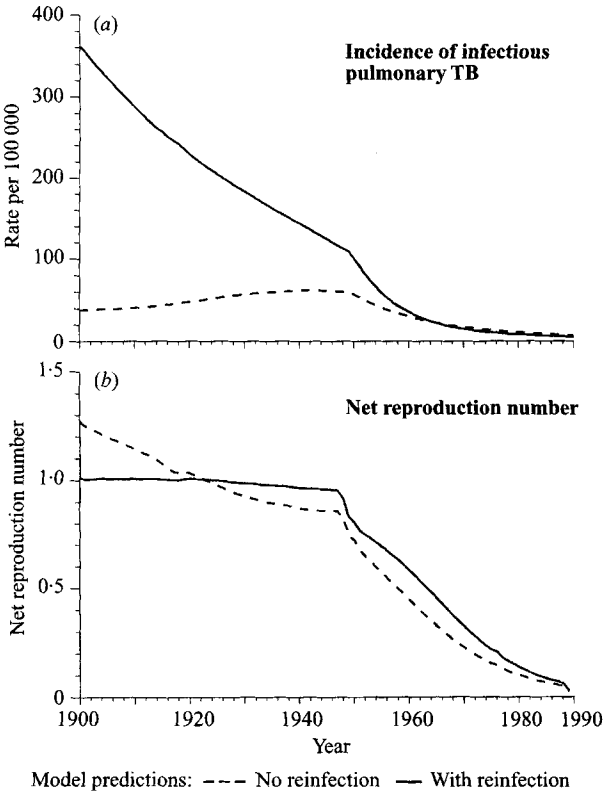

Fig. 6. (*a*) Estimated incidence of infectious pulmonary disease, and (*b*) net reproduction number for tuberculosis, derived assuming that disease could and could not occur following reinfection.

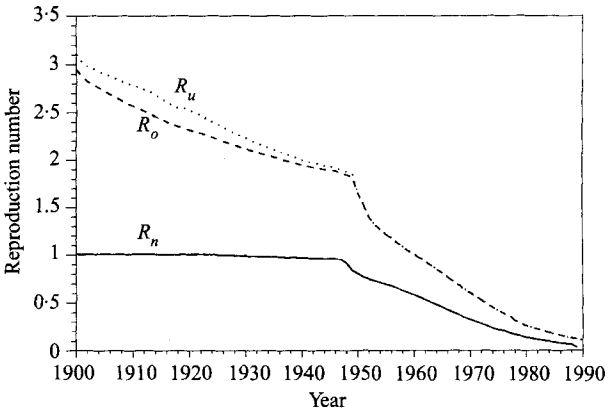

Fig. 7. Comparison between the estimated net reproduction number and the simple and ultimate basic reproduction numbers (R_n, R_0 and R_u respectively) for tuberculosis in England and Wales during the period 1900–90.

immense morbidity and mortality associated with this disease, it is useful to have suitable measures for assessing whether these increases are likely to continue. These analyses provide the first detailed evaluation of the utility of the net and basic reproduction numbers, which are conventionally used for predicting trends in disease incidence and describing the transmission potential of infections, in describing trends in tuberculosis. It is important to recognize that in analysing these measures, we have made several simplifications of the transmission dynamics of *M. tuberculosis*. Among the most important of these is the assumption that individuals mixed randomly in the population. This is unlikely to be realistic for recent years in developed countries, where tuberculosis has increasingly been confined to particular segments of society (e.g. immigrant populations). We have also assumed that individuals develop disease as a consequence of the most recent (re)infection – the implications of which is discussed below. Despite these simplifications, the analyses raise important issues concerning the interpretation of trends in infectious diseases with long serial intervals.

The net reproduction number for tuberculosis

The decline in tuberculosis during the past century in developed countries has been interpreted by some as evidence that each infectious tuberculosis case was leading, on average, to less than one other such case [32], which would imply that the net reproduction number was below one. Our analyses suggest otherwise – that even though the overall morbidity declined during this century, the net reproduction number has been below one only since the late 1920s. This counter-intuitive finding illustrates that neither the definition nor the interpretation of reproduction numbers is straightforward for tuberculosis.

We began by noting that the fact that reinfection can occur complicates the definitions of the reproduction numbers for tuberculosis. The implications of reinfection for reproduction numbers have not been discussed extensively in the literature (except in relation to the evolution of virulence or diversity of pathogens [33–35], though they are crucial for definitions of secondary infectious cases and for interpreting patterns in transmission. For example, an individual would be a secondary case of just one source case if bacilli from just one (e.g. the most recent) infection contributes to each disease episode, but could be attributable to multiple prior cases if bacilli from several (re)infections contribute to each disease episode. The definition of a secondary infectious case is even more complicated for tuberculosis, given that, despite the evidence indicating the importance of reinfection under conditions of high infection risk, the extent to which bacilli from

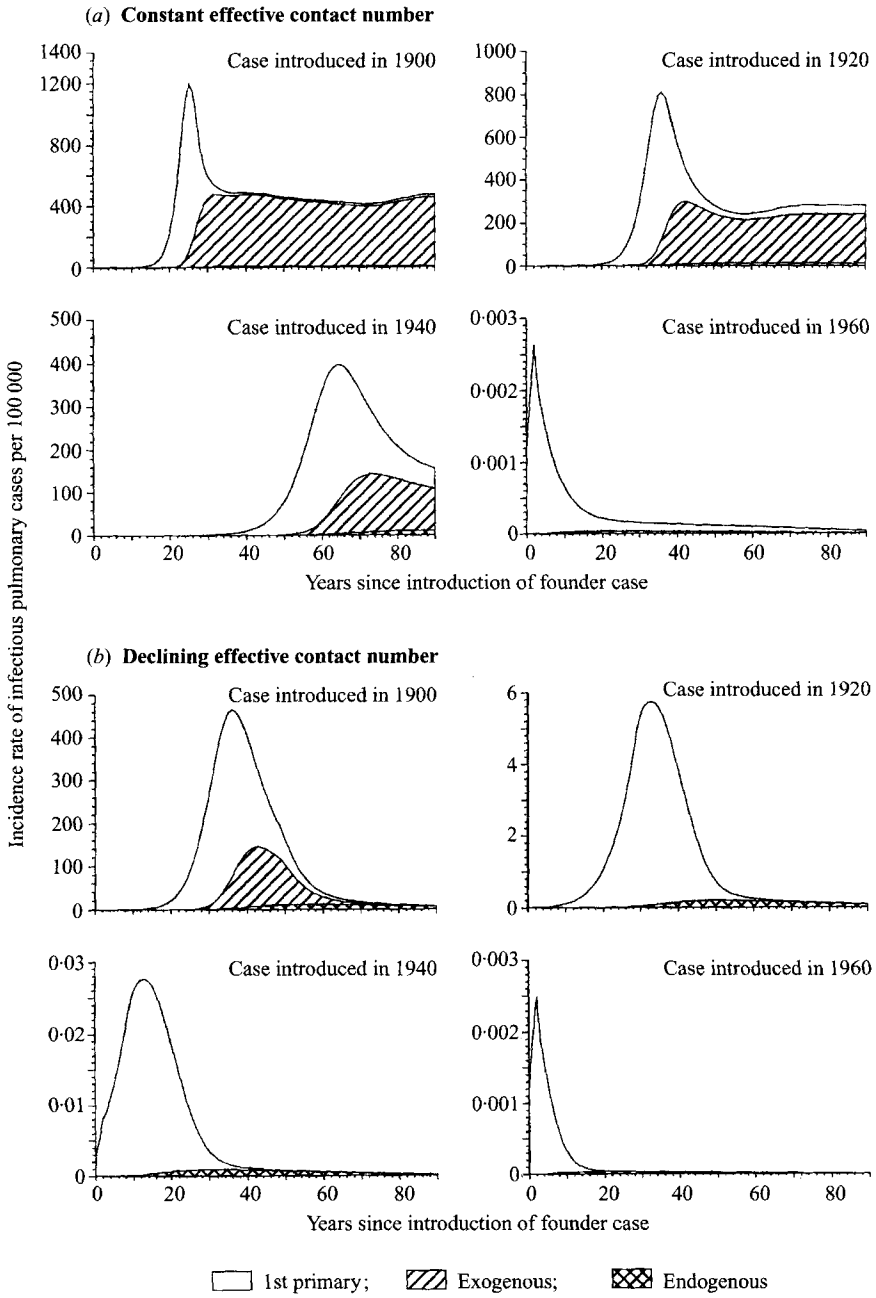

Fig. 8. Predicted annual incidence of infectious pulmonary tuberculosis cases following the introduction of a founder case into a population similar in structure to that in England and Wales in 1900, 1920, 1940 and 1960, assuming that the effective contact number (*a*) did not subsequently decline and (*b*) declined subsequently. Note that the linear scale employed suggests that, in some instances, cases would have first occurred more than 5 years following the introduction of the infectious case in the population. In fact, cases would have occurred immediately after the introduction of the infectious case, though the incidence rate is too low to register on the current scale.

successive reinfection events contribute to a particular disease episode is unknown. Only three studies to date have demonstrated simultaneous infection with multiple strains of tubercle bacilli [28–30], but very few investigators have looked for such evidence. Until more conclusive evidence is provided (e.g. through the application of molecular fingerprinting techniques), our assumption that each disease episode is attributable to the most recent (re)infection appears to be the most appropriate.

The amount of contact between infectious and other individuals is important in determining the

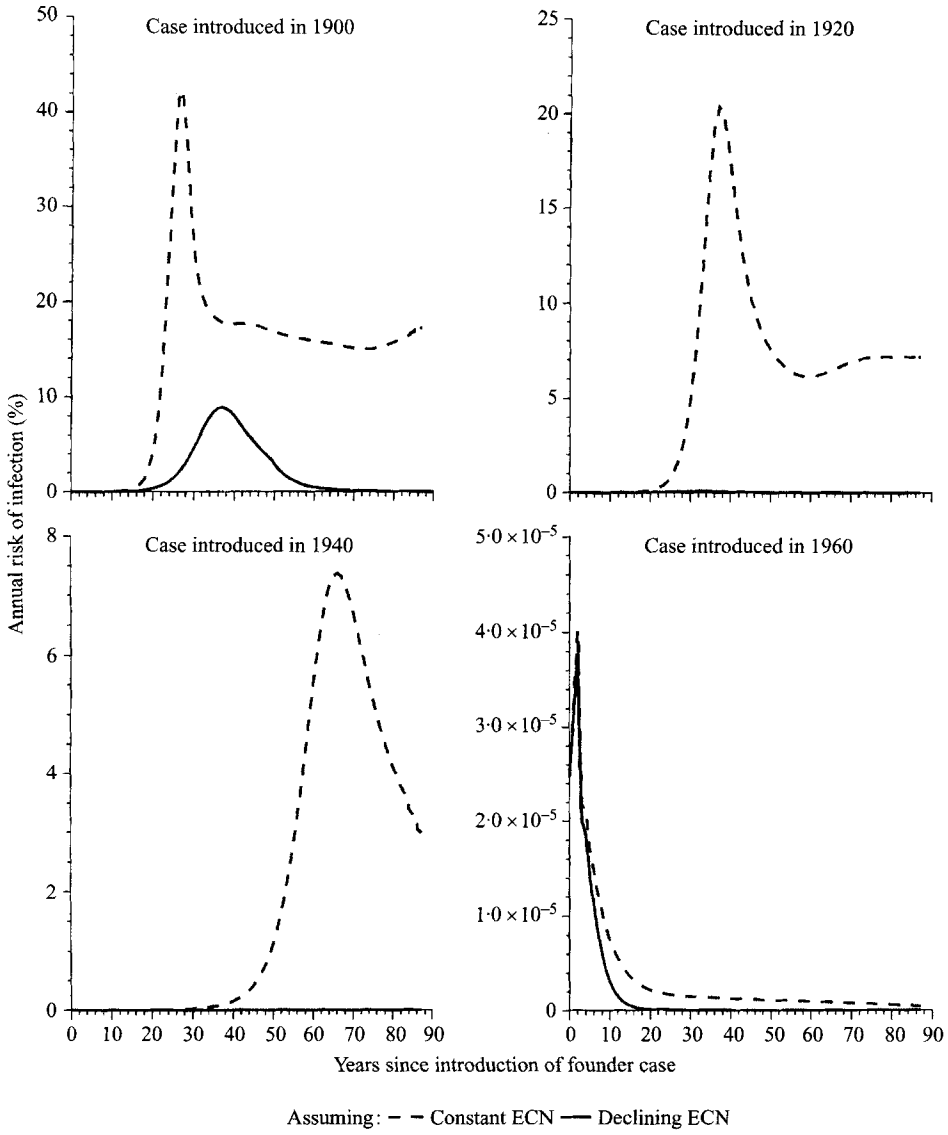

Fig. 9. Predicted annual risk of infection following the introduction of a founder case into a population similar in structure to that in England and Wales in 1900, 1920, 1940 and 1960, assuming that the effective contact number did and did not subsequently decline.

transmission dynamics and reproduction numbers of any infectious disease. Given the dramatic decline in the effective contact number during this century [14] and the declining disease incidence, it is surprising that the estimated net reproduction number scarcely changed until 1950. This result follows from the fact that the *decline* in the number of new infections resulting from each infectious case was compensated by *increases* in the overall lifetime risk of developing infectious pulmonary tuberculosis attributable to each (re)infection event. *These increases occurred because of the decline in the effective contact number, which, in turn, increased the average age at which individuals were either initially infected or were reinfected, coupled with the fact that older individuals have higher risks of developing infectious pulmonary tuberculosis than do younger individuals* (see e.g. Fig. 2*c*).

That the net reproduction number should have been above unity between 1900 and 1930, despite the decline in tuberculosis morbidity throughout this time, is counter-intuitive, as it is conventionally assumed that a net reproduction number of above one translates directly into an increasing disease incidence. The result is attributable to the fact that the effective contact number declined appreciably during the long and variable interval between infection and onset of

infectiousness. Thus of those individuals initially infected, e.g. during the early 1900s, some did not develop disease until many years thereafter, at a time when the effective contact number had fallen very low and when only a small proportion of disease was attributable to recent transmission. It is important to note that this finding is not an artifact of the model's treatment of demography or reinfection, as we discuss below.

First, it is not a consequence of the fact that the population in England and Wales grew rapidly during this century, since both the *numbers* and *incidence rates* of tuberculosis cases are estimated to have decreased over time. It is obvious that in populations growing dramatically over a time period comparable to the serial interval, the net reproduction number need not reflect trends in disease incidence *rates*. Secondly, it is not an artifact of the assumption that each disease episode is attributable to the most recent (re)infection event (i.e. the bacilli from the most recent reinfection event are those causing the current disease episode), as a similar paradox arises when we estimate the disease incidence and net reproduction number assuming that reinfection did *not* lead to disease in the past (Fig. 6). During the period 1920–50, for example, the net reproduction number under these circumstances would have been below one even though the disease incidence increased concurrently.

Some of the complexities which may affect reproduction numbers of infections with long incubation periods and/or serial intervals have been discussed in the literature, largely in relation to hepatitis [7]. The analyses presented here raise the broad issue of how to interpret the net reproduction number for any infectious disease at a given time. For diseases for which the serial interval (e.g. of a given duration s) varies little between individuals, and the epidemiological conditions do not change over this time interval, trends in morbidity at a given time reflect the magnitude of the net reproduction number at time s previously. *Thus the net reproduction number at a given time is an accurate predictor of future incidence only for diseases for which the serial interval varies little between individuals and for which epidemiological conditions do not change over this time interval.* Analogous complications (and the fact that reinfection can occur) also arise for diseases involving parasites with long life spans (for example, 8–10 years for *Onchocerca volvulus*) [7]. For diseases for which these conditions do not hold, variants of the reproduction number concept e.g. describing the number of secondary infectious cases arising from each infectious case *per unit time* (a 'net reproduction number *function*'), are perhaps required to predict trends in morbidity.

The basic reproduction numbers for tuberculosis

These analyses demonstrate that the basic reproduction number concept is particularly complicated for tuberculosis, and that the (simple) basic reproduction number must have declined over time, largely as a result of the decline in the effective contact number (the average number of infection transmissions per case). They also demonstrate that the conventional basic reproduction number measure at a particular time fails to reflect the actual transmission potential of *M. tuberculosis* infection, and suggest that this is best described by what we have called here the 'ultimate basic reproduction number'.

The fact that the (simple) basic reproduction number of an infection can change over time has been recognized in other contexts. For example, it is obvious that the basic reproduction number of an infection can change as a result of an intervention (e.g. one which alters the behaviour of individuals, or which reduces the duration of infectiousness). In such instances, the measure has sometimes been described as an 'effective reproduction number' [5] to reflect the fact that the intervention changed the 'intrinsic' properties of the infection. The basic reproduction number for several parasitic infections (e.g. schistosomiasis, hookworm) [7, 36] varies seasonally, through changes in social behaviour of human hosts or in environmental conditions which affect the infectious agent or its vector. There is also a literature on secular changes in the basic reproduction number as a result of genetic co-evolution of host-parasite systems [33, 34]. Studies of myxoma virus infection in rabbits provide an example of such changes [7]).

The analyses presented here provide another illustration of how the (simple) basic reproduction number of an infection may change over time. The decline in the basic reproduction number for tuberculosis is related to the decline in the effective contact number which, in turn, resulted from many changes in England and Wales, e.g. improvements in sanitation, water supply, housing and many aspects of health-related behaviour. These factors have undoubtedly influenced the (simple) basic reproduction numbers for many infections during this century in developed countries. Basic reproduction numbers may

also have changed through reductions in the risks of developing disease given infection and in the duration of infectiousness of diseased individuals, resulting from improvements in nutritional or general health status, as well as improved medical care. Actual trends in basic reproduction numbers will be disease-specific, and may even have increased in recent decades for some childhood infections as a result of increasing proportions of children attending crêches or pre-school day-care nurseries. This increased mixing among children is less important for the basic reproduction number for tuberculosis than for acute viral respiratory infections (e.g. measles or chickenpox), as the risk of infectious pulmonary tuberculosis subsequent to infection is low among infants (see Figs 2*c*, 5*a*).

The conventional interpretation of the (simple) basic reproduction number as a measure of the 'actual transmission potential' of an infection stems from analyses of acute childhood infections for which the time interval between successive cases in a chain of transmission is short, typically only a few days. For tuberculosis, at least in wealthy countries, we see that this interpretation of the (simple) basic reproduction number for a given year is inappropriate, since the effective contact number has declined over each serial interval, affecting both the number of individuals infected by each infectious case and the number of individuals who develop disease attributable to infection by a given case. Instead, the actual transmission potential is best reflected by what we have called here the 'ultimate basic reproduction number'. The distinction between the simple and ultimate basic reproduction numbers will be relevant for other infectious diseases for which epidemiological conditions affecting contact or the risk of developing disease change over intervals comparable to or shorter than the infectious or incubation period. Among the obvious examples is HIV/AIDS, for which the *infectious period* may be many years, during which transmission-related behaviours may change greatly Herpes varicella-zoster provides another example, as its *incubation period* may be many decades, and it is possible that the actual lifetime risks of developing disease may have changed over time, as a consequence of increased lifespan.

Our finding that the estimated ultimate basic reproduction number for tuberculosis exceeded the simple basic reproduction number during the pre-chemotherapy era may appear counter-intuitive, as a founder case should have infected fewer individuals following his/her introduction into an uninfected population with a declining effective contact number than if she/he was introduced into a population in which the effective contact number remained unchanged. The result follows from the fact that the annual risk of infection would have been *lower* following the introduction of an infectious case into a population in which the effective contact number declined over time (see Fig. 9), which means – if we accept that reinfection does occur at all – that fewer individuals would have been *re*infected (and thus a greater proportion would have developed disease *attributable to the initial infection*), than in the situation in which the effective contact number did not decline. An analogous explanation contributes to our finding that the lifetime risks of developing disease attributable to a given infection/reinfection at a given age *increased* slightly with calendar year of infection in the past (Fig. 5*b*).

It is interesting that the magnitudes of the simple and ultimate basic reproduction numbers for tuberculosis estimated here (i.e. below 3 throughout this century) are far lower than those for many acute infections. The basic reproduction numbers for measles and rubella, for example, are generally considered to be on the order of about 13 and 7 respectively [7]. This finding is consistent with the view that tuberculosis is less 'transmissible' than these infections [26, 32], despite its long infectious period. For example, of those exposed to an infectious tuberculosis case in a household, less than 65% appear to become 'infected' (on the basis of the tuberculin test) [37]; of these, only 5–15% develop disease during their lifetime (depending on age and on subsequent trends in the risk of (re)infection [13]), of whom only an age-dependent proportion are infectious (Fig. 2*c*). In contrast, for many acute 'childhood' infections, most uninfected individuals exposed to a case become infected, develop disease and become infectious.

The magnitude of the basic reproduction number is often assumed to reflect the likelihood of continued transmission of infection, and to have implications for eradication (as reflected in the herd immunity threshold) or control (see reviews in [5, 7, 38]) in a given population. This interpretation implies that, given the low values for the simple and ultimate basic reproduction numbers, control should be easier to attain for tuberculosis than for acute infections such as measles and rubella. As demonstrated here, such conclusions may be inappropriate, given the long

serial interval, which means that long term trends in morbidity are constrained by many factors such as demography, behaviour and contact between individuals which change over time. This raises the issue of whether the basic reproduction number concept can be extended to define a measure which describes the likelihood of continued transmission or eradication of an infection for which epidemiological conditions change over time. Such measures (sometimes called basic reproduction quotients [39, 40]) have been derived recently for infections for which transmission varies seasonally [39–41]. For tuberculosis, however, the derivation of any such measure is complicated by the problem of predicting changes before they occur, and hence before the measure itself can be derived. It is interesting that there has been so little discussion of this problem for other infectious diseases, though analogous issues have been raised by ecologists [42].

The basic and net reproduction numbers are two of the more widely applied concepts in infectious disease epidemiology, and have contributed greatly to our understanding of the impact of different control strategies on morbidity from acute infectious diseases. For tuberculosis we see that, as a result of the long and variable interval between infection and disease, and changes in epidemiological conditions which occurred during these intervals, and also the potential for reinfection to occur, the traditional interpretation of these concepts is inappropriate. The magnitude of the net reproduction number fails to reflect contemporary trends in disease incidence, and in order to describe the actual transmission potential of infection, we require a variant of the conventional (or 'simple') basic reproduction number measure, which we have defined here as the 'ultimate basic reproduction number'. Given the current interest and concern for new 'emerging' infections, let alone increased efforts to control 'old pathogens' in a rapidly-changing world, it is important that epidemiological theory incorporate the implications of changes in environment and society.

ACKNOWLEDGEMENTS

We thank the British Medical Research Council for financial support, Dr K. Styblo (Tuberculosis Surveillance Research Unit, The Hague) and Dr K. Bjartveit (Tuberculosis Screening Service, Norway) for supplying tuberculosis data from Norway, Dr J. Watson (PHLS Communicable Disease Surveillance Centre) for data from the national tuberculosis surveys in England and Wales, and the UK Government Actuary's Department for providing age and time-specific mortality rates for England and Wales. We are also grateful to Nigel Gay (CDSC) and Robert May (Dept of Zoology, Oxford University) for helpful comments.

APPENDIX

Expressions for the net reproduction number for tuberculosis

The model assumes that cases are infectious for 2 years, unless they die in the meantime (see Table 1). The net reproduction number for a given year was derived by considering the number of secondary cases which result from the first and second years of infectiousness of cases who have onset in that year, as follows.

Suppose that $D_n(t)$ infectious pulmonary cases have onset in year t. Given a total prevalence of infectious pulmonary cases in year t of $D_+(t)$, a proportion $D_n(t)/D_+(t)$ of the new infections and reinfections occurring in that year will be attributable to individuals who developed infectious pulmonary tuberculosis in year t. Similarly if a total of $D_u(t)$ individuals ultimately develop infectious pulmonary tuberculosis *attributable* to their infection/reinfection event in year t, then the total number of secondary cases which result from *the first year of infectiousness* of cases with onset of infectious pulmonary tuberculosis in year t is given by:

$$\frac{D_n(t)}{D_+(t)} D_u(t) \tag{1}$$

By an analogous argument, the total number of secondary cases resulting from the *second year of infectiousness* of cases who have onset of infectious pulmonary tuberculosis in year t is given by:

$$\frac{D_s(t+1)}{D_+(t+1)} D_u(t+1) \tag{2}$$

where $D_s(t+1)$ is the number of cases who had onset of infectious pulmonary disease in year t who *survive* until the subsequent year, $t+1$.

Summing equations 1 and 2, the total number of

cases resulting from individuals who developed infectious pulmonary tuberculosis in year t is given by:

$$\frac{D_n(t)}{D_+(t)}D_u(t)+\frac{D_s(t+1)}{D_+(t+1)}D_u(t+1) \quad (3)$$

Dividing through by $D_n(t)$, the average number of secondary cases attributable to individuals who developed infectious pulmonary tuberculosis in year $t(R_n(t))$ is given by

$$R_n(t)=\frac{1}{D_n(t)}\left\{\frac{D_n(t)}{D_+(t)}D_u(t)+\frac{D_s(t+1)}{D_+(t+1)}D_u(t+1)\right\} \quad (4)$$

This expression simplifies to:

$$R_n(t)=\frac{D_u(t)}{D_+(t)}+\frac{D_s(t+1)}{D_n(t)}\frac{D_u(t+1)}{D_+(t+1)} \quad (5)$$

REFERENCES

1. Raviglione MC, Sudre P, Rieder HL, Spinaci S, Kochi A. Secular trends of tuberculosis in Western Europe. *Bull WHO* 1993; **71**: 297–306.
2. Raviglione MC, Snider DE, Kochi A. Global epidemiology of tuberculosis. Morbidity and mortality of a worldwide epidemic. JAMA 1995; **273**: 220–6.
3. Macdonald G. The analysis of equilibrium in malaria. *Trop Dis Bull* 1952; **49**: 813–29.
4. Macdonald G. The epidemiology and control of malaria. London: Oxford University Press, 1956.
5. Dietz K. The estimation of the basic reproduction number for infectious diseases. Statist Meth Med Res 1993; 2: 23–41.
6. Diekmann O, Heesterbeek JAP, Metz JAJ. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. J Math Biol 1990; **28**: 365–82.
7. Anderson RM, May RM. Infectious diseases of humans – dynamics and control. Oxford: Oxford University Press, 1992.
8. Heesterbeek H. R_0 [thesis] Centrum voor Wiskunde en Informatica, Amsterdam 1992.
9. Canetti G. Endogenous reactivation and exogenous reinfection. Their relative importance with regard to the development of non-primary tuberculosis. Bull Int Union Tuberc 1972; **47**: 116–22.
10. Styblo K. Epidemiology of tuberculosis. Selected Papers. KNCV; The Hague, The Netherlands 1991.
11. Vynnycky E, Fine PEM. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. Epidemiol Infect 1997; **119**: 183–201.
12. Hope-Simpson RE. The period of transmission of certain epidemic diseases. Lancet 1948; ii: 755–60.
13. Vynnycky E, Fine PEM. The incubation period, serial interval and lifetime risks of developing tuberculosis. (Submitted).
14. Vynnycky E, Fine PEM. Interpreting the decline in tuberculosis during the last century – the roles of secular trends in effective contact. (Submitted).
15. Blower SM, Small PM, Hopewell PC. Control strategies for tuberculosis epidemics: new models for old problems. Science 1996; **273**: 497–500.
16. Blower SM, McLean AR, Porco TC, et al. The intrinsic transmission dynamics of tuberculosis epidemics. Nature Med 1995; 1: 815–21.
17. Sutherland I, Švandová E, Radhakrishna SE. The development of clinical tuberculosis following infection with tubercle bacilli. Tubercle 1982; **62**: 255–68.
18. Lowe CR. An association between smoking and respiratory tuberculosis. BMJ 1956; **2**: 1081–6.
19. Springett VH, Darbyshire JH, Nunn AJ, Sutherland I. Changes in tuberculosis notification rates in the white ethnic group in England and Wales between 1953 and 1983. J Epidemiol Community Health 1988; **42**: 370–6.
20. Medical Research Council Cardiothoracic Epidemiology Group. National survey of notifications of tuberculosis in England and Wales in 1988. Thorax 1992; **47**: 770–5.
21. Holm J. Development from tuberculous infection to tuberculous disease. TSRU Progress Report; KNCV; The Hague, The Netherlands, 1969.
22. National Tuberculosis Institute, Bangalore. Tuberculosis in a rural population of South India: a five-year epidemiological study. Bull WHO 1974; **51**, 473–88.
23 Meyer SN. Statistical investigations of the relationship of tuberculosis morbidity and mortality to infection. Copenhagen: Munksgaards Forlag, 1949.
24. Zeidberg LD, Dillon A, Gass RS. Risk of developing tuberculosis among children of tuberculous parents. Am J Hyg 1954; **70**: 1009–19.
25. Ferebee SH. Controlled chemoprophylaxis studies in tuberculosis: a general review. Vol 17. *Adv Tuberc Res.* Basel; Karger, 1970: 28–106.
26. Comstock GW, Livesay VT, Woolpert SF. The prognosis of a positive tuberculin reaction in childhood and adolescence. Am J Epidemiol 1974; **99**: 131–8.
27. Abbey H. An examination of the Reed-Frost theory of epidemics. Hum Biol 1952; **24**: 201–33.
28. Bates JH, Stead WW, Rado TA. Phage types of tubercle bacilli isolated from patients with two or more sites of organ involvement. Am Rev Respir Dis 1976; **114**: 353–8.
29. Mankiewicz E, Liivak M. Phage types of *M. tuberculosis* isolates in cultures isolated from Eskimo patients. Am Rev Respir Dis 1975; **111**: 307–12.
30. Challu V, Mahadev B, Rajalakshmi R, Chaudhuri K. Recovery of tubercle bacilli from urine of pulmonary tuberculosis patients and its comparison with the corresponding sputum isolates. Indian J Tuberc 1989; **36**: 107–11.
31. Press WH, Teukolsky SA, Vetterling WT, Flannery BP.

Numerical recipes in C. The art of scientific computing. 2nd ed. Cambridge: Cambridge University Press, 1992.
32. Frost WH. The age selection of mortality from tuberculosis in successive generations. Am J Publ Hlth 1939; **30**: 91–6.
33. Nowak MA, May RM. Superinfection and the evolution of parasite virulence. Proc Roy Soc Lond 1994; **255**: 81–9.
34. May RM, Nowak MA. Superinfection, metapopulation dynamics and the evolution of diversity. J Theor Biol 1994; **170**: 95–114.
35. May RM, Nowak MA. Coevolution and the evolution of parasite virulence. Proc Roy Soc Lond 1995; **261**: 209–15.
36. Woolhouse ME. On the application of mathematical models of schistosome transmission dynamics. I. Natural Transmission. *Acta Trop* 1991; **49**: 241–70.
37. van Geuns HA, Meijer J, Styblo K. Results of contact examination in Rotterdam, 1967–1969. Bull Int Union Tuberc 1975; **50**: 107–21.
38. Fine PEM. Herd immunity: history, theory, practice. Epidemiol Rev 1993; **15**: 265–302.
39. Roberts MG, Heesterbeek JAP. The dynamics of nematode infections of farmed ruminants. Parasitol 1995; **10**: 493–502.
40. Heesterbeek JAP, Roberts MG. Threshold quantities for helminth infections. J Math Biol 1995; **33**: 415–34.
41. Williams BG, Dye C. Infectious disease persistence when transmission varies seasonally. Math Biosci 1997; **145**: 77–88.
42. Metz JAJ, Nisbet RM, Geritz SAH. How should we define 'fitness' for general ecological scenarios. TREE 1992; **7**: 198–202.
43. Sutherland I. The ten-year incidence of clinical tuberculosis following 'conversion' in 2,550 individuals aged 14 to 19 years. TSRU Progress Report; KNCV; The Hague, The Netherlands 1968.
44. Vynnycky E, Fine PEM. The annual risk of infection with *Mycobacterium tuberculosis* in England and Wales since 1901. Int J Tuberc Lung Dis 1997; **1**: 389–96.