Lifetime Risks, Incubation Period, and Serial Interval of Tuberculosis

Emilia Vynnycky and Paul E. M. Fine

The lifetime risk of developing disease, the incubation period, and the time period between infection and transmission (the serial interval) are three important measures for interpreting trends in tuberculous infection and disease but are complicated by strong age dependencies regarding disease risk and by the potential for reinfection to occur. By using a model of the epidemiology of tuberculosis in England and Wales, the authors demonstrated that all three measures changed dramatically during the 20th century largely as a result of declines in the risk of infection. The estimated lifetime risk was highest following infection in early adulthood and declined with year of infection; the age-weighted average was approximately 12% during the last 50 years. Incubation period distributions depend on whether they are viewed prospectively (from infection to disease onset) or retrospectively (since infection for cases with disease onset at a particular time). As children rarely develop infectious forms of tuberculosis, infections acquired in childhood are associated with considerably longer serial intervals than those acquired in adulthood. These unusual properties are probably shared by other infections with long intervals between infection and disease. The results are important for interpreting data on transmission patterns, as are now being derived from molecular epidemiologic studies. Am J Epidemiol 2000;152:247–63.

It is often stated that the lifetime risk of developing clinical tuberculosis following infection with Mycobacterium tuberculosis is approximately 10 percent (1, 2) and that of those persons who develop disease, approximately half do so during the first few years after infection (3). However, to our knowledge, neither this lifetime risk nor the incubation period (the time interval between initial infection and disease onset) of tuberculosis have been measured directly. Little is known about the serial interval or generation time of tuberculosis (4), defined formally as the time interval between identical stages of disease in successive cases in a chain of transmission (5). These are three of the most important parameters defining the epidemiology of any infectious disease: the lifetime risks and the incubation period determine whether and how soon after initial infection someone is likely to develop disease, and the serial interval reflects how soon that person is likely to infect others.

For tuberculosis, the derivation of these measures is complicated by the fact that clinical disease may follow soon after initial infection (“primary” disease) or many years later, either through endogenous reactivation or after exogenous reinfection. As the risks of developing disease are age dependent (6–10) and because the risk of tuberculous infection has declined in many societies (11, 12), it is likely that the lifetime risks, incubation period, and serial interval measures are themselves age dependent and have changed over time. These possibilities have important implications for designing and interpreting studies to reveal patterns of transmission, for example, through molecular fingerprinting of M. tuberculosis strains collected from tuberculosis cases. For instance, infected children are more likely to develop noninfectious (nonpulmonary or smear-/culture-negative) than infectious (smear-/culture-positive) pulmonary forms of tuberculosis. Thus, many years may elapse before a newly infected child transmits infection to others, thereby potentially leading to secondary infectious cases.

To our knowledge, no studies to date have assessed in detail the implications of age and calendar year for the lifetime risks, incubation period, and serial interval of tuberculosis. The only discussion that we found in the literature of the effect of age on lifetime risk (8, 13) suggested that infected infants may have higher risks than adults, as they have many more years available in which to develop disease. The most detailed data on the incubation period of tuberculosis, those obtained during the United Kingdom Medical Research Council BCG trial in the 1950s (14) and those from the US Public Health Service chemoprophylaxis studies (9), are based on only 10 years of follow-up. Data on serial intervals for tuberculosis come from reports of outbreaks, which have limited utility because they involve relatively few and only short-time-linked cases. In addition,
many of the recent outbreaks in developed countries have involved persons positive for human immunodeficiency virus (HIV) (15–18) who face greater risks of developing disease than do those who are HIV negative (19). The most reliable insights into the serial interval for tuberculosis come from a recent DNA fingerprinting study (20), which was based on cases with disease onset in the Netherlands during a relatively short time period (4 years).

Direct estimation of the lifetime risks of developing tuberculosis and of the full incubation period and serial interval would require lifelong follow-up of those infected; thus, observational studies are impractical. However, these parameters can be estimated by using appropriate modeling techniques. This paper reports on our application of a general transmission model for tuberculosis (10) to derive age- and time-dependent estimates of the lifetime risks of developing disease and the incubation periods and serial intervals of tuberculosis, and it discusses general issues regarding how these measures are interpreted and applied to other infectious diseases.

MATERIALS AND METHODS

Basic assumptions of the model

Figure 1 shows the general structure of the model, which was designed to describe the dynamics of all forms of pulmonary tuberculosis in the population of England and Wales over the past century. The model’s input parameters and their values are summarized in table 1. Derivation and validation of the model have been described in detail elsewhere (10). To avoid the complications related to gender differences, immigration, and the HIV epidemic on the recent epidemiology of tuberculosis in England and Wales, our analyses related only to White males not infected with HIV (10).

The model assumes that persons are born uninfected and face an annual risk of infection \(i(t)\) that depends on calendar year (12) (declining from approximately 12 percent in 1900 by 4 percent per annum until 1950 and by 13 percent per annum thereafter (figure 2, top left)). Infection risk estimates for 1900–1949 were based on tuberculous meningitis mortality data (12); given the absence of reliable data, the decline in the annual risk of infection after 1950 was assumed to be the same as that found in the Netherlands (which has the most reliable estimates of secular trends in infection risks) (21).

Following the convention established by Holm (22), infected persons are categorized into those who have been infected for less than 5 years and who have not experienced (primary) disease \((I(a,t,s))\) and those in the “latent” class \((L(a,t))\) who are at risk of endogenous disease or of reinfection followed by exogenous disease. The risk of reinfection is assumed to be identical to that of first infection, although reinfection is less likely to lead to disease than is an initial infection due to some immunity induced by previous infection (10). These assumptions are based on results from our previous analyses (10), which found that the assumption that infection conferred no protection against reinfection (although it did reduce the probability that the reinfection event led to disease) led to the best fit to the data. For simplicity, the reinfection risk is assumed to be negligible for a period of 5 years after each (re)infection during the period of high risk for either the first primary episode or exogenous disease (10).

Risks of developing the first primary episode \((d_p(a,s))\) and of exogenous disease \((d_e(a,s))\) are assumed to depend on age and to decline with time since (re)infection (figure 2, bottom left and top right). The relation between age and the risks of developing disease is consistent with the available data. The decline in the risk of developing the first primary episode is based on data on the time interval between tuberculin “conversion” and disease onset for participants in the United Kingdom Medical Research Council BCG trial during the 1950s (14). There appear to be no data showing the relation between the risk of exogenous disease and time since reinfection, although the one assumed here appears the most reasonable. The risk of developing endogenous disease \((d_e(a))\) is age dependent. These risks were estimated by fitting disease incidence predictions from the model to age-specific notifications of pulmonary tuberculosis in England and
TABLE 1. Summary of parameter values used in the model designed to describe the dynamics of all forms of pulmonary tuberculosis in the population of England and Wales since 1900

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i(t))</td>
<td>Risk of infection and reinfection at time (t)</td>
<td>Twenty percent until 1880, declining by 2% per annum until 1901, by 4% per annum until 1950, and by 13% per annum thereafter (12) (figure 2, top left)</td>
</tr>
<tr>
<td>(v(a,t))</td>
<td>Proportion of uninfected persons aged (a) immunized at time (t)</td>
<td>Vaccination introduced in 1954 and restricted to 13-year-olds; vaccine efficacy assumed to be 77%; vaccine coverage increasing to approximately 80% since 1960 (10)</td>
</tr>
<tr>
<td>(d_s(a,s))</td>
<td>Risk of developing the first primary episode at time (s) after infection at age (a)</td>
<td>Depends on age at and time since first infection (figure 2, bottom left and top right); cumulative risks within 5 years of initial infection: 4.06%, 8.98%, and 13.8% for 0–10-year-olds, 15-year-olds, and those aged over 20 years, respectively (10)</td>
</tr>
<tr>
<td>(d_e(a,s))</td>
<td>Risk of developing exogenous disease at time (s) after reinfection at age (a)</td>
<td>Relation between age at and time since reinfection identical to that between (d_s(a,s)) and age at and time since first infection (figure 2, bottom left and top right); cumulative risks within 5 years of reinfection: 6.89%, 7.57%, and 8.25% for 0–10-year-olds, 15-year-olds, and those aged over 20 years, respectively (10)</td>
</tr>
<tr>
<td>(d_a(a))</td>
<td>Annual risk of developing endogenous disease at age (a)</td>
<td>Refer to figure 2, bottom left; assumed to be 9.82 × 10⁻⁶%, 0.0150%, and 0.0299% for 0–10-year-olds, 15-year-olds, and those aged over 20 years, respectively (10)</td>
</tr>
<tr>
<td>(d_i(a))</td>
<td>Proportion of total disease incidence among cases aged (a) assumed to be infectious</td>
<td>Ten percent for 0–10-year-olds, increasing linearly to 65% for 20-year-olds and increasing linearly to 85% for 90-year-olds (figure 2, bottom right)</td>
</tr>
<tr>
<td>(k_s(s))</td>
<td>Rate at which persons who have been infected or reinfected for time (s) without developing disease move into the “latent” class</td>
<td>Transition occurs exactly 5 years after infection/reinfection, i.e., (k_s(s) = 0) if (0 &lt; s &lt; 5) and (\infty) for (s = 5) years</td>
</tr>
<tr>
<td>(r(a,t,\hat{s}))</td>
<td>Recovery rate for cases aged (a) at time (t) at time (\hat{s}) after disease onset</td>
<td>Cases are diseased for 2 years unless they die in the meantime (see below)</td>
</tr>
<tr>
<td>(m_s(t,\hat{s}))</td>
<td>Case fatality of infectious pulmonary cases at time (t) and time (\hat{s}) since disease onset.</td>
<td>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 (10)</td>
</tr>
<tr>
<td>(m_{n}(a,t))</td>
<td>Mortality rate of noninfectious and non-diseased persons in the general population aged (a) at time (t).</td>
<td>Identical to all-cause mortality (after subtracting deaths of infectious cases, estimated in the model); annual age-specific all-cause mortality rates obtained from the Government’s Actuary Department (United Kingdom) since 1841; data until 1841 obtained by back extrapolation</td>
</tr>
</tbody>
</table>

Wales during 1953–1988 (10); model predictions made by using these disease risks also compared well with observed age-specific mortality rates during the prechemotherapy era (10, 23). The sensitivity of the model to the various assumptions has been explored elsewhere (10).

An age-specific proportion of disease is assumed to be “sputum positive,” that is, infectious \((d_s(a))\), as was observed in a classic data set from Norway (figure 2, bottom right; also refer to Murray et al. (24)). This age-specific pattern is assumed to be independent of the mechanism of disease onset and of whether it is the first or a subsequent disease episode. Cases are assumed to be infectious for 2 years unless they die in the meantime. The case-fatality rate for infectious cases is assumed to be 50 percent during the prechemotherapy era, as was found in several observational studies (refer to the reviews by Styblo (11) and Murray et al.) and in a major longitudinal study of the natural history of tuberculosis in the absence of treatment (25). BCG vaccination has also been incorporated, although those effectively protected by vaccination do not develop disease and hence do not contribute any information to the incubation period, serial interval, and lifetime risks of developing disease. Age- and timespecific mortality rates were obtained from the United Kingdom Government’s Actuary Department and were used in the model to simulate the historical demographic structure. The only simplification was to assume that no one survived after age 100 years.

Appendix table 1 provides detailed definitions of the several disease categories used in the model. Figure 3 illustrates how well the model’s predictions compare with observed mortality and notifications in England and Wales in the past.
Complications regarding the incubation period, lifetime risks, and serial interval of tuberculosis

Incubation period and lifetime risks. The incubation period of an infection is defined as the time interval between initial infection and disease onset. There are several complications regarding the incubation period of tuberculosis.

For many infections, incubation period distributions are derived directly by considering cases whose onset of both infection and disease occurred during a given time period. For tuberculosis, however, any distribution derived by considering such cases (e.g., over a 10- or 20-year period) is likely to be “censored,” as some persons will have incubation periods of several decades and will thus be excluded.

Because the risks of developing disease are lifelong, the incubation period distribution depends on whether it is derived prospectively (i.e., as the time interval from infection until disease onset for those persons newly infected) or retrospectively (i.e., as the time interval since infection for new cases). Persons who acquire infection late in life, for example, have few years in which to develop disease and hence have short incubation periods derived prospectively in comparison to those who acquire infection when young: cases who are old have had many years in which to have developed disease and can have long incubation periods derived retrospectively. The fact that the risks of developing disease are age dependent (6–10) further complicates these distributions. In contrast, for acute viral and bacterial infections, the incubation period can be considered a (relatively) fixed characteristic of the infection.

The incubation period typically is interpreted as the time required for the infecting pathogen to multiply in the host until it leads to disease. For tuberculosis, the fact that reinfection can occur leads to further complications, since the bacilli that cause disease in reinfected cases may not be from the initial infection event but from a subsequent reinfection. For reinfectected cases, the incubation period is perhaps best defined as the time between the “causal” reinfection event and disease onset. This distribution, derived either prospectively or retrospectively, depends on how frequently reinfection occurs and thus on the annual risk of infection. Thus, in developed countries, where the annual risk of infection declined dramatically during the 20th century (12, 21), the incubation period also must depend on the calendar year of disease onset (from the retrospective perspective) or of infection (from the prospective perspective).
Several of these complications also affect the lifetime risk of developing disease. For example, given age-dependent risks of disease, it is obvious that lifetime risks should depend on age at infection; given the decline in the annual risk of infection during the 20th century in developed countries, this lifetime risk also may have declined over time.

**Serial interval.** The distinction between the incubation period and the serial interval of an infection is rarely appreciated. The serial interval is defined as the time between homologous stages of disease in successive cases in a chain of transmission (5) and depends on two factors: 1) the time interval between infection and onset of infectiousness and 2) the duration of infectiousness. For tuberculosis, the serial interval is complicated by the fact that only an age-dependent proportion of those infected develop infectious pulmonary disease and thus transmit bacilli to others. As for the incubation period, the serial interval can be examined either retrospectively as the time interval since cases infecting others were themselves infected or prospectively as the time interval until those infected at a given time infect others.

**Analyses of the incubation period, serial interval, and lifetime risks**

In our analyses, we illustrate the differences between the incubation period derived prospectively and retrospectively and explore the effect of age and calendar year on these distributions. We also examine how the incubation period differs from the serial interval and derive estimates of the lifetime risks of developing disease.

**Retrospective analyses of the incubation period.** The incubation period was first derived retrospectively as the time interval since initial infection for cases with onset of any form of pulmonary tuberculosis at a given age at a given time, considering cases with onset at different ages in the years 1900, 1930, 1960, and 1990 in England and Wales. To illustrate how this distribution may differ from that of the time interval since the (re)infection event “causing” the disease in these cases, we also derived the distribution of the time interval since the most recent (re)infection event for cases with onset of any form of pulmonary tuberculosis in these years on the assumption that bacilli causing disease in reinfected cases were from the most recent reinfection event.
To derive these distributions for cases whose onset of disease occurred at age \( a \) and at time \( t \), we first used the model to obtain the number of persons in this birth cohort (born in year \( t-a \)) infected during each year of life until age \( a \). The corresponding equations in the Appendix were then applied to these numbers to obtain 1) \( D(a,t,s) \), the number of these persons who developed disease at age \( a \) after having been infected for duration \( s \) years and 2) \( D^0(a,t,s) \) and \( D^0_1(a,t,s) \), the number of these persons who developed disease at age \( a \) after having been infected or reinfected, respectively, for duration \( s \) years, without having being reinfected in the meantime. We then obtained the distribution of the incubation period, derived retrospectively as the time interval since initial infection for cases whose disease onset occurred at age \( a \) at time \( t \) (\( F_r(a,t) \)) by dividing \( D(a,t,s) \) by the total number of cases whose onset began at age \( a \) at time \( t \), as follows:

\[
F_r(a,t) = \frac{D(a,t,s)}{\sum_{s=0}^{100-a} D(a,t,s)} \quad (0 \leq s \leq a)
\]

The equation used to determine the distributions of the time interval since the most recent (re)infection for cases whose disease onset occurred at age \( a \) at time \( t \) (\( F_r(a,t) \)) is analogous:

\[
F_s(a,t) = \frac{D^0(a,t,s) + D^0_1(a,t,s)}{\sum_{s=0}^{100-a} D^0(a,t,s) + D^0_1(a,t,s)} \quad (0 \leq s \leq a)
\]

Prospective analyses of the incubation period and lifetime risks. In these analyses, the incubation period, as derived prospectively, was defined as the time interval until the first episode of any form of pulmonary tuberculosis for persons initially infected at a given age at a given time. These distributions were derived for persons infected at different ages in 1900 and 1940 who developed disease some time during their lives by using methods similar to those described for the retrospective analyses, as follows. The equations in the Appendix were applied to the number of persons newly infected at the given age \( a \) in the given year \( t \) (\( I_a,t \)) (as estimated by using the model) to derive the number of persons who developed disease for the first time during each subsequent year of life after having been infected for duration \( s \) (\( D_1(a,t,s) \)). The overall distribution of the incubation period (\( F_1(a,t) \)) was then derived by dividing \( D_1(a,t,s) \) by the total number of persons infected at age \( a \) at time \( t \) who developed disease some time during their lives, as follows:

\[
F_1(a,t) = \frac{D_1(a,t,s)}{\sum_{s=0}^{100-a} D_1(a,t,s)} \quad (0 \leq s \leq 100 - a)
\]

Lifetime risks of developing infectious pulmonary tuberculosis were derived analogously.

Analysis of the serial interval. In these analyses, we defined the serial interval for tuberculosis as the time interval between receipt (onset) of infection in one (primary) case and receipt of infection or reinfection in a secondary “case” in a chain of transmission. The distributions were derived prospectively for persons infected at different ages in different years (1900 and 1940), as follows. Model predictions of the number of persons infected at a given age \( a \) at time \( t \) and the corresponding equations (Appendix), were used to derive \( S_s(a,t,s) \), the number of these persons who infect others after time \( s \). The distribution of the serial interval (\( F_s(a,t) \)) was then obtained by dividing \( S_s(a,t,s) \) by the total number of persons infected at age \( a \) and time \( t \) who infect others sometime thereafter, as follows:

\[
F_s(a,t) = \frac{S_s(a,t,s)}{\sum_{s=0}^{100-a} S_s(a,t,s)} \quad (0 \leq s \leq 100 - a)
\]

In these analyses, we assumed that cases are infectious for 2 years unless they die in the meantime. It is recognized that, in reality, the duration of infectiousness also depends on treatment and on factors such as the time course of the degree of infectiousness in relation to disease onset (which is still inadequately known (26)) and that the serial interval depends further on mixing patterns in the population.

RESULTS

Retrospective analyses of the incubation period

Figure 4 (top) shows the results of the retrospective analysis of incubation periods, defined as the time interval since initial infection for cases whose disease onset occurred at age 20 years in 1900, 1930, 1960, and 1990. The majority of cases with disease onset early in the century are estimated to have been initially infected during early childhood and relatively few during the year preceding clinical onset (e.g., only 4 percent of those whose disease onset occurred in 1900). The pattern changed during the century as smaller and smaller proportions of persons were infected in childhood and an increasing proportion of cases followed recent initial infection.

Figure 4 (bottom) shows the analysis of the time intervals since the most recent (re)infection event for cases whose disease onset occurred at age 20 years. About 60 percent of those whose disease onset occurred in 1900 are estimated to have received their most recent (re)infection during the previous year, despite having first been infected in childhood (figure 4, top). On the other hand, for those diseased in 1990, there was little difference between the distributions since initial (figure 4, top) and most recent (re)infection (figure 4, bottom). Similar patterns were seen in the corresponding distributions for cases whose disease onset occurred at age 50 years (figure 5).

Prospective analyses of the incubation period and serial interval

Figures 6 and 7 present results from our prospective analysis of the incubation period and serial interval of
tuberculosis. Infection acquired at a young age was associated with appreciably longer incubation periods and serial intervals—time to transmission to a subsequent "case"—than was infection acquired in adulthood. For example, of those who developed respiratory tuberculosis following initial infection in 1900, the proportion who did

---

**FIGURE 4.** Predicted distributions of the time intervals since initial tuberculous infection (top panel) and since the most recent (re)infection (bottom panel) for all cases whose disease onset occurred at age 20 years in 1900, 1930, 1960, and 1990 in England and Wales. Shaded areas reflect the proportion of disease attributed to the first primary episode, exogenous disease, and reactivation of the primary infection or a reinfection event.
FIGURE 5. Predicted distributions of the time intervals since initial tuberculous infection (top panel) and since the most recent (re)infection (bottom panel) for all cases whose disease onset occurred at age 50 years in 1900, 1930, 1960, and 1990 in England and Wales. Shaded areas reflect the proportion of disease attributed to the first primary episode, exogenous disease, and reactivation of the primary infection or a reinfection event.

so within 5 years of infection was far greater if infection occurred at age 40 years than at age 10 years (i.e., 60 vs. 20 percent, respectively (figure 6, top)). These figures also show the extent to which the incubation periods were
shorter than the serial intervals for all ages at infection. This difference was greatest when measured for infections acquired in childhood; for example, 25 percent of the incubation periods and 5 percent of the serial intervals associated with infections in early life were less than 5 years in duration.
FIGURE 7. Top panel: predicted distributions of the time interval between initial infection and onset of pulmonary tuberculosis (all forms) (the incubation period) for persons who developed disease sometime during their lives following infection at different ages in 1940 in England and Wales. Bottom panel: predicted distributions of the serial interval for persons infected in 1940 in England and Wales, by age at infection. Shaded areas reflect the proportion of disease attributed to the first primary episode and to endogenous and exogenous disease.

The implications of calendar year of infection on these distributions are also illustrated in figures 6 and 7, which show that infection in 1940 was associated with appreciably shorter incubation periods and serial intervals than was infection in 1900. Although a smaller proportion of persons infected in infancy in 1940 developed disease than did those
who were infected in infancy in 1900 (figure 8), 55 percent of the disease episodes are estimated to have occurred within 5 years of infection in 1940 (figure 7, top) compared with only 23 percent of episodes after infection in infancy in 1900 (figure 6, top).

**Lifetime risks of developing disease**

Figure 8 summarizes the estimated lifetime risks of developing any form of pulmonary tuberculosis (top) and infectious (smear-/culture-positive) pulmonary tuberculosis (bottom) as a function of age and time of initial infection. For all ages at infection, lifetime risks of developing pulmonary tuberculosis (all forms) are estimated to have declined over time, for example, from about 27 percent in 1900 to 16 percent in 1950 for persons infected at age 20 years (figure 8, top). Estimated lifetime risks for infected 20-year-olds are higher than those for any other age group. Similar secular trends are seen in lifetime risks of developing infectious pulmonary disease (figure 8, top) except that, from 1940, the highest risks were for those infected when aged more than 70 years. The average lifetime risk of developing pulmonary tuberculosis (all forms), weighted according to the predicted number of persons initially infected at each age, is estimated to have declined during the early years of the 20th century, plateauing at about 12 percent by 1930 for all forms of pulmonary disease (figure 8, top) and at about 6 percent for infectious pulmonary forms (figure 8, bottom).

**DISCUSSION**

The work described in this paper probably constitutes the first attempt to estimate the full lifetime risks of developing disease, the incubation period, and the serial interval for pulmonary tuberculosis and to relate them to age and calendar year of tuberculous infection. Such estimates are important for forming the basis of detailed predictions about trends in tuberculosis. Our estimates were based on a model that extends the classic work of Sutherland et al. (10, 27). Although the model’s basic assumptions are realistic (e.g., the risks of developing disease are age dependent, and reinfection can occur), it still oversimplifies the natural history of tuberculosis. For example, we assumed that reinfection cannot occur if persons are either already diseased or at high risk of developing their first primary episode or exogenous disease. When analyzing the serial interval, we did not account for any variability in infectiousness between cases or in nonrandom mixing patterns in the population (which may encourage selective transmission within certain age groups). Our analyses also did not include the effect of chemoprophylaxis (which has not been used extensively in England and Wales), which would reduce the lifetime risks of developing disease and perhaps lengthen incubation periods and serial intervals. In spite of these simplifications, these analyses provide unique insights into the changes that occurred in these parameters during the decline in tuberculosis that occurred in developed countries during the 20th century.

**Incubation periods and serial intervals**

Most of the literature on the incubation periods of infectious diseases is based on acute infections, for which the incubation period is short (i.e., measured in hours or days) and the infection induces a permanent and solid immunity (so that a single disease episode is attributable to a single infection event). Tuberculosis presents a much more complicated picture.

The incubation period distribution depends on whether it is derived prospectively (as the time until disease onset for those persons newly infected at some defined time) or retrospectively (as the time since infection for cases whose disease onset occurred at some defined time). This will be true of any infectious disease for which the risk of infection or of disease changes over a time scale similar to that of the time interval between infection and disease onset. In addition, for tuberculosis (and for diseases such as acquired immunodeficiency syndrome (AIDS), leprosy, and perhaps the transmissible spongiform encephalopathies such as kuru and Creutzfeldt-Jakob disease), many years—even decades—can elapse between infection and disease onset; therefore, the background life expectancy of persons infected at different ages will also influence incubation period distributions obtained from the two perspectives. Thus, for old cases today, most tuberculous disease is a consequence of epi-
demiologic events that took place long ago, when the risk of infection was far higher and when both the lifetime risks of developing disease and the life expectancies were different from what they are today.

The incubation period of tuberculosis is complicated further by the fact that reinfection can occur. The implications of reinfection for the definition of incubation periods do not appear to have been considered at all in the infectious disease literature. With tuberculosis, the situation is particularly complicated because the extent of reinfection disease in different populations is itself contentious, and we are rarely able to identify a particular disease episode as being “attributable” to an initial or subsequent (re)infection. Recent applications of molecular epidemiology have provided examples of reinfection, or “superinfection,” disease (28), but too few data have accumulated to date to provide frequency distributions of the incubation periods of such episodes. The distributions provided here of time intervals to disease from either initial infection or most recent (re)infection probably represent the first efforts to distinguish these processes, and they provide theoretical estimates against which to infer the historical background of contemporary disease.

The serial interval has received little attention in the infectious disease literature and is often confused with the incubation period, largely because most studies have focused on acute infections for which the two measures are very similar. Our analyses demonstrate that the two measures are very different for tuberculosis, in large part because the relative infectiousness of cases (proportion culture-/smear-positive) is a function of age. As the serial interval defines the time scale at which infection spreads through a population (5, 29), a basic understanding of its properties is helpful for interpreting transmission patterns observed during a given time interval (e.g., as are now being derived from molecular epidemiology studies (30)). For example, most investigations of outbreaks look through a relatively narrow time window of a few months and thus will miss chains of transmission involving links of many years.

**Lifetime risks of developing tuberculosis**

Our analyses indicate that the lifetime risks of developing tuberculosis have declined over time and were consistently higher for those infected as adolescents and adults rather than as infants. The effects mirror the age patterns of the risks of developing the three types of disease (“primary,” endogenous, and exogenous disease) as derived when we fitted the model to observed notification data (10). It is interesting that the age-weighted average of the lifetime risk of developing tuberculosis derived in these analyses also declined over time and that, since 1930, it has been close to the 10 percent figure often quoted in the tuberculosis literature.

The secular decline in the lifetime risks predicted in our study is consistent with estimates of a declining incidence of exogenous disease in England and Wales (10), which in turn was a consequence of the reduction in the risk of infection (i.e., from about 14 percent per annum in 1900 to about 2 percent in 1950 and to less than 0.01 percent by 1990 (12)). Thus, for example, a person first infected in 1950 who did not develop disease shortly after infection was much less likely to experience reinfection and (exogenous) disease sometime thereafter than was someone first infected in 1900. The decline in risk associated with reinfection may have been confounded by improvements in living standards and the nutritional and health status of the population, which could have reduced the risk of developing disease following infection, and by the increased prevalence of smoking, which could have increased the risk of developing disease (31).

It is often stated that the lifetime risk of developing tuberculosis is 10 percent and that of those who do develop disease, approximately half do so during the first few years after infection and the remainder do so sometime thereafter (3). Our analyses suggest that both the lifetime risks of developing disease and the distribution of the morbidity after infection are strongly age (and time) dependent. For example, the estimated lifetime risks of developing disease following infection in most years were higher for the elderly and adults than for infants and children. Of those infected in their first year of life early in the 20th century who ultimately developed disease, only about 20 percent are estimated to have done so within 5 years of infection as compared with about 60 percent of those infected at age 40 years.

Although these age patterns (especially regarding lifetime risks) may appear counterintuitive (also refer to Comstock et al. (8) and Comstock (13)), they reflect the fact that the high risk of developing (primary) disease in adulthood (e.g., a 14 percent risk following infection at age 20 years (10)) greatly outweighs the cumulation of lower risks of developing disease experienced over many years (a 5 percent risk of developing primary disease following infection in childhood and a 0.03 percent annual risk of developing endogenous disease in adulthood (table 1)). Analogous processes (e.g., increasing risks of developing infectious pulmonary tuberculosis with age (figure 2, bottom right)) account for the higher lifetime risks of developing infectious pulmonary tuberculosis estimated for newly infected 70-year-olds (figure 8) as compared with the risks for younger people.

Although the lifetime risk of disease following infection estimated for the early 20th century appears high (e.g., weighted average of 18.5 percent for all age groups (figure 8)), it is not unrealistic. Tuberculosis was responsible for about 10 percent of all deaths during the early years of this century in England and Wales (32), and the mortality rates of respiratory tuberculosis exceeded 200 per 100,000 per year for young adults (figure 3). One of the first longitudinal morbidity studies ever carried out, in Williamson County (Tennessee) (7), found that 19.7 percent of household contacts of sputum-positive cases developed clinical tuberculosis during 10 years of follow-up. Although that result was based on relatively few cases, the magnitude of the risk is remarkable given that not all of those persons exposed had necessarily been infected. Other studies from the prechemotherapy era involving shorter follow-ups (refer to reviews in Styblo (11)) also found very high risks of disease in adulthood.

These analyses provide insights into the properties of the incubation period, serial interval, and lifetime risks of tuberculosis. Some of these properties also hold for other infec-
tious diseases with long intervals between infection and disease, such as AIDS, herpes varicella-zoster, leprosy, and perhaps the spongiform encephalopathies. AIDS is also similar to tuberculosis in that its incubation period appears to be strongly age dependent (33). Zoster is a particularly interesting analogy, in that it is thought often to represent recrudescence of infection acquired in early childhood and may possibly also involve reinfec tion (34). Leprosy is a problematic analogy, as the absence of any test for infection has meant that most estimates of the incubation period are based on disease in people who visit endemic areas for short periods of time (35). The epidemiology of tuberculosis has changed recently, largely as a result of HIV. The implications of these changes depend greatly on the extent to which the HIV epidemic has increased transmission of \textit{M. tuberculosis} to those who are HIV negative; thus, they are difficult to assess without having estimates of the lifetime risks, incubation periods, and serial intervals of tuberculosis. Given that much of the increased morbidity observed so far has occurred among HIV-positive persons who experience considerably shorter incubation periods (and higher lifetime risks) of tuberculosis than do those who are HIV negative (19), it is possible that so far we have witnessed only a fraction of the total tuberculosis morbidity that will ultimately occur as a result of the HIV epidemic. A detailed understanding of the properties and behavior of the lifetime risks, incubation period, and serial interval of tuberculosis in the past in developed countries provides a foundation for interpreting patterns and trends in tuberculous infection and disease over the years to come.

**ACKNOWLEDGMENTS**

The authors thank the British Medical Research Council for financial support, the late Dr. K. Styblo (Tuberculosis Surveillance Research Unit, The Hague) and Dr. K. Bjartveit (Norwegian Health Screening Service) for supplying tuberculosis data from Norway, Dr. J. Watson (Public Health Laboratory Service, Communicable Disease Surveillance Centre) for data from the national tuberculosis surveys in England and Wales, and the United Kingdom Government Actuary's Department for providing age- and time-specific mortality rates for England and Wales.

**REFERENCES**


APPENDIX

Definitions of disease categories used in the model

APPENDIX TABLE 1. Definitions of state variables used in the model designed to describe the dynamics of all forms of pulmonary tuberculosis in the populations of England and Wales since 1900 (refer to figure 1)

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>B(t)</td>
<td>Number of livebirths at time t. Obtained from the Office for Population and Census Surveys (OPCS) (36, 37) since 1841; estimates for the years prior to 1841 derived by back extrapolation</td>
</tr>
<tr>
<td>U(a,t)</td>
<td>Number of uninfected persons aged a at time t</td>
</tr>
<tr>
<td>V(a,t)</td>
<td>Number of persons aged a at time t protected by BCG vaccination</td>
</tr>
<tr>
<td>I(a,t,s)</td>
<td>Number of persons aged a at time t who have been infected for time s (≤5 years) without having yet developed disease</td>
</tr>
<tr>
<td>P(a,t,s)</td>
<td>Number of persons aged a experiencing their first primary episode at time t who have been diseased for time s</td>
</tr>
<tr>
<td>L(a,t)</td>
<td>Number of persons aged a at time t in the “latent” class, i.e., those who have either just recovered from their first primary episode or who have been infected for more than 5 years</td>
</tr>
<tr>
<td>I1(a,t,s)</td>
<td>Number of persons aged a at time t who have been reinfected for time s (≤5 years) and who have not yet developed exogenous disease</td>
</tr>
<tr>
<td>E1(a,t,s)</td>
<td>Number of persons aged a with exogenous disease at time t who have been diseased for time s</td>
</tr>
<tr>
<td>E2(a,t,s)</td>
<td>Number of persons aged a with endogenous disease at time t who have been diseased for time s</td>
</tr>
</tbody>
</table>
Estimation of the incubation period and serial interval

We used subscript $A,T$ to denote persons who were first infected at age $A$ at time $T$. We used the following disease classes to describe the disease dynamics in these persons:

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{s,i}(s)$</td>
<td>Number of persons infected for duration $s$</td>
</tr>
<tr>
<td>$P_{A,i}(s,\hat{s})$</td>
<td>Number of persons infected for duration $s$ who have been experiencing their first primary episode for $\hat{s}$ years</td>
</tr>
<tr>
<td>$L_{s,i}(s)$</td>
<td>Number of persons in the “latent” class infected for duration $s$</td>
</tr>
<tr>
<td>$E_{m,i}(s,\hat{s})$</td>
<td>Number of persons infected for duration $s$ and who have been experiencing endogenous disease for $\hat{s}$ years</td>
</tr>
<tr>
<td>$l_{s,i}(s,s)$</td>
<td>Number of persons infected for duration $s$ and for whom $s_i(\leq 5)$ years have elapsed since the most recent reinfection</td>
</tr>
<tr>
<td>$I_{s,i}(s)$</td>
<td>Number of persons reinfected for duration $s$</td>
</tr>
<tr>
<td>$E_{m,i}(s,\hat{s})$</td>
<td>Number of persons infected for duration $s$ and who have been experiencing exogenous disease for $\hat{s}$ years</td>
</tr>
</tbody>
</table>

We used the following shorthand notation for several of the parameters (also refer to text table 1):

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_i(A + S, T + s)$</td>
<td>$m_i(s)$</td>
</tr>
<tr>
<td>$m_i(T + s, \hat{s})$</td>
<td>$m_i(s)$</td>
</tr>
<tr>
<td>$d_i(A + s)$</td>
<td>$d_i(s)$</td>
</tr>
<tr>
<td>$d_i(A + s)$</td>
<td>$d_i(s)$</td>
</tr>
<tr>
<td>$i(T + s)$</td>
<td>$i(s)$</td>
</tr>
<tr>
<td>$r(A + S, t + s, \hat{s})$</td>
<td>$r_i(s)$</td>
</tr>
</tbody>
</table>

Retrospective analyses of the incubation period

Distributions of the time interval since initial infection. The system of equations describing the disease dynamics, according to time since infection $s$, for persons infected at age $A$ at time $T$ is as follows:

\[
\frac{dI_{A,T}}{ds} = -(d_i(A,s) + m_i(s) + k_i(s))I_{A,T}(s) \quad (0 < s \leq 5) \quad (1)
\]

\[
\frac{\partial P_{A,T}(s,\hat{s})}{\partial s} + \frac{\partial P_{A,T}(s,\hat{s})}{\partial \hat{s}} = -(m_{s}(s)d_{s}(s) + m_{s}(s)d_{-}(s) + r_{s}(s))P_{A,T}(s,\hat{s}) + d_i(A,s)I_{A,T}(s) \quad (0 < s \leq 5) \quad (2)
\]

\[
\frac{dL_{A,T}}{ds} = (I_{A,T}(s) + I_{m,i}(s,5))k_i(5) + r_i(2)(P_{A,T}(s,2) + E_{m,i}(s,2) + E_{e,i}(s,2))
- (i(s) + d_{n}(A + s) + m_{s}(s))L_{A,T}(s) \quad (3)
\]
The equations were implemented by using the C computer programming language using Forward Euler differencing (38), with time-steps of 1 year for both age and time. The total number of cases infected for duration \( s \), who were first infected at age \( A \) at time \( T \), \( D_A(A,T,s) \), is given by:

\[
D_A(A,T,s) = P_{A,T}(s,0) + E_{n_A}(s,0) + E_{x_A}(s,0)
\]

**Distributions of the time since most recent (re)infection.** The system of equations describing the disease dynamics for persons infected at time \( T \) at age \( A \) and who subsequently escape reinfection is given by equations 1–4, with equation 3 amended as follows to exclude persons reinfected since their initial infection:

\[
\frac{dL_{A,T}}{ds} = I_{A,T}(5)k_L(5) + r_s(2)(E_{n_A}(s,2) + E_{x_A}(s,2)) - (i(s) + d_A(A + s) + m_g(s))L_{A,T}(s)
\]

Based on these equations, the number of persons developing disease at time \( s \) following infection at age \( A \) at time \( T \), without having been reinfected in the meantime, is given by: \( D'_R(A,T,s) = P_{A,T}(s,0) + E_{n_A}(s,0) \).

The equations that describe the disease dynamics among persons reinfected at age \( A \) and time \( T \) and who subsequently escape reinfection are given below. In each equation, \( s \) denotes time since reinfection.

\[
\frac{dI_{x_A}}{ds} = - (d_A(A,s) + m_g(s) + k_L(s))I_{x_A}(s)
\]

\[
\frac{dE_{n_A}(s,\hat{s})}{ds} + \frac{dE_{x_A}(s,\hat{s})}{ds} = - (m_A(s)d_v(s) + m_g(s)d_v(s) + r_s(\hat{s}))E_{n_A}(s,\hat{s}) + d_A(A,s)I_{x_A}(s)
\]

\[
\frac{dL_{A,T}}{ds} = I_{x_A}(5)k_L(5) + r_s(2)(E_{n_A}(s,2) + E_{x_A}(s,2)) - (i(s) + d_A(A + s) + m_g(s))L_{A,T}(s)
\]

\[
\frac{dE_{n_A}(s,\hat{s})}{ds} + \frac{dE_{x_A}(s,\hat{s})}{ds} = - (m_A(s)d_v(s) + m_g(s)d_v(s) + r_s(\hat{s}))E_{n_A}(s,\hat{s}) + d_A(A + s)I_{x_A}(s)
\]

Based on these equations, the number of persons reinfected at age \( A \) and at time \( T \) who develop disease after time \( s \) without having been reinfected in the meantime is given by: \( D'_R(A,T,s) = E_{n_A}(s,0) + E_{x_A}(s,0) \).

**Prospective analyses**

**Distributions of the incubation period.** The system of equations describing the disease dynamics among persons infected at age \( A \) and at time \( T \), who have not yet experienced any form of pulmonary tuberculosis, is given by equations 1–6, with equation 3 amended as follows to exclude persons who experienced any form of pulmonary disease in the meantime:

\[
\frac{dL_{A,T}}{ds} = (I_{A,T}(5) + I_{n_A}(s,5))k_L(5) - (i(s) + d_A(A + s) + m_g(s))L_{A,T}(s)
\]
The number of persons experiencing any form of pulmonary tuberculosis at time $s$ after being infected at age $A$ and at time $T$ is given by $D_i(A,T,s) = P_{A,T}(s,0) + E_{s_i}(s,0) + E_{s_{i+1}}(s,0)$.

**Distributions of the serial interval.** The system of equations describing the disease dynamics among persons infected at age $A$ and at time $T$, who have not yet experienced infectious pulmonary tuberculosis, is given by equations 1–6, with equation 3 amended as follows to exclude persons who experienced infectious pulmonary disease in the meantime:

$$\frac{dL_{A,T}}{ds} = (I_{A,T}(5) + I_{s_i}(s,5))k_e(5) + r_i(s)P_{A,T}(s,2) + E_{s_i}(s,2) + E_{s_{i+1}}(s,2)) - (i(s) + d_i(A + s) + m_i(s))L_{A,T}(s)$$

The number of persons experiencing infectious pulmonary tuberculosis for the first time at time $s$ after being infected at age $A$ and at time $T$ is given by: $D_i(A,T,s) = d_i(s)(P_{A,T}(s,0) + E_{s_i}(s,0) + E_{s_{i+1}}(s,0))$. Of these persons, a proportion $(1 - m_i(s))$ survive for 2 years (and are infectious an average of 2 years) and a proportion $m_i(s)$ survived for 1 year only (and were infectious for only 1 year). The number of persons who infect others at time $s$ after having been infected at age $A$ and at time $T$ is given by:

$$S_i(A,T,s) = D_i(A,T,s - 1)m_i(s - 1) + D_i(A,T,s - 2)(1 - m_i(s - 2)).$$