Epidemiologic Basis of Tuberculosis Control

First edition
1999

Hans L. Rieder

International Union Against Tuberculosis and Lung Disease
68, boulevard Saint-Michel, 75006 Paris

The publication of this monograph was made possible thanks to the support of the World Health Organization, the Norwegian Heart and Lung Association, the British Columbia Lung Association, the Anti-Tuberculosis Association of Thailand, the Finnish Lung Health Association and the French Ministry of Foreign Affairs.
# Table of Contents

**Preface** ............................................................................................................................... 5

**Acknowledgements** ............................................................................................................ 7

**Introduction** .......................................................................................................................... 9

1. **Exposure to tubercle bacilli** ............................................................................................. 11

   - Number of incident cases .............................................................................................. 11
   - Duration of infectiousness ............................................................................................... 12
   - Number of case-contact interactions per unit of time .................................................. 12
     - Population density
     - Family size
     - Differences in climatic conditions
     - Age of sources of infection
     - Gender

2. **Infection with tubercle bacilli** ........................................................................................ 17

   - Etiologic epidemiology: risk factors for becoming infected given that exposure has occurred .................................................................................................................. 17
     - Airborne transmission through infectious droplet nuclei
     - Characteristics of an infectious patient
     - Air circulation and ventilation
     - Reducing expulsion of infectious material from source cases
     - Host immune response
     - Other modes of transmission: *M. bovis*

   - Descriptive epidemiology: incidence and prevalence of infection ............................... 26
     - Methodological issues in measuring infection .............................................................. 26
       - Evaluation of the sensitivity of the tuberculin skin test
       - Problems with the specificity of the tuberculin skin test in veterinary medicine
       - Standardization of tuberculin and dosage in humans
       - Variation in the specificity of the tuberculin skin test with a standard tuberculin dose
Approaches to determining the prevalence of tuberculous infection in the presence of cross-reactions attributable to environmental mycobacteria

Prevalence of infection .............................................................................................................................. 43

Prevalence of infection by age and sex
Prevalence of infection and population density
Tuberculous infection and socioeconomic indicators

Predictive epidemiology: modeling risk of infection ................................................ 49

Deriving risk of infection from prevalence of infection
Trends in risk of infection over time
Extrapolation from risk of infection to age-specific prevalence of infection
Risk of infection and infectious cases

3. Tuberculosis ............................................................................................................................... ............... 63

Etiologic epidemiology: risk factors for disease given that infection has occurred ............................................................................................................................... 63

Time elapsed since becoming infected
Infection with the human immunodeficiency virus (HIV)
Spontaneously healed tuberculosis with fibrotic residuals
Age
Genetic factors ............................................................................................................................... ............... 72

Sex
Body build
HLA types
Blood groups
Hemophilia
Virgin populations
Other genetic factors

Environmental factors ............................................................................................................................... 77

Substance abuse
Smoking
Alcohol abuse
Injection drug use
Nutrition
Malnutrition
Diet

Medical conditions ............................................................................................................................... 79

Silicosis
Diabetes mellitus
Malignancies
Renal failure
Measles
Gastrectomy
Jejunoileal bypass
Corticosteroid treatment

Pregnancy

Factors associated with the etiologic agent .................................................................................. 83
  Infecting dose effect
  Strain virulence
  Infection with *M. bovis*

Re-infection

**Descriptive epidemiology: morbidity** .................................................................................. 87

Secular trends in industrialized countries
Geographic differences in disease frequency
Differences in disease frequency in the community ................................................................. 94
  Age
  Sex
  Socioeconomic status
  Race and ethnicity
  Migration
  Population density
  Marital status
  Substance abuse
  Other risk groups

Impact of HIV infection............................................................................................................ 108
  United States and Western Europe
  Sub-Saharan Africa
  Asia
  Latin America

Contribution of primary progressive disease, exogenous re-infection, and endogenous reactivation to morbidity

Tuberculosis attributable to *M. bovis*

**Predictive epidemiology: prospects for the course of the epidemic** ...... 119

  Industrialized countries
  Low income countries

4. **Tuberculosis deaths** ........................................................................................................ 123

  Etiologic epidemiology: risk factors for death from tuberculosis
  (case fatality) .................................................................................................................. 123
Efficient tuberculosis control can be achieved in the absence of an extensive theoretical background. However, a thorough understanding of the etiologic agent, the clinical presentation of tuberculosis, the epidemiology of tuberculosis, the role of various intervention strategies, and of how to efficiently apply the tools currently available for the control of tuberculosis, is likely to increase the efficiency of a national tuberculosis program. A theoretical background will also help program managers at all levels to base their practice on modern concepts of tuberculosis control, and to justify their actions when these are questioned by others.

The International Union Against Tuberculosis and Lung Disease (IUATLD) organizes international courses on tuberculosis control. These courses are largely aimed at health care personnel with responsibility at regional or national level within national tuberculosis programs. An attempt is made to provide the participants with a thorough theoretical knowledge of the basis of modern tuberculosis control strategies. The course content encompasses the following major components:

- Bacteriologic basis of tuberculosis control
- Clinical presentation and diagnosis of tuberculosis
- Epidemiologic basis of tuberculosis control
- Intervention strategies in tuberculosis control
- Principles of tuberculosis control (figure 1).

Figure 1. The five elements which are necessary for the understanding of modern tuberculosis control.
Course participants are provided with a selected tuberculosis library containing approximately 100 original articles (continuously updated) which constitute the essential basis of the modules.

The five components are logical and coherent, each one building upon the others (figure 1). Course participants first learn in depth about the etiologic agent, *Mycobacterium tuberculosis* complex, i.e., the bacteriologic basis of tuberculosis control. Second, they see what the etiologic agent does to the individual, i.e., the clinical presentation and diagnosis of tuberculosis. Third, they are taught how it affects the community, i.e., the epidemiologic basis of tuberculosis control. From these first three core facets, interventions to protect the individual and the community, i.e., intervention strategies, are discussed. Finally, all four components are combined with implementation in mind and integrated into a single package for control: the principles of tuberculosis control.

Despite the wealth of original literature provided, course participants have repeatedly voiced their interest in obtaining illustrated monographs with the most important figures/graphics from each module. This monograph on the epidemiologic basis of tuberculosis control is an attempt to meet this need for this particular module.
Acknowledgements

The author wishes to express his deepest appreciation to the hundreds of participants in the courses for which this material was developed and in which it was tested. The interaction among the participants over a number of years has strengthened the logic, improved the flow and increased the comprehension of epidemiology as a basic science for the control of tuberculosis.

Martien W Borgdorff, George W Comstock, and Kevin M De Cock made major contributions to improve the content of this monograph. Thuridur Arnadottir, Christopher Dye, Donald A Enarson, and Arnaud Trébucq have critically reviewed the document and provided valuable suggestions for its improvement. Robert L Cowie and Robert Steffen positively commented on the text of the monograph. Clare Pierard provided support in the final edition of the manuscript.

The production and distribution of this monograph would not have been possible without the generous financial contributions of the World Health Organization’s Global Tuberculosis Programme, the Norwegian Heart and Lung Association, the British Columbia Lung Association, the Anti-Tuberculosis Association of Thailand, the Finnish Lung Health Association and the French Ministry of Foreign Affairs.

Paris and Berne, July 1999

First reprint with correction of minor errors.

The first edition went out of print swiftly and required a reprint. The occasion was seized to eliminate a few blatant errors that had crept into the first printing. However, no attempt was made at this time to update the material, as it was felt to be still relevant as presented.

Paris and Berne, April 2003
Introduction

Epidemiology is the basic science of preventive medicine and public health. It is relevant for those involved in clinical practice, research, and public health policy.

This monograph deals with the epidemiology of tuberculosis. It attempts to delineate the determinants of exposure, infection, disease, and death from tuberculosis. To convey the conceptual framework of the epidemiology of tuberculosis, a selection of pertinent examples published in the biomedical literature is provided. This approach should give the basis for an intuitive and rational comprehension of the dynamics of the epidemic.

To facilitate the understanding of the relevance of the epidemiology of tuberculosis as a basis for implementing a successful national tuberculosis control program, a model following the pathogenesis of tuberculosis from exposure to death is useful. The model should be complex enough to allow the distinction of the major determinants of the epidemiology of tuberculosis, but simple enough to delineate the major indicators that are essential for comprehending the dynamics of the disease. Such a model, derived from the tuberculosis classification of the American Thoracic Society and the United States Centers for Disease Control (CDC, now Centers for Disease Control and Prevention) [1], is proposed in figure 2 [2]. According to this model, four distinct steps in the pathogenesis of tuberculosis can be identified: exposure, infection, disease, and death.

![Figure 2](image-url)
Exposure to a potentially infectious case is a prerequisite for becoming infected. Once an individual is exposed, there are factors which determine the risk of becoming infected, risk factors which determine the probability that an infected individual will develop tuberculosis, and risk factors which determine the probability that a diseased individual will die from tuberculosis.

The epidemiologic framework for the approach to understanding the dynamics of tuberculosis in a community that is proposed here pertains, simplified, to three types of epidemiologic questions. These are denoted here as analytic epidemiology or, to use an intuitively more appealing term, etiologic epidemiology [3]. Etiologic epidemiology aims at disentangling and identifying factors which increase the likelihood of progression from one stage to the next. Descriptive epidemiology of tuberculosis outlines the frequency and distribution of infection, disease, and death from tuberculosis in different populations. Finally, predictive epidemiology uses modeling techniques to forecast the likely course of the tuberculosis epidemic in a given community, based on observations from the past. In this monograph, exposure, infection, disease, and death are discussed sequentially. While in other texts descriptive epidemiology is often discussed before etiologic epidemiology, it seemed more appealing to reverse this order here and begin each pathogenetic step with etiologic epidemiology, followed by descriptive epidemiology, and finally predictive epidemiology. This approach is chosen because one needs first to know the factors leading up to the current situation to understand the reasons for its magnitude, from which the prospects for the epidemic might then be derived.
The major factors that determine the risk of becoming exposed to tubercle bacilli include the number of incident infectious cases in the community, the duration of their infectiousness, and the number and nature of interactions between a case and a susceptible contact per unit of time of infectiousness.

Given a defined number of infectious sources, the number of persons who might be exposed to the tubercle bacilli from these cases may vary considerably, depending on the duration of their infectiousness and the number and nature of possible case-contact interactions per unit of time of infectiousness. It is not easy to define “exposure”, because in its widest sense all human beings are exposed to the same air space. What is needed, therefore, is a pragmatic definition of exposure. In this monograph, the underlying definition of significant exposure is contact between two individuals in sufficient proximity to allow conversation between them, or, within confined spaces, where the air exchange (ventilation) of the space has been incomplete between the visits of two persons.

In rural areas, for example, the number of persons who become exposed to a single case might be considerably lower than in urban areas. Similarly, a case occurring in a household may expose few or many persons, depending on the size of the household. Cases are more likely to expose people with whom they tend to socialize, e.g., people of a similar age or social group. Climatic conditions may affect behavior substantially by influencing the amount of time spent outside confined spaces.

**Number of incident cases**

Without the presence and emergence of infectious cases, no relevant exposure will occur. However, while the incidence level of infectious cases is a prerequisite in determining exposure, the risk to which incident cases of tuberculosis expose susceptible persons is modified by a number of factors. Thus, at the same level of
incidence, risk of exposure may still vary greatly. The following modifiers of exposure risk can be identified, given a constant number of incident sources of infection in the community.

**Duration of infectiousness**

The duration of infectiousness of an incident infectious case is of crucial importance for the risk of the general population becoming exposed to such a case. The risk of becoming exposed is greatly enhanced if infectiousness is prolonged, as compared with a short duration of infectiousness. The point in time at which infection during the transmission period occurs is, of course, important, as repeated superinfections from the same source case might be redundant and of little epidemiologic significance. However, as the proportion of contacts found to be infected at the time of the diagnosis of a sputum smear-positive index case is around 30 to 40 per cent [4], a large proportion remains who would still be susceptible to new infection if the case were not detected (and placed on adequate chemotherapy).

An untreated case of infectious tuberculosis will remain infectious for a longer period of time than a patient who is diagnosed and treated appropriately in a timely manner. Early intervention with appropriate chemotherapy reduces the time of infectiousness, and thus the potential that members of the community continue to be exposed. Of similar concern for prolonged infectiousness are patients who are diagnosed and started on chemotherapy, but where chemotherapy is inadequate. Inadequate chemotherapy includes improper combinations of drugs or inadequate dosages of anti-tuberculosis drugs, or occurs when patients have strains resistant to one or more of the antimicrobial agents used [5-7] or when patients become irregular or selective in ingesting the prescribed medications.

**Number of case-contact interactions per unit of time**

The number and nature of possible case-contact interactions will vary greatly according to individual behavior and opportunities for interacting with other people in the community. Obvious factors, which vary by time and geographic location, include the following.

**Population density**

Population density varies greatly between and within countries. The nature of the dwelling in which people live and the number of persons who share that dwelling have an important impact on the risk of exposure, given that a case of tuberculosis
lives in that dwelling. In addition, urban areas have a much higher population density than rural areas. The number of possible contacts of a tuberculosis patient in a rural area may be smaller per unit of time than that of a patient living in an urban setting (given a similar dwelling and family size). Thus, the likelihood that a susceptible person will be exposed to an infectious tuberculosis patient increases with population density, even if the incidence is the same.

**Family size**

Family size changed considerably in industrialized countries during the twentieth century. In 1900, the average private household in Switzerland, for example, had close to two children present, and by 1990 the number of children present per family household had decreased below one [8]. Clearly, the number of exposed children must decrease under such circumstances, even if the probability of infectious tuberculosis occurring in a parent remains the same. Of further importance are social arrangements within the family, i.e., the sleeping arrangements or responsibilities for childcare.

**Differences in climatic conditions**

In a warm climate, outdoor social activities are much more common than in a colder climate, such as in countries in the north of Europe, which have long, cold winters. Tubercle bacilli expelled outdoors are dispersed rapidly, and exposed to sunlight they die very quickly due to the sun’s ultraviolet rays. In contrast, tubercle bacilli expelled indoors by a patient in a confined space with poor ventilation may retain viability, and thus the potential to cause infection, for a prolonged period of time. People entering such a room may become exposed even after the source patient producing the droplets has left the room. A cold climate causes people to congregate inside, thus increasing the likelihood of exposure if there is a case of tuberculosis in the group. In contrast, indoor activities might be reduced in temperate or tropical climates, and indoor ventilation might be better, because windows can be kept open for much longer periods than in colder climates.

**Age of sources of infection**

The most intense exposure is likely to occur among persons who share the same household or who spend long periods of time in the same room with an infectious source case (for example, those living in a closed institution). In industrialized countries, parents expose their children to a much larger extent than do their grandparents, who rarely live in the same household. Similarly, as the median age of tuberculosis patients in industrialized countries has progressively increased over the
past few decades, with most cases of disease occurring now in the elderly (see chapter 3), the chance that children will be exposed must necessarily have decreased. This is quite clearly depicted in figure 3 [9]. Within just 30 to 40 years the median age of tuberculosis patients in Finland increased by 30 years, moving from the generation of parents to that of grandparents.

Patterns of social participation are key to the transmission dynamics of tuberculosis [10, 11]. It is a common notion that people tend to socialize with those of their own age: it is highly unusual to see a septuagenarian mingle with teenagers heading for a disco. A notable exception is a household of parents with small children. Here, two generations closely interact, and this social interaction goes beyond the family: the children’s friends closely interact, and this social interaction goes beyond the family: the children’s friends and their parents often join in as well.

There is no reason to believe that tuberculosis patients will behave differently in this respect from the non-affected population. Indeed, there is evidence that the risk of tuberculous infection varies considerably with age [12], an important observation, yet not in itself sufficiently conclusive to accept the hypothesis of preferential intra-generational transmission. The generation raising children has both important inter- and intra-generational contacts, but with the aging of this birth cohort an increasingly larger proportion of their children is released from parental custody, and contacts

\[ \text{Figure 3. The median age of new cases of respiratory tuberculosis among males and females in Finland, 1954 – 1995, adjusted for changes in the age structure of the population. Data from [9].} \]
become dominated by intra-generational social interaction. The epidemiology of tuberculosis morbidity in Europe has shown a rapid increase in the median age of tuberculosis patients. Even if the crude incidence of tuberculosis were the same, the risk of transmission to the youngest generation must surely decrease in this case, since elderly patients are less involved in child-rearing than younger patients. Furthermore, while the risk of tuberculous infection may remain high in the oldest generation, as has been demonstrated in nursing home residents in Arkansas, United States [13], the consequences are not nearly as grave as when tuberculosis predominantly occurs among the young, for two reasons. First, transmission is more likely to be predominantly intra-generational, and second, the cumulative probability of ultimate progression to disease is markedly reduced by the shorter remaining potential life span.

**Gender**

The degree of social interactions differs by gender to a great extent in different societies. In some countries, women and men take part almost equally in public activities, while in other countries women’s lives are very secluded. The opportunity of becoming exposed to an infectious case both inside and outside of the home will thus differ for men and women.
Etiologic epidemiology: risk factors for becoming infected given that exposure has occurred

The probability of becoming infected with *Mycobacterium tuberculosis* depends on the number of infectious droplet nuclei per volume of air (infectious particle density) and the duration of exposure of a susceptible individual to that particle density.

*Airborne transmission through infectious droplet nuclei*

The risk of becoming infected with tubercle bacilli (*M. tuberculosis* complex, which includes the pathogenic species *M. tuberculosis*, *M. bovis*, and *M. africanum*) is largely exogenous in nature.

That primarily an airborne route transmits tuberculosis was suspected by Koch in his seminal line of reasoning for the proof of tubercle bacilli as the cause of tuberculosis [14, 15]. Because he recognized that the majority of cases began in the respiratory tract, he deduced that “…bacilli are usually inspired with the air…” [15]. The experimental demonstration was nevertheless made only decades later [16].

A major breakthrough in the understanding of airborne transmission came with the work of Wells at the Harvard School of Public Health [17]. To be transmissible through the air, the infectious agent must remain buoyant in the air. The velocity of a droplet falling to the ground is proportional to the surface area of the droplet or to the square of its diameter. In moisture-saturated air all but very small droplets would fall to the ground from a height of two meters in less than 10 seconds (figure 4) [17]. The time such larger droplets spend in the air is thus too short for them to be inhaled by a susceptible person. The most important characteristic of liquid droplets is their tendency to evaporate. Evaporation causes the volume of water droplets to diminish in size, and this change becomes more rapid as the droplet decreases in size. The duration of time droplets remain in unsaturated air is inversely proportional to
the square of their diameter. It is consequently very short for small droplets, much shorter than the time it takes for them to fall from a height of two meters if they settled without evaporation. Approximate evaporation times for droplets of different sizes in unsaturated air are shown in figure 5 [17].

Very small droplets thus settle slowly and evaporate almost immediately, whereas large drops settle rapidly and reach the ground without appreciable loss by evaporation. Wells gives the example of raindrops, which are sufficiently large in size to travel from the clouds to the earth, while in contrast spray drying is used in industry using very small droplets. Wells’ concept is depicted schematically in figure 6 [17]. The curves of evaporation and falling time (falling two meters in unsaturated air) are very similar in shape but in the opposite direction. Somewhere between 0.1 mm and 0.2 mm lies the droplet size which reaches the ground from a height of 2 meters, while smaller particles evaporate before reaching the ground.

The falling times in figures 5 and 6 assume droplets of fixed size. However, evaporation creates droplets of diminishing size, which may become droplet nuclei. Wells demonstrates that the distance a droplet will fall before ceasing to be a droplet is proportional to the square of the surface or the fourth power of the diameter [17].

![Graph](image.png)

**Figure 4.** Settling of water droplets in moisture-saturated air. Data from [17].
**Figure 5.** Evaporation time of water droplets in unsaturated air at 18° C. Data from [17].

**Figure 6.** Falling time of droplets and evaporation as a function of droplet size diameter. Figure reproduced with the permission of the American Journal of Epidemiology from [17].
Other factors, principally temperature and humidity, also enter into the equation. In high humidity, evaporation is less pronounced than in low humidity: thus small droplets of a given size are more likely to settle in humid than in dry air. The speed of fall depends on the first power of factors of humidity and temperature, but on the fourth power of the surface area. Thus, a minor change in droplet size is more important than a major change in atmospheric conditions.

Droplets produced by a patient with tuberculosis may contain tubercle bacilli. Tubercle bacilli may still be contained in droplets below the critical size of settling before evaporation. Thus, such droplets may evaporate to the size of infectious droplet nuclei containing one or more tubercle bacilli that remain buoyant in the ambient air for a prolonged period of time. It is these infectious droplet nuclei which are the principal source of transmission of *M. tuberculosis*.

Talking, coughing, sneezing or singing produces droplets. The landmark experiment of Loudon and Roberts has greatly contributed to the understanding of the transmission of *M. tuberculosis* [18]. They demonstrated that one cough was the equivalent of about 5 minutes of loud talking in terms of the resulting number of droplet nuclei, about half of which were still suspended in air 30 minutes after coughing.

These physical properties are important, because *M. tuberculosis* is almost exclusively transmitted by the airborne route, apart from accidental direct inoculation of infectious material into the skin [19-22]. Successful transmission requires airborne infectious droplet nuclei, small enough to reach an alveolus in the periphery of the lung. Such particles may remain suspended in the air for several hours. Larger particles generally drop more rapidly to the ground or, if inhaled, are trapped in the mucociliary system of the tracheobronchial tree, swept up and rendered harmless through swallowing. Sonkin found that the majority of particles above 5 µm in diameter are trapped in the nose, while those smaller than 0.1 µm tended to remain suspended and able to reach the alveoli [23]. Conversely, retention of particles decreased with diminishing size. Thus, there is a critical diameter range that maximizes the probability of inhalation and retention of infectious particles leading to the establishment of infection. This diameter size lies somewhere between 1 µm and 5 µm.

Riley and co-workers then conducted a series of studies on experimental airborne tuberculosis [16, 24]. A tuberculosis ward with single rooms in which there were tuberculosis patients was connected through a carefully controlled and calibrated closed circuit ventilating system to a large animal exposure chamber with healthy guinea pigs located in the exhaust of these ducts. Monitoring of these animals and examination of their organs once they were killed allowed quantification of the number of infectious droplets in the air that led to successful transmission. On average, one
infectious droplet nucleus was present in 340 cubic meters of air. The time was estimated for a nurse to breathe this amount of ward air (and inhale one infectious particle). It was estimated that infection (and hence tuberculin conversion from negative to positive) would take, on average, one [25] to one and a half years [26].

**Characteristics of an infectious patient**

For all practical purposes, for successful transmission to take place, a tuberculosis patient must be able to produce airborne infectious droplets. This limits the potential of transmission, as a rule, to patients with tuberculosis of the respiratory tract. Among patients with tuberculosis of the respiratory tract, not all are equally efficient at transmission.

The number of bacilli found in sputum specimens correlates well with the potential of infectiousness. It requires some 5,000 bacilli in 1 mL of sputum to yield a positive smear examination with a reasonable chance [27] and some 10,000 to identify a smear as positive with a 95 per cent probability [28]. For this reason, smear microscopy is a good (sensitive) test for identifying the most infectious cases; it is not, however, a very sensitive test for diagnosing tuberculosis. Among several

![Figure 7. Infectiousness of pulmonary tuberculosis by bacteriologic status of source case. s+/c+ indicates smear- and culture-positive; s-/c+ indicates smear-negative, culture-positive; and s-/c- indicates negative on both smear and culture. Data from [29-31].](image)

- 21 -
studies, three major ones are exemplary in demonstrating the relative infectiousness of sputum smear-positive, sputum smear-negative, culture-only positive, and culture-negative tuberculosis [29-31]. Each of these studies confirmed that patients with sputum smear-positive tuberculosis of the respiratory tract are by far more infectious than patients whose smears were negative, but whose cultures were positive (figure 7). This was evidenced in the proportion of children in close contact with the source case who had a positive tuberculin skin test (figure 8). The absolute differences in prevalence of infection in the three studies are not relevant; they can be explained by differences in the background prevalence and different definitions of a contact. The important point is the relative difference between the prevalence of infection among those children exposed to a sputum smear-positive versus a sputum smear-negative, culture-only positive case. These differences are large and are consistent across all three studies. It appeared, in fact, that patients with culture-only positive tuberculosis were not much more infectious than patients with neither smear- nor culture-positive tuberculosis or control cases. These studies demonstrated convincingly that patients with sputum smear-positive tuberculosis are by far the most potent sources of infection in the community.

In a study in Finland, contacts of patients with sputum smear-positive and sputum smear-negative tuberculosis (all confirmed by culture) were followed up for 2 years [32]. Among the 609 contacts of 134 index cases, four developed tuberculosis during the follow-up period, all of them among contacts of the 69 cases with heavily sputum smear-positive tuberculosis. However, it can not be concluded from these studies that patients whose sputum is negative on direct microscopic examination are not transmitters at all. A study in San Francisco, California, using molecular fingerprinting techniques demonstrated that 17 per cent of transmissions were attributable to index cases with sputum smears negative at diagnosis, and that the relative transmission rate from smear-negative compared to smear-positive tuberculosis was 0.22, or roughly one fifth of all transmissions [33].

While the probability of being infected following contact with an infectious source decreases with decreasing proximity of the contact person to the index case [34], the absolute number of persons an infectious case may infect outside the context of close contact might be considerable, and may exceed the number of infected close contacts. This occurs when the number of casual contacts of an infectious case far exceeds that of susceptible close contacts. This is illustrated schematically in figure 9, which shows that the probability of infection among the closest, inner-circle contacts was three out of 10, in the next to closest contact circle three out of 20, and much lower in casual contacts beyond these two inner circles. The absolute number of people infected outside the closest contacts, however, exceeded the number among close contacts.
Figure 8. Infectiousness of tuberculosis by bacteriologic status of and proximity to source case. Data from [30].

Figure 9. Example for the risk of infection among close contacts (inner circle), intermediate contacts (outer circle), and casual contacts. The index case is indicated by a cross, infected contacts by filled circles, and uninfected contacts by hollow circles.
Air circulation and ventilation

Given a defined number of tubercle bacilli expelled into the air, the volume of air into which the bacilli are expelled determines the probability that a susceptible individual breathing that air will become infected. Droplet nuclei have negligible settling tendency, disperse rapidly throughout the air of a room, and are carried wherever air currents take them [35].

Tuberculosis is not as easily transmitted as are certain other airborne infectious diseases [36]. Riley has estimated that on average 12 to 18 months of exposure to tuberculosis patients were required for a student nurse to become infected in the pre-chemotherapy era [25, 26].

A study of a tuberculosis outbreak on a United States Navy ship provided information on the importance of ventilation and air circulation [37, 38]. On this ship, 139 of 308 persons previously known to be tuberculin skin test negative ultimately became infected from a single source. Men sharing the same compartment with the case had the highest frequency of skin test conversions. Nevertheless, a very high proportion of those living in other compartments and of those who did not work or significantly socialize with the index case, but whose compartments were connected through a close circuit ventilation system, also became infected. Men with similar social exposure but living in compartments without direct exposure to the air of the ventilation system converted much less frequently. Conversions continued for a prolonged period of time after removal of the infectious source case, suggesting that air in closed circuits may contain viable infectious droplet nuclei for a prolonged period of time.

Ventilation dramatically dilutes the concentration of infectious droplet nuclei. Wherever possible, opening windows during and after cough-inducing procedures or in hospital wards is one of the most, if not the most, efficient means of reducing the probability of exposed persons becoming infected.

Wearing surgical masks by exposed persons is likely to be of low efficiency, because most masks neither filter out particles of less than 5 µm, nor do they generally fit snugly enough around both mouth and nose. In industrialized countries, the emergence of multidrug-resistant tuberculosis, and exposure of staff in hospitals with often poor ventilation and usually little to no direct exchange with fresh outdoor air, has led to the recommendation for staff to wear a special mask, called a high-efficiency particulate air-filter respirator [39]. This mask is designed to filter out particles in the droplet nucleus size of 1 µm to 5 µm. It costs several times the price of usual surgical masks, and data on how snugly they fit to prevent droplet nuclei from
entering between skin and mask are scarce. The financial repercussions of implementing a policy of wearing particulate respirators may be huge [40].

**Reducing expulsion of infectious material from source cases**

Preventing infectious source cases from producing infectious droplets is the most effective means of reducing the probability of transmission. Covering the mouth (and nose) during coughing is likely to be very efficient in reducing the number of infectious droplets which can reach the air. The most effective intervention for reducing infectiousness is treatment of cases, which reduces the infectiousness even of sputum smear-positive cases within a matter of a few weeks [4, 41-44].

**Host immune response**

All of the factors already discussed relate to an underlying hypothesis that the risk of becoming infected is entirely based on factors exogenous to the host. In its narrowest sense this means that for the establishment of tuberculous infection it is sufficient for a single tubercle bacillus to adhere to the alveolar cell wall. However, this is not necessarily true, and actually quite unlikely. A tubercle bacillus adhering to the alveolar cell wall faces two potential outcomes: either it manages to establish latent tuberculous infection, or it is removed before being able to do so. After being ingested by alveolar macrophages, bacilli are incorporated into phagosomes and subject to killing via a variety of mechanisms, including phagosome-lysosome fusion, generation of oxygen radicals, and generation of reactive nitrogen intermediates [45]. Quite obviously, macrophage function may vary from individual to individual and over time, thus being a modifier of the risk associated with implantation of tubercle bacilli in their ability to establish tuberculous infection.

**Other modes of transmission: M. bovis**

Milk, an essential part of the diet in many cultures, has been responsible for much infant mortality in the past, and *M. bovis* as a cause of tuberculosis is one of the most prominent pathogens carried in this vehicle [46].

As early as 1846 it was observed that peripheral lymphatic tuberculosis was more frequent in children who were given cow’s milk than in children who were breast fed [47]. Koch ran the danger of increasingly isolating himself in the academic community at the beginning of the twentieth century by downplaying the role of bovine tubercle bacilli in the creation of disease in man [47, 48]. Veterinarians, particularly in the United States, had already in the first decade of the twentieth century begun a program of tuberculin skin testing dairy cattle and slaughtering those reactors
identified [48]. In Europe, the battle against bovine tuberculosis began much later. In Britain, in 1932, 40 per cent of animals had evidence of tuberculosis, and 0.5 per cent of all dairy cows produced milk containing tubercle bacilli [47]. Tuberculin skin testing and slaughtering of infected animals in Britain was introduced systematically in 1935 [47]. By 1965, only 1 per cent of herds remained infected [47]. The site of disease in cattle was thoracic in the majority of cases (80 to 90 per cent) and, in descending order of frequency, liver, kidneys, spleen, uterus, and udders, the latter occurring in 1 to 2 per cent of diseased animals [47]. Dairy cattle with tuberculous mastitis are the source of infected milk that can cause bovine bacilli to be transmitted to humans by ingestion of milk if it is not pasteurized or boiled.

The relation between the frequency of infected herds and tuberculosis in humans was studied extensively by Magnus in Denmark [49-51], where, in 1937, almost 50 per cent of dairy herds were infected with tubercle bacilli, rapidly decreasing to 0.3 per cent in 1950 (figure 10) [49]. The figure also demonstrates that the occurrence of tuberculosis in cattle varied greatly throughout the districts of the country. A very strong correlation between the proportion of infected herds and human tuberculin reactors was demonstrated (figure 11) [50]. Thus, \textit{M. bovis} was an important contributor to the overall prevalence of tuberculous infection.

Conversely, cattle, sheep, goats, and horses may be infected with \textit{M. tuberculosis}, but in most cases the disease is self-limiting in animals [28, 52].

There is strong evidence that \textit{M. bovis} is much more difficult to transmit from man to man than \textit{M. tuberculosis} [53, 54].

**Descriptive epidemiology: incidence and prevalence of infection**

**Methodological issues in measuring infection**

Tuberculin, a concoction of antigens produced from killed tubercle bacilli, was first introduced by Koch, who thought he had identified a specific treatment for tuberculosis [55, 56]. He also recognized the potential of this tuberculin as a diagnostic tool [56]. While the hopes for tuberculin as a treatment modality were soon to disappear, tuberculin has remained an important diagnostic tool for the identification of infection with \textit{M. tuberculosis}. Because the tuberculin skin test is the only currently available tool to determine infection, and tuberculous infection takes center stage for providing an understanding of the dynamics of the tuberculosis epidemic in the community, it is essential that this test and its development are discussed in some detail.
Figure 10. Frequency of tuberculous herds in Denmark in two selected counties, 1937–1950. Figure reproduced from [49].

Figure 11. Correlation between percentage of tuberculous herds (1937–1939) and percentage of tuberculin reactors in the human population (1950–1952), Denmark. Figure reproduced with the permission of the World Health Organization from [50].
The history of the tuberculin skin test can be broadly grouped into three periods. In a first period at the beginning of the twentieth century, research on tuberculin concentrated on the sensitivity of the test. In a second phase, research sought to address the problem of the test’s specificity, and today, research is concentrating on both characteristics depending on the expected and observed prevalence of infection, and thus the predictive value of the test under different circumstances. Attempts to identify a better tool to determine infection more accurately, and to distinguish infection from disease with such a test, have so far remained unsuccessful.

**Evaluation of the sensitivity of the tuberculin skin test**

The sensitivity of a test is defined as the proportion of people correctly identified with a given characteristic (table 1) [57]. A high sensitivity is of particular importance if a characteristic needs to be excluded with high certainty. For example, provision of a blood transfusion supply free of transmissible agents requires tests of a high sensitivity.

At the beginning of the twentieth century, von Pirquet introduced a cutaneous tuberculin skin test [58]. In his technique, tuberculin was dropped on the cleaned skin, and the skin was subsequently scratched. He found the test’s sensitivity to be less than 100 per cent in serious forms of tuberculosis.

In the same decade, Moro introduced a simplified cutaneous test, using tuberculin cream rubbed onto the skin [59]; he found the sensitivity of results to be comparable with the von Pirquet method.

Mendel [60] and Mantoux [61] almost simultaneously introduced the intradermal technique, allowing the administration of an exact dose of tuberculin with needle and syringe.

While all three techniques have survived in modified form, worldwide preference is now given to the intradermal technique. Official recommendations list only the intradermal technique for epidemiologic use in surveys [62-66]. The reasons are obvious if one examines the difficulties even with this best quantifiable technique.

**Problems with the specificity of the tuberculin skin test in veterinary medicine**

The specificity of a test is defined as the proportion of people correctly identified without a given characteristic [57]. Conversely, a high specificity is of particular concern if one needs to be certain about the presence of a condition. For example, before informing a patient of infection with the human immunodeficiency virus (HIV), one must be certain that the patient is truly infected, to prevent any
consequences for the patient of the knowledge of having a lethal disease when the diagnosis is based on an erroneous test result.

The importance of tuberculosis in cattle was early recognized in the United States, and a rigorous program of culling infected cattle was implemented. The decision to slaughter an animal was made based on the presence of a positive tuberculin skin test. With the progression of the eradication program, it was noted that an increasing proportion of tuberculin-positive cattle had no detectable tuberculous lesion [67-69]. These observations indicate early considerations of the dependence of the predictive value of a positive test result on the prevalence of the condition: Hastings recognized that the frequency of no-lesion reactors increased as the prevalence of reactors decreased (figure 12) [68]. This finding was interpreted as the result of reactions caused by the presence of environmental mycobacteria.

<table>
<thead>
<tr>
<th>Test result</th>
<th>True presence or absence of characteristic</th>
<th>Present</th>
<th>Absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td></td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>a+c</td>
<td>b+d</td>
<td>N=a+b+c+d</td>
</tr>
</tbody>
</table>

\[
\begin{aligned}
a / (a+c): & \quad \text{Sensitivity of test (proportion correctly identified among those with the characteristic)} \\
d / (b+d): & \quad \text{Specificity of test (proportion correctly identified among those without the characteristic)} \\
(a+c) / N: & \quad \text{Prevalence of characteristic} \\
a / (a+b): & \quad \text{Predictive value of a positive test} \\
d / (c+d): & \quad \text{Predictive value of a negative test} \\
b / (a+b): & \quad \text{Proportion false positive among positive test results (relative false positive)} \\
b / (b+d): & \quad \text{Proportion false positive among those without the characteristic (absolute false positive)} \\
c / (c+d): & \quad \text{Proportion false negative among negative test results (relative false negative)} \\
c / (a+c): & \quad \text{Proportion false negative among those with the characteristic (absolute false negative)} 
\end{aligned}

Table 1. Operating characteristics of a test.
Standardization of tuberculin and dosage in humans

With some delay, it was recognized that the tuberculin skin test posed similar problems in humans as in veterinary medicine. To increase the specificity, a purification of tuberculin was attempted, because it was recognized that the preparation then in use (Old Tuberculin) contained varying proportions of a number of different proteins depending on the production technique. Florence Seibert is credited with the purification of tuberculin and of establishing a mechanism for international standardization [70-72]. The tuberculin produced by her was recommended by the World Health Organization (WHO) in 1951 to be the international standard [73]. Lot No. 49608 of tuberculin Purified Protein Derivative (PPD) is the international standard, called PPD-S (S denoting “Standard”) or PPD-M (M denoting “Mammalian”) [74].

The international unit of tuberculin is defined as the biological activity contained in 0.000028 mg of the standard preparation, Lot 49608, consisting of 0.00002 mg PPD plus 0.000008 mg salts. A portion of the referent is kept in Copenhagen, Denmark, at the Statens Serum Institut, and is distributed in volumes of 500,000 units in lyophilized form. Each new tuberculin produced for sale should be calibrated against this standard in guinea pigs.

Because sensitivity and specificity are interdependent, and a gain in sensitivity is offset by a loss of specificity and vice versa, it was necessary to determine the optimal dosage with the already standardized form of application (intradermal technique) and the standardized tuberculin PPD. Persons stratified into groups with and without a history of contact with a tuberculosis patient received sequentially increasing doses of tuberculin, calculating the cumulative proportion of reactors (figure 13) [75]. Persons with a history of contact reacted to relatively small doses of tuberculin, and their reactivity could not be enhanced significantly after a certain dose was administered. In contrast, only a small proportion of persons without a history of contact reacted to small doses, but could be provoked to react to larger doses.

Perhaps even more convincing was the demonstration of the differences in the frequency of reactions among healthy children and tuberculosis patients [75]. The children were chosen from an orphanage in Ohio, United States, an environment with little chance for the children of exposure to tuberculosis [76]. The tuberculosis patient group consisted of the entire patient group with confirmed tuberculosis but who were not critically ill. One ten-thousandth of a milligram of tuberculin PPD-S caused a positive reaction in nearly every tuberculosis patient, but in only 20 per cent of the unexposed children who, incidentally, had been raised in an area where cross-reactions due to environmental mycobacteria were not uncommon (figure 14) [75].
Figure 12. Correlation between prevalence of tuberculin-positive cattle and the frequency of negative autopsy findings. Data from [68].

Figure 13. Cumulative frequency of reactors responding to increasing doses of tuberculin, stratified by history of contact with a tuberculosis patient. Data from [75].
These findings were confirmed in another study [77], where the relative risk of reacting to a higher than the established standard dose of 5 TU (denoting “Tuberculin Units”) of PPD-S was much greater among persons without a history of contact than among persons with such a history of contact (figure 15). From these studies it was concluded that the dose that struck the best balance between sensitivity and specificity was 5 TU PPD-S.

In 1955, the Statens Serum Institut in Copenhagen, Denmark, agreed at the request of UNICEF to produce a large batch of tuberculin for international use; it was called Tuberculin PPD RT 23 [78, 79]. A stabilizing agent, Tween 80 (a polyoxyethylene derivative of sorbitan mono-oleate) was added to the tuberculin. This resulted in a stronger potency, approximately double that of tuberculin PPD-S. In a subsequent study in the United States, a comparison was made between 5 TU tuberculin PPD-S and different doses of tuberculin PPD RT 23 [80]. In this study it was demonstrated that 2 TU tuberculin PPD RT 23 and 5 TU tuberculin PPD-S resulted in a similar distribution of tuberculin skin test reaction sizes in healthy Eskimo children with a high prevalence of tuberculous infection and a low prevalence of infection with environmental mycobacteria (figure 16). The two doses also elicited similar distributions in United States Navy recruits with a relatively low prevalence of tuberculous infection and a relatively high prevalence of infection with environmental

![Figure 14](image-url)

**Figure 14.** Cumulative frequency of reactors responding to increasing doses of tuberculin, among healthy children and patients with tuberculosis. Data from [75].
**Figure 15.** Relative risk of reacting to a higher than the standard dose of tuberculin by history of contact. Data from [77].

**Figure 16.** Frequency distribution of tuberculin skin test reaction sizes with 5 TU Tuberculin PPD-S (solid line) and 2 TU Tuberculin PPD RT 23 (dotted line) among Eskimo children and United States Navy recruits. Figure reproduced with the permission of the World Health Organization from [80].
mycobacteria. However, the authors noted that 2 TU tuberculin PPD RT 23 identified a considerably larger proportion (40 per cent) of persons with intermediate sized reactions (6 mm and more), among recruits from the south-eastern States, indicating a lower specificity of tuberculin PPD RT 23 in areas with a particularly high prevalence of infection with environmental mycobacteria [80].

**Variation in the specificity of the tuberculin skin test with a standard tuberculin dose**

Despite a standardized technique of administration (the intradermal technique) and a standardized dose (equivalent to 5 TU tuberculin PPD-S), large variations in the distribution of reaction sizes can be observed, often within the same country [81]. This is exemplified in the surveys in Minnesota and North Carolina, United States (figure 17) [82]. While in Minnesota the infected were fairly easily separated from the non-infected in virtually every age group, it was much more difficult to do so in North Carolina.

The largest tuberculin survey ever carried out was conducted among United States Navy recruits from 1958 to 1965, when almost 700,000 recruits were tested [76]. As part of the study, a group of 5,544 tuberculosis patients were also tested with 5 TU tuberculin PPD-S, and the distribution of reactions amongst these patients was very close to normal, with a peak frequency of around 16 mm to 17 mm (figure 18). Similar results of a normal distribution among tuberculosis patients with a mode located at a similar point were found in earlier WHO-sponsored tuberculin surveys among tuberculosis patients [83]. In the United States study Navy recruits were asked whether they had ever had previous contact with a tuberculosis patient. The distribution of tuberculin skin test reactions in the respective groups confirming or denying such a history is shown in figure 19 [76]. The number of recruits with no contact was almost 400,000, while that of recruits with a history of contact was over 10,000. These data give a very solid basis upon which to draw conclusions. The difference between the two groups is striking. In particular, the distribution of those with no known history of contact illustrates the impossibility of separating those who were infected (despite a negative history) from those who were not.

As part of the program, almost 300,000 white male recruits who had been residents of a single county (the United States has over 3,000 counties) during their lifetime were also tested with tuberculin PPD-B. This antigen complex (antigen complexes obtained from environmental mycobacteria are nowadays denoted as “sensitins” [84-88]) had been prepared from a *M. intracellulare* strain first isolated at the Battey Hospital (thus PPD-B) in Rome, Georgia, United States. The distribution of reaction sizes shows striking geographic differences in the frequency of
Figure 17. Frequency distribution of tuberculin skin test results by age, in Minnesota and North Carolina, United States. Figure reproduced with the permission of the European Respiratory Society Journals Ltd. from [82].

Figure 18. Frequency distribution of tuberculin skin test results (5 TU PPD-S) in 5,544 tuberculosis patients, United States. Figure reproduced with the permission of the American Thoracic Society / American Lung Association from [76].
recruits reacting to the sensitin obtained from that environmental *Mycobacterium* (figure 20) [76].

The same diameter of tuberculin skin test induration can indicate a different probability of tuberculous infection, depending on the situation. This is exemplified in the model developed by Rust and Thomas using the United States Navy recruit data [89]. The probability of tuberculous infection is greatly increased if there is a history of contact with tuberculosis as compared with those with no such history (figure 21). While a 10 mm induration size, for example, indicates a probability of over 70 per cent for tuberculous infection in a person with a history of contact with a tuberculosis patient, the same induration indicates a probability of only 10 per cent in a person without such a history. This study reveals the limitations of the predictive value of a positive test result, if the test specificity and the prevalence of the condition sought are low. Applying the test to persons with a history of contact raises the expected prevalence of tuberculous infection in comparison to a population without such a history: it increases the predictive value of a positive test (its ability when positive to correctly separate the infected from the non-infected).

The specificity of the tuberculin skin test is not predictable in different areas of the world, as there are large variations in the presence of environmental mycobacteria,
Figure 20. Frequency of reactions to PPD-B in white United States Navy recruits who were lifetime residents of a single county, by county of origin. Figure reproduced with the permission of the American Thoracic Society / American Lung Association from [76].

Figure 21. Probability of infection with *M. tuberculosis* as a function of diameter of reaction size to tuberculin and history of contact. Data from [89].
which induce cross-reactions. In Djibouti, for example, cross-reactions to tuberculin PPD are virtually absent (figure 22, data kindly provided by Dr. Arnaud Trébucq, IUATLD). Thus, in this situation, the error is minimal, even if all reactions of a size above 0 mm are considered to be indicative of infection with *M. tuberculosis*. In contrast, data from the large tuberculin skin test survey in Tanzania show that the mode of reactions is below 10 mm of induration, and that the expected mode resulting from tuberculous infection is hidden [90]. An intermediate picture emerges from a survey conducted in Korea in 1975 [91]. Among males aged 10 to 14 years, a clear dichotomy of reaction sizes can be observed (figure 23) [91]. The fraction of reactors with indurations of 5 mm to 9 mm is fairly small, and the misclassification, considering as infected those with 10 mm induration or greater, is likely to be very small. The ability to distinguish infected from non-infected persons is largely dependent on the ratio of the prevalence of infection with environmental mycobacteria, usually causing smaller reactions, to the prevalence of infection with *M. tuberculosis*, usually causing larger reactions. This ratio may change over time, making the distinction between the infected and non-infected more or less difficult. This is exemplified in two surveys

![Figure 22. Distribution of tuberculin skin test reaction sizes to 2 TU PPD RT23 compared in surveys conducted in Djibouti and Tanzania. Persons with a reaction size of 0 mm were excluded for easier display. Data for Tanzania from [90]. Data for Djibouti courtesy :Trébucq A, IUATLD.](image-url)
Figure 23. Distribution of tuberculin skin test reaction sizes to 1 TU PPD RT23 in Korea, 1975. Persons with a reaction size of 0 mm were excluded for easier display. Data from [91].

Figure 24. Distribution of tuberculin skin test reaction sizes to 1 TU PPD RT23 in Korea, results of two surveys from 1965 and 1995. Data from [93, 94].
from Korea, one carried out in 1965, and the other in 1995 (figure 24) [92-94]. In 1965, differentiating those who were infected with *M. tuberculosis* from those who were not was straightforward. By 1995, when the prevalence of infection with *M. tuberculosis* had markedly decreased, the distinction had become much more difficult, as the gray zone at that time contributed a much larger fraction of all reactors.

Vaccination with Bacille Calmette-Guérin (BCG) induces tuberculin sensitivity to various degrees [95]. The largest trial on BCG vaccination ever conducted was carried out in Chingleput, India [96]. This trial also provided follow-up information on tuberculin sensitivity following vaccination. An example of the distribution of reaction sizes to tuberculin 2.5 months and 4 years following vaccination in children aged less than 5 years is shown in figure 25. Shortly following vaccination, the distribution is indistinguishable from that caused by natural infection with *M. tuberculosis*. Within a few years, the mode shifted to the left, but a considerable proportion of children still showed large reactions. The influence of BCG vaccination on the results of tuberculin skin testing is clearly related to the time elapsed since

**Figure 25.** Distribution of tuberculin skin test reaction sizes following BCG vaccination 2.5 months and 4 years after vaccination, Chingleput, India, and 4 to 5 years after vaccination in Puerto Rico. Figure reproduced with the permission of the Indian Council of Medical Research from [96] and with the permission of the American Public Health Association from [97].
vaccination. In Puerto Rico, a tuberculin skin test showed considerably less waning 4 to 5 years after vaccination (figure 25) [97]. The age of the children in this case was between 1 and 18 years at intake. In Chingleput, BCG strains 1331 (“Paris” strain) and 1173 P (“Copenhagen” strain) were both used, and in Puerto Rico the Birkhaug-Albany strain was used. It is documented that post-vaccination reactivity to tuberculin varies not only with the strain [95], but also with age at vaccination and time elapsed since vaccination [96].

**Approaches to determining the prevalence of tuberculous infection in the presence of cross-reactions attributable to environmental mycobacteria**

The problem of determining the prevalence of tuberculous infection in the presence of cross-reactions attributable to environmental mycobacteria is summarized in figure 26 [98]. At any chosen cut-off point a certain proportion of tested persons

<table>
<thead>
<tr>
<th>Infection</th>
<th>yes</th>
<th>no</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pos</td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>neg</td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
<tr>
<td>total</td>
<td>a+c</td>
<td>b+d</td>
<td>N</td>
</tr>
</tbody>
</table>

**Figure 26.** Model of the composite distribution of tuberculin skin test reactions reaction sizes in the presence of non-specific reactions. The observed composite distribution (solid histogram) is attributable to at least three distributions: a distribution from those infected with *M. tuberculosis* (solid line), a distribution of those infected with mycobacteria other than *M. tuberculosis* (solid line with a mode at around 2 to 3 mm), and *M. tuberculosis* (dotted line with a mode at around 2 to 3 mm). Zero mm reactors are excluded. Figure reproduced with the permission of Churchill Livingstone from [98].
is misclassified. Moving to the left, i.e., towards a lower cut-off point, increases sensitivity by reducing the proportion with false-negative results among those with the characteristic, while decreasing specificity by increasing the proportion of false-positive results among those without the characteristic. Moving to the right, i.e., towards a higher cut-off point, has the opposite effect. To determine the prevalence of infection is possible only if both the sensitivity and the specificity of the test are known. However, while the sensitivity of the tuberculin test is fairly well characterized, the specificity of the test remains largely unpredictable.

Arbitrary cut-off points to denote the presence or absence of infection, such as 10 mm, 14 mm, or 18 mm of induration or more, might be adequate under certain circumstances, but not at all under others.

In a tuberculin skin test survey on mainland Tanzania in 1957, the frequency of non-specific sensitization varied considerably in different districts in the southeastern part of the country [99]. Nevertheless, it was assumed that non-specific reactions did not exceed 14 mm induration. Under this assumption, the sensitivity at this cut-off (based on areas with little cross-reaction) was 82 per cent, i.e., 18 per cent of infected persons were missed with this cut-off point. For this reason, the prevalence of infection was subsequently determined by dividing the number of those with ≥14 mm reaction sizes by 0.82 to account for the loss of sensitivity. As stated in the paper, this approach was dependent on the assumption that the specificity of the test at this cut-off point was 100 per cent, an assumption not borne out by findings in the much larger surveys later carried out in Tanzania [90].

The sensitivity of the tuberculin skin test is much better defined than its specificity. This is indicated by similar, close to normal distributions of reaction sizes in tuberculosis patients [76, 83] and healthy persons in areas with little interference from environmental mycobacteria (Djibouti survey). In such settings, the mode occurs at a size from 16 mm to 19 mm [81]. This has led to the use of the so-called mirror image technique [100]. This approach assumes that the number of persons with reaction sizes above the mode equals the number of those with reaction sizes below the mode. The number of infected persons is calculated by doubling the number with reaction sizes larger than the mode plus the number of those with a reaction size located at the mode. Again, the assumption is made that the specificity at this diameter is 100 per cent, an assumption that is certainly more justified than with 14 mm. Nevertheless, there are several problems with the approach. First, a large amount of information is lost. Second, a mode can not always be easily identified, as shown in the example of the Tanzania survey. However, the choice of the exact location of the mode is crucial, as the calculated prevalence of infection is very sensitive to the location of the mode [98].
The problems in determining the prevalence of infection are seemingly insurmountable in some settings, while they are minor in others. An appealing method, so-called mixture analysis, has been proposed by statisticians to evaluate the contribution of two finite distributions to an observable composite distribution [101]. Preliminary attempts are being made to apply mixture analysis to tuberculin skin test surveys. An example from the Tanzania survey is shown in figure 27 (Neuenschwander BE, Zwahlen M, Rieder HL, unpublished data). The method is statistically and computationally demanding, and will need more application to evaluate its usefulness in the interpretation of tuberculin skin test surveys that show important contributions from sensitization from environmental mycobacteria and/or BCG vaccination.

**Prevalence of infection**

**Prevalence of infection by age and sex**

Despite the precautions that need to be taken in the interpretation of tuberculin skin test surveys, a wealth of information is available that can be interpreted sensibly. As the predictive value of the tuberculin skin test is much improved when the
prevalence of infection is higher, findings from early surveys in industrialized countries and from surveys in countries with a currently high prevalence of tuberculous infection are particularly meaningful.

Mantoux conducted the first survey giving age-specific prevalence of tuberculous infection using the intradermal technique (with Old Tuberculin) in healthy children in Paris, France [61]. Figure 28 shows that at the age of 5 years, about half the children were already infected with \textit{M. tuberculosis}. Figures 29 and 30 show the age-specific prevalence of infection, stratified by sex, in Denmark [102] and India [103], respectively. In both countries, the prevalence of infection among males and females increases in children at approximately the same increments, but increases more rapidly in males than in females after childhood. The resulting difference in the prevalence of infection among adult males and females is noticeable in Denmark and rather large in India. This is best demonstrated in the male-to-female ratio of age-specific prevalence of infection in the two surveys (figure 31) [102, 103].

The age-specific prevalence of infection in industrialized countries markedly decreased during this century, as exemplified in Switzerland (figure 32) [104-107]. In 1920, virtually everybody was infected with \textit{M. tuberculosis} by the age of 20 years. By 1978, the proportion infected at 20 years had dropped to 10 per cent. Similarly, in a middle-income country, Korea, the age-specific prevalence of persons

![Figure 28. Age-specific prevalence of tuberculous infection in children, Paris, France, 1910. Data from [61].](image-url)
Figure 29. Age- and sex-specific prevalence of tuberculosis infection in Denmark, 1950–1952. Data from [102].

Figure 30. Age- and sex-specific prevalence of tuberculosis infection in India, 1961–1962. Data from [103].
with reaction sizes of 10 mm or more decreased rapidly from 1965 to 1995 (figure 33) [91, 93, 94, 108, 109]).

**Prevalence of infection and population density**

As population density varies in rural and urban settings, opportunity of exposure and thus risk of infection is expected to differ. It is thus not surprising that infection prevalence differs in rural areas from that observed in urban settings. A higher prevalence of tuberculous infection in urban than in rural areas was shown in Korea (figure 34) [93], and was even more pronounced in Zanzibar (figure 35) [110].

**Tuberculous infection and socioeconomic indicators**

The prevalence of infection varies not only by sex and age in the community, but also by socioeconomic stratum of the population. Studies indicating this were first carried out in the early part of the twentieth century. D’Arcy Hart summarized three of these studies done in Trondheim in Norway, Heidelberg in Germany, and Hué in Vietnam (figure 36) [111]. The differences in age-specific prevalence of tuberculous infection between the wealthier and the poorer segments of the population are substantial. It is noteworthy that these differences are actually larger than the differences between countries.

![Figure 31. Age-specific male-to-female ratio in prevalence of tuberculous infection in Denmark and India. Data from [102] and [103].](image)
Figure 32. Age-specific prevalence of tuberculous infection in Switzerland from 1920–1990. Data from [104-107].

Figure 33. Age-specific prevalence of tuberculous infection, Korea, 1965–1995. Data from [91, 93, 94, 108, 109].
**Figure 34.** Age-specific prevalence of tuberculous infection (≥10 mm induration of tuberculin skin test) in urban and rural areas of Korea, 1965. Data from [93].

**Figure 35.** Age-specific prevalence of tuberculous infection (only reactors with ≥16 mm induration of tuberculin skin test included), in urban and rural areas of Zanzibar, 1961. Data from [110].
Kuemmerer and Comstock examined the variation in prevalence of infection among high-school students in Washington County, United States, according to the education level of parents and housing characteristics (figure 37) [112]. Crowding was a particularly important predictor of large tuberculin skin test reactions.

**Predictive epidemiology: modeling risk of infection**

The incidence of infection with *M. tuberculosis* is virtually impossible to measure. It requires large numbers of individuals to be tested repeatedly, and both boosting by repeat testing of the same individual [113] (leading to successively larger reactions) and reversion [114-116] (leading to smaller or negative reactions) make the test difficult to interpret [117]. In addition, there are problems associated with the operating characteristics of the test. For these reasons, preference is usually given to measuring the prevalence of infection at a given age, and then algebraically deriving the average annual probability (average annual risk) that would produce this prevalence. The calculation of the average annual risk is fairly straightforward. The calculation is based on the average annual probability of escaping infection. The derivation of risk of infection from prevalence of infection is summarized in table 2. For example, if the risk of infection is 10 per cent per year, then 90 per cent

![Graph showing age-specific prevalence of tuberculous infection in three countries.](image)

**Figure 36.** Age-specific prevalence of tuberculous infection in three countries. Surveys from 1910 to 1920 in Heidelberg (Germany), Trondheim (Norway), and Hué (Vietnam) by socioeconomic status. Figure reproduced with the permission of the British Medical Research Council from [111].
of children will have escaped infection during the first year of life. A subsequent 90 per cent of 90 per cent, i.e., 81 per cent, will remain free of infection at the end of the second year, and 90 per cent of 81 per cent, i.e., 72.9 per cent, will remain uninfected by their third birthday, etc. This is the principle behind deriving risk of infection from prevalence of infection, elaborated in more detail below.

**Deriving risk of infection from prevalence of infection**

Assuming that the prevalence of infection with *M. tuberculosis* has been satisfactorily estimated, the essence of the analysis of a tuberculin skin test survey is the estimation of the average annual risk of infection. The annual risk of infection refers to a risk at a specified calendar time \( b+x \), where \( b \) indicates the calendar time at which the cohort in the survey was born and \( x \) is a number between 0 and \( a \), where \( a \) is the age of the cohort at calendar time \( b+a \), the time when the survey was conducted. It cannot be known at exactly what calendar time this risk existed without inferences from serial surveys [118, 119]. Because the risk may change over calendar time, \( x \) has been approximated to lie at the midpoint between the year the cohort was born and the year the survey was conducted, if data from a single survey only are available [119, 120]:

\[
R_{b+a/2} = 1 - (1 - P_{b+a})^{1/a}
\]

where \( R_{b+a/2} \) denotes the annual risk of infection at the midpoint in calendar time between the year the cohort was born and the year of the survey, and \( P_{b+a} \) the
prevalence of infection at the time of the survey, where both risk and prevalence are expressed as fractions. Thus, if the prevalence of infection among 10.5-year-old children is found to be 10.0 per cent at the midpoint of the survey (assumed to be at the end of June 1997, for example, i.e., 1997.5,) then the risk of infection is:

$$R_{1992.25} = 1 - (1 - 0.1)^{1/10.5} = 0.010$$

i.e., 1 per cent at the approximate calendar time of the end of March 1992.

Even sequential surveys a few years apart will not necessarily provide information on the change in infection risk. If an earlier survey conducted in 1988, for example, had also provided an estimated average annual risk of infection of 1.0 per cent (approximated at calendar time 1983), the risk of infection has not necessarily remained unchanged over calendar time up to the time of the second survey in 1997. It may well be that the risk decreased in the first years after the birth of the second

**Table 2. Derivation of the average annual risk of infection.**

<table>
<thead>
<tr>
<th>Notation:</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>Calendar year cohort was born</td>
</tr>
<tr>
<td>a</td>
<td>Age of cohort at calendar time survey was conducted</td>
</tr>
<tr>
<td>$P_{a+b}$</td>
<td>Prevalence of infection in cohort at time of survey</td>
</tr>
<tr>
<td>R</td>
<td>Average annual risk of infection between calendar time $b$ and calendar time $b+a$</td>
</tr>
</tbody>
</table>

NB: All calculations must be based on the probability of escaping infection!

at birth: $1 - P_b = 1$
at age = 1: $1 - P_{b+1} = 1 (1 - R)_1$
at age = 2: $1 - P_{b+2} = 1 (1 - R)_1 (1 - R)_2$
...at age = a: $1 - P_{b+a} = 1 (1 - R)_1 (1 - R)_2 \ldots (1 - R)_a$

$1 - P_{b+a} = (1 - R)^a$

$(1 - P_{b+a})^{1/a} = (1 - R)$

$R = 1 - (1 - P_{b+a})^{1/a}$
cohort, for example as a result of implementation of an efficient program for identifying and curing infectious cases. But the number of infectious cases may subsequently have begun to increase because of increasing prevalence of HIV infection among persons with tuberculous infection, leading to an increasing number of transmitters and thus an increased risk of infection in the community. The net effect of an initial decline, followed by an increase in risk of infection, may be that a similar prevalence of infection is measured in the two surveys. Only serial estimates will allow the determination of a trend. Because tuberculin surveys are by their very nature not capable of identifying short-term changes in the risk of infection in a community, spacing of surveys should be sufficiently large to economize on resources.

**Trends in risk of infection over time**

That the risk of infection at a particular point in time (intercept parameter) is of less importance than its change over time (slope parameter) in determining the dynamics of the epidemic becomes apparent if one looks at the changes that have occurred in Europe. Secular trends in risk of infection were first analyzed for Europe by Styblo, Sutherland, and collaborators of the Tuberculosis Surveillance Research Unit [118, 119, 121].

Styblo *et al.* recognized that the calendar change in risk of infection was of overriding importance in shaping the course of the tuberculosis epidemic. Even if risk of infection is initially very high, if a regular decrease can be achieved the cohort effect will set in with tremendous force. By cohort effect it is meant here that a group of persons born in a given calendar year (birth cohort) will be succeeded by a birth cohort (born in the following calendar year) with a slightly smaller risk of infection, while at the same time a portion of persons from the oldest birth cohorts (who had experienced the highest risk of infection and acquired the largest prevalence of infection in the community) is removed by death. With an average annual decline of 5 per cent in risk of infection, the risk will be reduced to 60 per cent of its starting value within just 10 years, and thus a considerably lower proportion of a newly born generation will become infected. In other words, the effect is comparable to compound interest, but it goes in the other direction if the risk of infection declines. Over time, cohorts with successively decreasing levels of infection will replace older birth cohorts in this scenario. Styblo [122] back-calculated the course of cohort-specific prevalence of infection, based on a long series of derivations of risk of infection from the Netherlands, and was thus able to project the future course of infection using the data from the Netherlands, provided that no major changes occur. It can be shown [122] that a rapid change will set in once the risk of infection begins to decrease, even if modestly so.

The annual risk of infection at the beginning of the twentieth century was so large that it was very unlikely that a person could escape infection by the time of
Figure 38. Secular trends in the annual risk of infection in selected European countries. Data from [107, 118, 121-125].

Figure 39. Recent estimates of the annual risk of infection in different regions and countries. Figure reproduced with the permission of the World Health Organization from [120].
reaching adulthood. In the Netherlands the risk of infection was in the order of magnitude of 10 or more per cent [122], and it was probably very similar in most European countries, as a comparison of analyses from different countries shows (figure 38) [107, 118, 121-125]. The slope of decline in all countries is parallel, and the average annual decline is in the order of magnitude of 10 to 12 per cent. While a similar level may not be reached in the different countries until as much as 15 to 20 years later, this is less important than the rate of the decline. This has, for example, resulted in a reduction in the annual risk of infection in the Netherlands from around 10 per 100 in 1910 to less than 10 per 100,000 population in 1990.

Cauthen and collaborators analyzed the trends in the average annual risk of infection from available tuberculin skin test surveys in several low income countries [120]. A summary of the most recent estimates of the risk of infection shows large differences between regions and countries (figure 39) [120]. Examples of the results of this analysis are summarized in figure 40 for the WHO African Region, in figure 41 for the WHO Eastern Mediterranean Region and in figure 42 for the WHO

![Figure 40](image-url)

**Figure 40.** Trends in the annual risk of infection in the World Health Organization Africa Region. Dark symbols and lines are used to display risk estimates and trends for national populations and large parts of national populations; light symbols and lines are used to display risk estimates for sub-national units; solid lines connect surveys that closely represent the same population; dashed lines connect surveys that represent approximately the same population. Figure reproduced with the permission of the World Health Organization from [120].
Figure 41. Trends in the annual risk of infection in the World Health Organization Eastern Mediterranean Region. Dark symbols and lines are used to display risk estimates and trends for national populations and large parts of national populations; light symbols and lines are used to display risk estimates for sub-national units; solid lines connect surveys that closely represent the same population; dashed lines connect surveys that represent approximately the same population. Figure reproduced with the permission of the World Health Organization from [120].

Figure 42. Trends in the annual risk of infection in the World Health Organization South-East Asia Region. Dark symbols and lines are used to display risk estimates and trends for national populations and large parts of national populations; light symbols and lines are used to display risk estimates for sub-national units; solid lines connect surveys that closely represent the same population; dashed lines connect surveys that represent approximately the same population. Figure reproduced with the permission of the World Health Organization from [120].
South-East Asia Region. In the WHO Africa Region, the annual decline is less than 5 per cent in all countries for which data were available. In the Eastern Mediterranean Region the average annual decline was between 5 and 10 per cent, and in the South-East Asia Region the decline was difficult to ascertain, but the data also suggest an annual decline of less than 5 per cent.

Thus, this analysis shows that the decline in the annual risk of infection varies greatly in different areas of the world. In no country is the risk currently known to be as large as it was in Europe at the beginning of the twentieth century. However, the rate of decline is so slow in some countries that the prospects for an acceleration of the decline in the risk of infection in the near future are small. Furthermore, the rapid increase in the number of infectious cases observed in many countries, particularly in sub-Saharan Africa, is most likely to lead to an increase in infection risk. Because the risk of infection and its change over time will affect the prevalence of infection as people enter into adulthood, the speed of decline in risk of infection has important repercussions on subsequent morbidity in different regions of the world. To achieve this goal, a much more aggressive approach to control than currently applied will be required [126].

**Extrapolating from risk of infection to age-specific prevalence of infection**

The results of surveys of the age-specific prevalence of infection are often not available, but it is important to know what proportion of each segment in the age pyramid of a country’s population is infected. Utilizing the analyses by Cauthen and collaborators and data from Europe, ten Dam back-calculated the expected underlying age-specific infection prevalence for 1990 in different regions of the world (data from Gerard ten Dam 1990, WHO, kindly provided by Mario C Raviglione, WHO, written communication May 17, 1993).

For Western Europe, for example, a very high proportion of those over the age of 60 years remains infected (figure 43). This is because these birth cohorts were born at a time when the risk of infection was very high and because they lived for a sufficiently long time to have only a minimal chance of escaping infection or, conversely, a high risk of having become infected. In contrast, the youngest generation is virtually free of tuberculous infection, because the probability of acquiring infection during one year has become minute. With the current trend in risk of infection, each birth cohort that dies is being replaced by a cohort with a lower level of infection: the hatched area in figure 43 is moving steadily to the right on the abscissa of the graph.

In sub-Saharan Africa the situation is entirely different (figure 44). First, the age distribution of the population is strongly skewed towards the young, with
Figure 43. Estimated age-specific prevalence of tuberculous infection in Western Europe in 1990. Data courtesy: ten Dam HG, 1990.

Figure 44. Estimated age-specific prevalence of tuberculous infection in sub-Saharan Africa in 1990. Data courtesy: ten Dam HG, 1990.
approximately 50 per cent of the population younger than 15 years of age. Half or more of the population aged 15 to 49 years is infected with tubercle bacilli according to this model. Taking into account the slow rate of decline in infection risk, and perhaps even an increase in some of the countries, it is unlikely that this picture will change substantially for the better over the coming decades.

In South-East Asia, the age distribution of the population resembles the distribution in sub-Saharan Africa, but with a slightly lower proportion of the population found in the youngest age groups (figure 45). Here, the proportion infected in the economically productive years of life (15 to 49 years of age) is even larger than in sub-Saharan Africa. Any change in this picture will depend on the trend in risk of infection.

In all of these analyses, the role that HIV may impart on the transmission dynamics of \textit{M. tuberculosis} has not been accounted for. As the discussion on increased morbidity in high HIV incidence and prevalence countries will show, the risk and prevalence of tuberculous infection is likely on the increase in some countries. This will provoke a vicious cycle: an increase in the prevalence of tuberculous infection will lead to ever more incident cases (and thus transmitters), which will lead to a further increase in the risk of infection. The situation is further aggravated if some of the tuberculosis infected are, or are becoming, co-infected with HIV infection.

![Figure 45. Estimated age-specific prevalence of tuberculous infection in South-East Asia / Oceania (excluding Australia and New Zealand) in 1990. Data courtesy: ten Dam HG, 1990.](image-url)
These modeled estimates of the prevalence of infection from risk of infection must nevertheless be interpreted with caution. Estimates of the risk of infection were usually derived from surveys in a single age group, and have thus assumed a constant risk of infection across age groups. That this is not necessarily the case has been demonstrated by Sutherland [12], and more recently by Nagelkerke [127]. If the risk of infection varies considerably by age, extrapolations from a single age group might thus be erroneous.

Overall, available estimates suggest that approximately one third of the world population is currently infected with *M. tuberculosis*. This crude proportion is very similar in different regions of the world, but the age-specific analyses demonstrate important differences. In most industrialized countries (as exemplified with Western Europe), tuberculous infection is concentrated among the elderly, while in Latin America, South-East Asia, and sub-Saharan Africa, the young economically productive and reproductive age groups constitute the most important population segment with tuberculous infection. Furthermore, the rapid decline in risk of infection is likely to accelerate the disappearance of tuberculous infection as a public health problem in industrialized countries, while the slow decline or probable increase in some parts of the world will result in a continued high prevalence of infection in young adults.

**Risk of infection and infectious cases**

It is apparent that the rate of transmission of tubercle bacilli is dependent on the number of sources of infection in a society. The number of successful transmissions from infectious cases to susceptible individuals over a defined period of time (usually one year) at a certain calendar time determines the risk of infection in the community during that period. Styblo estimated that in the pre-chemotherapy era one infectious source infected, on average, about 20 persons during the 2-year period the case remained infectious before death intervened or spontaneous bacteriologic conversion had occurred [122]. If it was further assumed that the lifetime risk for development of infectious tuberculosis following infection was 5 per cent, then two prevalent cases correspond to one incident case per year, each producing 10 successful transmissions per year. Thus, in a community with 100 prevalent infectious cases per 100,000 population, 1,000 transmissions occur in one year, or 1 per cent of the population becomes infected. Because incidence is half the prevalence in this model, 1 per cent risk of infection corresponds to approximately 50 incident infectious cases [128].

It does not, however, necessarily follow that the incidence of tuberculosis can be accurately extrapolated from knowledge of the current risk of infection. Risk of
infection is intrinsically coupled to duration of undiagnosed, untreated, transmissible tuberculosis, and thus with person-time of infectiousness in the community. This person-time of infectiousness may be variable, except in the absence of intervention in a standardized population and in the presence of an epidemiologic equilibrium. The duration of infectiousness, coupled with the incidence of new infectious cases, is reflected in the prevalence of such cases (i.e., prevalence is a function of incidence and duration). As soon as an effective intervention is introduced, the duration of infectiousness is reduced, transmission is decreased and the relation between prevalence and incidence is disturbed.

This is demonstrated in figures 46 and 47. In the first example, the point prevalence is double (24 cases) the incidence (12 cases), and the person-time of infectiousness or period prevalence (the sum of the periods of infectiousness contributed by each case) in year three is 288 person-months. In the second example, the average duration of each case is only 6 months, and the period prevalence in year three is 72 person-months. If the cases in these examples are infectious tuberculosis cases, the first situation might reflect no intervention, while in the second situation

![Figure 46](image.png)

*Figure 46.* Schematic presentation of the relation between incidence, point prevalence, and period prevalence. This figure depicts a situation where the prevalence is double the incidence. The straight horizontal lines symbolize each incident case and its duration of infectiousness. The straight vertical lines symbolize the beginning / end of a calendar year, and the dotted vertical line an arbitrary chosen point of a prevalence survey on March 15.
interventions reduced the period of infectiousness. Although the incidence in both examples is the same, the period of infectiousness (summarized as person-time of infectiousness) in the case of intervention is only a quarter of that in the absence of intervention. Thus, despite the same incidence, the risk of infection to the community is likely to be considerably lower in the second as compared with the first situation. The relation between infectious cases and risk of infection is determined primarily by the duration of infectiousness rather than the incidence of cases.

A note of caution is indicated. The model here is an over-simplification, in that it may suggest that infectiousness is constant over time. This is unlikely to be the case, as infectiousness may increase over time (figure 48) [129] or bacillary excretion might be irregular and intermittent.

In countries with inadequate case management, it can be argued that the number of infectious (sputum smear-positive) patients may remain essentially the same after 2 years, because the principal impact of such an intervention lies with a reduction of case fatality at the expense of keeping infectious cases alive [130].

---

**Figure 47.** Schematic presentation of the relation between incidence, point prevalence, and period prevalence. This figure depicts a situation where the prevalence is half the incidence. The straight horizontal lines symbolize each incident case and its duration of infectiousness. The straight vertical lines symbolize the beginning / end of a calendar year, and the dotted vertical line an arbitrary chosen point of a prevalence survey on March 15.
Furthermore, the risk of tuberculosis following infection with *M. tuberculosis* may vary in different populations. It is, for example, greatly increased in persons with HIV infection compared to immunocompetent hosts (see chapter 3). Thus, the epidemiologic balance usually observed between host and bacillus is not maintained under these circumstances where each case of tuberculosis may produce more than one new infectious case in the HIV-infected segment of the population. Conversely, in countries where intervention effectively cuts the chain of transmission, the number of transmissions caused by one case will be reduced. Thus, to produce a 1 per cent annual risk of infection, a larger number of incident cases is required, because the person-time of infectiousness is reduced. This was shown for the United States before HIV had a noticeable impact on tuberculosis. Here, extrapolation required over 400 incident cases per 100,000 population to result in a 1 per cent risk of infection [131].
Etiologic epidemiology: risk factors for disease given that infection has occurred

The most important risk factor for tuberculosis is infection with tubercle bacilli. Tubercle bacilli are a necessary, but not a sufficient cause of tuberculosis. While the risk of becoming infected is largely exogenous in nature, determined by the characteristics of the source case, environment, and duration of exposure, the risk of developing tuberculosis given that infection has occurred is largely endogenous, determined by the integrity of the cellular immune system.

In most instances it cannot be determined why a particular person does or does not develop tuberculosis after becoming infected with tubercle bacilli. On the other hand, a multitude of factors has been identified which increase the risk of progression from sub-clinical infection with \textit{M. tuberculosis} to overt tuberculosis [132]. A selection of these factors is summarized in figure 49. Some of these may have considerable impact because they are not only strong risk factors but may also be highly prevalent in the general population. Although other factors appear to increase risk several-fold, they are to be regarded more as medical curiosities of minimal public health importance because they occur only rarely. The importance of any risk factor in public health is determined by both the strength of the association and the prevalence of the risk factor in the population.

\textit{Time elapsed since becoming infected}

The temporal association between infection and progression to disease has been recognized for quite some time [133]. This association is most striking for the development of primary tuberculosis (figure 50) and tuberculous meningitis (figure 51). Information on risk of tuberculosis following infection is also available from preventive chemotherapy trials among household contacts, clearly showing that the risk is elevated in the first years following infection, rapidly falls off and then remains
Figure 49. Selected risk factors for tuberculosis given that infection has occurred. Data from [132] and various other sources (see text). The diamond represents the referent (infection that has occurred > 7 years in the past), full circles are relative risks (from population-based studies), open circles are odds ratios (from case-control studies), and dotted lines connecting circles indicate ranges in different studies.

Figure 50. Development of primary tuberculosis after infection. Figure reproduced with the permission of Churchill Livingstone from [133].
low, but measurable for a prolonged period of time (figure 52) [134]. Information from the tuberculin-positive placebo group in BCG vaccination trials also indicated that the incidence of tuberculosis is highest in the first few years after infection, and then rapidly falls off (figure 52) [135]. Of the 140 cases observed in this group in the BCG trial conducted by the British Medical Research Council during a 20-year follow-up, 64 (45.7 per cent) developed within the first 2.5 years, and 95 (67.9 per cent) within

Figure 51. Development of tuberculous meningitis after infection. Figure reproduced with the permission of Churchill Livingstone from [133].

Figure 52. Incidence of tuberculosis among household contacts receiving placebo, compared to household contacts receiving preventive chemotherapy with isoniazid [134] and number of tuberculosis cases among unvaccinated persons with large tuberculin skin test reactions by time interval after beginning of observation in a controlled clinical trial on BCG vaccination in Great Britain [135].
the first 5 years. It must be stressed that it was not known in the latter study at what time infection had actually been acquired before intake into the study.

The quality of the information on the absolute incidence of tuberculosis varies considerably in different studies and at different times. These differences are attributable to differences in the case definition of primary tuberculosis, to differences of time elapsed since becoming infected, to differences in the age of the patients at the time of assessment, and other factors. A commonly used rule of thumb is that the lifetime risk of a newly infected young child might be 10 per cent [136], and that half of this risk falls within the first five years following infection. The qualifying words here are “young child” (1 to 3 years in this case), because it is apparent that the shorter the remaining life span, the smaller the cumulative lifetime risk will be. For operational purposes, the term “primary tuberculosis” has been proposed for disease developing within 5 years following infection [122, 137]. While a 5-year cut-off may seem somewhat arbitrary, it approximates the time required for achieving the median cumulative lifetime risk of a child.

The reason for the high risk of tuberculosis following shortly after acquisition of infection is most likely largely attributable to a “filter” effect: those with the least defenses against progression from sub-clinical infection to tuberculosis are the most likely to develop disease shortly after infection. Not developing tuberculosis immediately following infection, by contrast, may indicate superior defense mechanisms.

The period of time that has elapsed since becoming infected is probably one of the most important risk factors for tuberculosis. It is a strong factor, with recent infection being 10 times more likely to produce a case than a long-standing infection. Furthermore, virtually every new case of infectious tuberculosis is accompanied by new infections that may lead to new cases. The contribution of disease resulting from recent infection to overall tuberculosis morbidity in a community will, however, vary with the risk of infection in a community and the pre-existing prevalence of tuberculous infection. A qualification of the definition of recent infection as a risk factor is indicated here. It refers solely to the average annual risk of tuberculosis during each of the 5 years following primary infection defined above cumulatively as primary tuberculosis, as the cumulative risk of tuberculosis from remote infection (lasting for more than 5 years) might become considerable over a sufficiently long period of observation.

**Infection with the human immunodeficiency virus (HIV)**

In a prospective study among injection drug users in a methadone maintenance program in New York City, United States, the incidence of tuberculosis among
Initially tuberculin-positive persons was prospectively assessed [138]. Among 49 tuberculin-positive, HIV-seropositive persons, seven developed tuberculosis over an observation period of approximately 2 years (7.9 per 100 person-years). The risk of development of disease in previously infected persons might be considerably higher, however, because 13 of the 49 had received preventive therapy, and none of the cases appeared in this group. This study has demonstrated that HIV infection is the most powerful factor yet recognized in the progression to disease from pre-existing infection with *M. tuberculosis* [132]. In a retrospective cohort study in Zaire [139], the relative risk of tuberculosis among HIV-seropositive women was 26 compared with seronegative women.

The risk of development of disease in previously infected persons is not equally spread over the course of HIV infection. A study in Florida, United States, conducted before the definition of the acquired immune-deficiency syndrome (AIDS) included tuberculosis [140], and another one in New York City [141], have shed some light on the temporal association between tuberculosis and the occurrence of other opportunistic diseases. In the Florida study [140], tuberculosis occurred more than 1 month before other opportunistic diseases in 50 per cent; in another 30 per cent it occurred within 1 month before or 1 month after, and it was diagnosed more than 1 month after the occurrence of other opportunistic diseases in 20 per cent (figure 53).

![Figure 53. Temporal association between tuberculosis and other AIDS-defining opportunistic diseases in Florida, United States. Figure reproduced with the permission of the American Medical Association from [140].](image-url)
Similarly, in New York City [141], the occurrence of tuberculosis was temporarily skewed towards precedence before other opportunistic diseases by a median of 2 months. Tuberculosis started to occur with increasing frequency about 3 years before onset of other opportunistic diseases, and increased in frequency with diminishing interval to form a bell-shaped curve around the diagnosis of other opportunistic diseases. This is biologically plausible, because HIV infection tends to induce increasing immunosuppression with increasing duration of infection, thus increasing the likelihood of offsetting the precarious balance between the tubercle bacillus and the host's cellular immune system.

The risk of tuberculosis among HIV-infected persons is closely correlated with the number of CD4$^+$ lymphocytes, as shown in a prospective study among HIV-infected persons with tuberculin skin test reactions of 5 mm or more in Italy (figure 54) [142]. De Cock has reviewed the relation between degree of immunosuppression and manifestation of tuberculosis [143]. If tuberculosis in an HIV-infected individual develops relatively early on in the course of HIV infection, pulmonary tuberculosis is a common manifestation. With progression of immunosuppression, lymphatic tuberculosis and serous forms ensue (pleural, peritoneal, and pericardial tuberculosis), followed by tuberculous meningitis. With very low CD4 cell counts, disseminated tuberculosis is a common feature (figure 55). This is

![Figure 54](image_url)

**Figure 54.** Risk of tuberculosis (cases per 100 person-years) among HIV infected persons, by CD4$^+$ lymphocyte count. Data from [142].
consistent with studies in animal models conducted by Lurie [144]. Rabbits inbred for susceptibility to bovine tuberculosis showed that the more susceptible animals were considerably less likely to form cavities, but more likely to develop lymphatic and disseminated tuberculosis. Similarly, children who develop primary progressive tuberculosis are more likely to develop lymphatic and disseminated tuberculosis [145, 146].

The annual risk of progression from latent, sub-clinical tuberculous infection to disease is in the order of 5 to 15 per cent if HIV infection becomes superimposed [138, 139, 147, 148]. Because reactions of between 5 and 9 mm are less likely to be attributable to tuberculous infection, the risk might be larger in dually infected persons. This was shown in another prospective study from Italy (figure 56) [149]. Thus, an assessment of the tuberculosis risk among HIV-infected persons depends first on the certainty that there is co-existing tuberculous infection and then on the level of immunosuppression. The risk is likely to be much larger if HIV infection is followed by tuberculous infection: in that case, the host might be almost completely defenseless against the invading organism. Thus, risk differences in individuals might be hugely different, as shown in figure 57.

![Figure 55](image.png)

**Figure 55.** Clinical and immunopathological course of human immunodeficiency virus-associated tuberculosis. Figure reproduced with the permission of the American Medical Association from [143].
Figure 56. Risk of tuberculosis among HIV-infected persons by size of tuberculin skin test reaction. Data from [149].

Figure 57. Example of risk of tuberculosis following tuberculous infection in individuals with different risk factors.
Spontaneously healed tuberculosis with fibrotic residuals

Tuberculosis may spontaneously remit without treatment. However, if not treated, the disease is likely to recur in a large proportion of cases. Persons who had tuberculosis which healed spontaneously leaving fibrotic residuals (“fibrotic lesions” is another commonly used term) are at increased risk of developing tuberculosis again [150-152]. The 5-year follow-up in the trial on preventive therapy carried out by the International Union Against Tuberculosis provided data for estimating the incidence of tuberculosis among persons with fibrotic, stable lesions suggestive of old, healed tuberculosis [152]. From the 5-year follow-up in the untreated placebo group it can be estimated that the natural incidence was 2.9 per 1,000 person-years. This is similar to that obtained by Falk and Fuchs [151], who estimated an incidence of two cases per 1,000 person-years. However, the incidence varied considerably (2.6, 4.0, and 13.6 per 1,000 person-years) in three studies reviewed by Edwards et al. [150].

The size of the radiographic lesion resulting from healed tuberculosis is an important determinant of the risk for reactivation [152]. Persons with fibrotic lesions of 2 cm or more in diameter were twice as likely to develop tuberculosis compared with patients with smaller sized lesions.

Age

There are large differences in tuberculosis incidence by age. Theoretically, these disparities may be attributed to differences in risk or prevalence of infection, differences in disease risk once infected, or both. The trend of commonly observed higher incidence of disease with increasing age can be partly explained by the cumulatively increasing prevalence of tuberculous infection. Adolescents and young adults appear to be especially prone to progression from latent infection to disease (figure 58) [136], while children around the age of 10 years appear to be least prone. Variation with stage of maturity is not as likely an explanation of the steady increase in incidence rates among adults up to the age of 60 years. However, there are indications that the risk of tuberculosis following infection increases beyond the age of 60 years [153].

A study of the occurrence of different forms of extrapulmonary tuberculosis in the United States among patients with tuberculosis [145] opens anew questions of differences in susceptibility to the manifestation of tuberculosis among different age, sex, and race/ethnic groups. The relation of tuberculosis with age at one extreme showed the decreasing likelihood of lymphatic tuberculosis with increasing age and at the other the increasing likelihood of genitourinary tuberculosis with increasing age. Lymphatic tuberculosis, particularly intrathoracic, is commonly perceived as...
an expression of increased susceptibility because of immunosuppression [141, 154] or period of time elapsed since infection coupled with maturational factors. Conversely, genitourinary tuberculosis is believed to be most often the result of recrudescence of long-standing latent tuberculous foci [155]. Pleural tuberculosis is virtually absent in small children. Pleural tuberculosis is generally perceived as a delayed hypersensitivity reaction to antigens of *M. tuberculosis* [156], often leading to sequestration of antigen-reactive T-lymphocytes into the pleural space [157]. It is conceivable that the virtual absence of tuberculous pleurisy in very young children might be attributable to a lower sensitivity to tuberculin at that age.

**Genetic factors**

A re-analysis of the data from the Prophit survey, adjusting for various potential confounding factors, has shown that concordance for tuberculosis risk was significantly higher among monozygotic than dizygotic twin pairs [158]. This finding indicates that inherited susceptibility is an important risk factor for tuberculosis.

There are numerous discernable genetic factors that might affect the risk of tuberculosis, some of which are summarized here.

**Sex**

There appears to be a difference between males and females in tuberculosis incidence rates following infection. Among tuberculin-positive participants in the BCG vaccination trial in Puerto Rico, incidence was 18 per cent higher among females than among males [136]. In the Danish national survey [102], the risk of tuberculosis among infected females was also greater than among infected males from the age of 15 to 44 years, but was lower among females than among males beyond 44 years of age.

In a study of extrapulmonary tuberculosis in Cambodian refugees in Thailand, lymphatic tuberculosis was equally frequent among male and female children, but was much more common among adult females than among adult males (figure 59) [159], suggesting that underlying genetic and maturational factors may impact on the expression of tuberculosis.

A striking difference in the likelihood of developing extrapulmonary tuberculosis was found between male and female patients with tuberculosis in the study on extrapulmonary tuberculosis in the United States [145]. Among all forms of extrapulmonary tuberculosis, only pleural tuberculosis was equally likely to develop in male and female tuberculosis patients (adjusted for age, race/ethnicity, and country of origin). All other extrapulmonary forms (lymphatic, osteoarticular, peritoneal, pericardial, meningeal, and rarer forms) were consistently more likely to develop in female than
Figure 58. Incidence of tuberculosis among initial reactors to tuberculin by age when tuberculosis was first diagnosed. Figure reproduced with the permission of the American Journal of Epidemiology from [136].

Figure 59. Tuberculosis of peripheral lymph nodes by age and sex, Cambodian refugees, Thailand, 1981-1984. Data from [159].
in male patients. The preponderance of lymphatic tuberculosis among female patients has been noted in a large, detailed review [160]. The investigator noted that female preponderance was relatively greatest in the 20- to 24-year-old group, and absolutely greatest in the 10- to 14-year-old group. These data suggest that maturational and hormonal factors may play a role in the risk of tuberculosis and its manifestations.

**Body build**

There is considerable evidence that tuberculosis incidence is strongly associated with body build [161]. Data from studies in United States Navy recruits [162, 163] and from the BCG trial in Georgia/Alabama, United States [164], suggest that the incidence of tuberculosis among persons below ideal body weight is 2.2 to four times greater than among persons with normal weight for height. In these studies adjustments had been made for tuberculin status. In Norway, Tverdal [165] studied the association between body mass index and tuberculosis incidence among 1.2 million persons older than 14 years of age who were followed for periods of 8 to 19 years. With increased body mass index, the incidence of pulmonary (but not extrapulmonary) tuberculosis decreased by a roughly linear trend on the logarithmic scale. The age-adjusted incidence in the 17 per cent with the lowest body mass index was 3.1 times greater than that in the 12 per cent with the highest body mass index, and 2.5 times greater than in the group at the body mass index median (containing 22 per cent of persons followed). The incidence of infectious tuberculosis by body mass index among 30- to 49-year-old persons, stratified by sex, is shown in figure 60 [165].

**HLA types**

Among the genetic traits that have been implicated as tuberculosis risk factors, the role of human leukocyte antigen (HLA) types has been studied extensively. Among HLA types, A11-B15 and DR2 have been shown in some studies to increase the odds of tuberculosis 1.5- to 3.5-fold [166-168]. An association between HLA type and occurrence of tuberculosis is not consistently found, however. Hawkins *et al.* [169] found no statistically significant association in a large study among Hong Kong Chinese (256 pulmonary cases and 100 healthy control subjects). In addition, Hawkins *et al.* found no association by haplotype segregation analysis among 93 offspring and 38 parents in 21 multicase families.

**Blood groups**

In a study by Overfield and Klauber [170] of Eskimo tuberculosis patients, the disease was significantly more likely among those with blood group AB or B than among those with blood group O or A.
Hemophilia

In an outbreak of tuberculosis in a pediatric hospital, children with hemophilia were more likely than children with other disorders to develop the disease [171, 172]. Of the 16 children with hemophilia who were exposed, six developed tuberculosis, compared with three of 75 with other disorders. If the two boys in the study with HIV infection are excluded, the relative risk is 7.1. The incidence observed in exposed hemophiliacs was similar to that in children receiving cytotoxic chemotherapy for leukemia and solid tumors (10 of 21 exposed).

Virgin populations

There are indications that the risk of disease following infection might be considerably higher in populations which have not previously been in contact with tuberculosis [173, 174].

Other genetic factors

That genetic factors play a role in susceptibility to progression from infection to tuberculosis is evident. That some ill-defined factors may affect the clinical expression of tuberculosis has also been noted. An anecdote to that effect is the report of

![Figure 60. Incidence of infectious tuberculosis by body mass index and sex among adults, Norway. Data from [165].](image)

- 75 -
a family particularly prone to developing a rare form of skin tuberculosis, *lupus vulgaris*, while other forms were virtually absent (figure 61) [175]. Natural resistance to progression to disease following infection with tubercle bacilli has been postulated to be influenced by the gene for the natural-resistance-associated macrophage protein 1 (*nramp1*) as was shown in a case-control study of tuberculosis in the Gambia, West Africa [176]. However, the results of the study have been challenged because of the failure to adjust either for pre-existing tuberculous infection or for the probably ubiquitous presence of environmental mycobacteria in the country where the study was conducted [177]. Animal experiments suggest further that *nramp1*-determined defenses are subordinate to other defenses against tuberculosis [178].

While it has been argued that black Americans might be more susceptible to tuberculosis than white Americans given that tuberculous infection has occurred [179, 180], this has been challenged [181, 182]. Differences in disease frequency observed among different races appear mainly attributable to differences in incidence and prevalence of infection. Follow-up of large numbers of participants in a BCG vaccination trial in Puerto Rico revealed no major differences in the incidence of tuberculosis between black and white tuberculin reactors [97, 136]. In a follow-up study of tuberculin-positive United States Navy recruits, incidence rates for blacks and whites were also similar. However, a much higher incidence was observed in

---

**Figure 61.** Excess frequency of *lupus vulgaris* in a family. The filled symbols indicate death from tuberculosis, the double vertical line unspecified decendency, and circles indicate unknown sex. Figure reproduced with the permission of Springer-Verlag GmbH from [175].
Asians (nearly all of whom were from the Philippines) than in blacks or whites [183]. Black/white disparities in the incidence of disease following infection in the United States are relatively minor. Inequities between races in the long-term decline of tuberculosis incidence observed in adult age groups appear to reflect disparities between blacks and whites in the rate of decline in the risk of tuberculous infection.

However, in the study on extrapulmonary tuberculosis in the United States considerable differences were observed between the race/ethnic groups with regard to the likelihood of extrapulmonary tuberculosis at a particular site [145]. For instance, black patients with tuberculosis were only half as likely to develop genitourinary tuberculosis as non-Hispanic white patients, and twice as likely to develop miliary tuberculosis. American Indians with tuberculosis had the same likelihood of developing lymphatic tuberculosis as non-Hispanic white patients, but were 11 times more likely to develop peritoneal tuberculosis. These differences remain largely unexplained, but suggest that there may be underlying genetic or environmental factors that determine the site of disease.

**Environmental factors**

**Substance abuse**

**Smoking**

Two studies in England examined the relationship between cigarette smoking and risk of tuberculosis [184, 185]. From these case-control studies, there is evidence that the odds of tuberculosis increase with an increasing number of cigarettes smoked (figure 62) [185].

In a study in Shanghai, tuberculosis incidence was shown to be higher among smokers than among non-smokers [186]. The effect remained after adjustment for age, sex, type of work, history of contact, and area of housing.

**Alcohol abuse**

Clinicians often point out an association between alcohol consumption and tuberculosis incidence. Nevertheless, epidemiologic evidence of a causal association is inconclusive. The postulated association is confounded with environments (in industrialized countries at least) which are conducive for increased transmission and thus infection with *M. tuberculosis*. Nevertheless, because immune mechanisms affected by alcohol are also those which are essential for resistance to tuberculosis [187], excessive alcohol consumption might indeed increase the risk of tuberculosis.
Injecting drug use

Reichman and colleagues have postulated (2 years before the discovery of AIDS) a higher risk of tuberculosis among injecting drug users following tuberculous infection compared to non-users [188]. The reasoning was based on the finding that the proportion of tuberculin reactors was not as different among methadone-maintenance program enrollees compared to the general population as the respective disease incidence.

Nutrition

Malnutrition

It is a common notion that malnutrition adversely affects the immune system. In Germany, tuberculosis mortality increased rapidly during the First World War, declined subsequently, and increased again during the period of monetary inflation with serious food shortages in the years 1922–1923 [189]. This second peak has been attributed to malnutrition-associated tuberculosis. In the Netherlands, tuberculosis mortality increased with the beginning of the Second World War, but the increase was accentuated when a German food embargo led to a sharp drop in available calories [189]. Switzerland, spared from the Second World War, nevertheless had to introduce
food rationing; increases in tuberculosis mortality were smaller than in Germany or
the Netherlands, but nevertheless noticeable [189].

Diet

Vegetarian diet has been identified as a risk factor for tuberculosis [190]. In this
case-control study, an increasing trend in risk was found with decreasing meat or fish
consumption. The highest odds were found for lactovegetarians. In this context, it
has been recommended that vitamin D supplementation might be particularly impor-
tant in persons with reduced exposure to sunlight [191]. The active metabolite of
vitamin D, 1,25-hydroxy-vitamin D₃, promotes maturation and activation of human
monocytes and macrophages, and its inhibitory activity on multiplication of virulent
tubercle bacilli in human macrophages has been demonstrated [192]. It has been
hypothesized that some of the characteristics of tuberculosis among immigrants to
the United Kingdom (particularly the frequency of non-respiratory tuberculosis)
might be partially related to vitamin D deficiency [193]. However, the presence of
a large number of confounding effects precludes a firm conclusion.

Medical conditions

Silicosi

That tuberculosis is common among miners and patients with silicosi has long
been recognized [194]. Paul has estimated that the incidence of tuberculosis among
miners with silicosi was 26 times greater than among miners without silicosi [195].
Westerholm et al. published the results of a study of tuberculosis from the National
Swedish Pneumoconiosis register which found the odds of tuberculosis among
silicosi patients to be 30 times that of appropriate controls [196]. In a long-term
follow-up study of gold miners in South Africa, the relative risk for tuberculosis
was 2.8 for men with silicosi compared with that in men without silicosi [197].
Only about 55 per cent of those with the most severe form of silicosi remained
tuberculosis free after 7 years of follow-up, a 6.3 per cent average annual risk
(figure 63) [197]. A detailed study on risk factors among South African gold
miners has also shown the importance of the precise type of occupation of the
miners [198]: drilling, for example, was associated with a more than two-fold higher
risk of tuberculosis than working in low dust occupations within the mines.

Diabetes mellitus

The notion that diabetes mellitus and tuberculosis are associated is so old and
prevalent that it has long been taken for granted as more than a coincidence. There
are, however, only a few large surveys that permit a quantitative assessment of the incidence of tuberculosis among diabetics in comparison with appropriate controls. In a large survey in Philadelphia in the mid-1940s, the incidence was 8.4 per cent in diabetics, compared with 4.3 per cent in non-diabetic industrial employees [199]. In Sweden, Silwer and Oscarsson [200, 201] found pulmonary tuberculosis in 3.6 per cent of their diabetic patients, compared with 0.88 per cent in a general population control group. After adjustment for age and sex, the relative risk was estimated to be 3.6. Opsahl et al. found the incidence of tuberculosis among diabetics to be more than three times greater than in the general population [202]. These studies suffer in that they fail to adjust for tuberculous infection. It can thus not be excluded that tuberculous infection was more frequent among diabetics than non-diabetics. However, as the major studies were population-based, this potential bias might not be so severe as to invalidate the findings. Furthermore, it should be noted that diabetes is positively correlated, while tuberculosis risk is negatively correlated with increasing body mass index [162, 165, 203], thus leading to a potential underestimation of the risk of tuberculosis among diabetics without adjustment for body weight.

Malignancies

Malignant lymphomas are recognized as being associated with high tuberculosis morbidity [204]. Tuberculosis was reported to be common among patients with lung cancer, and (what was then termed) lymphosarcoma, and reticulum cell sarcoma [205]. Feld et al. investigated the occurrence of tuberculosis in a series of patients with malignant disease and found a strikingly high frequency of tuberculosis specifically in patients with carcinoma of the head and neck [206]. The occurrence of 10 tuberculosis cases among 1,366 (seven cases per 1,000) patients with squamous cell carcinoma of the head and neck would suggest a relative risk of the order of 16, compared with an annual incidence of 45 per 100,000 in the general population of similar age (45 to 65 years). Tuberculosis cases were ascertained from the time of primary cancer diagnosis, but the length of follow-up was not determined. However, tuberculosis occurred much less frequently among patients with other malignancies, which provides a degree of assurance through specificity of effect that the association is not the spurious result of bias associated with hospital-based studies.

Renal failure

Patients with end-stage renal failure and those on hemodialysis have consistently been shown to be at increased risk of tuberculosis [207-212]. The incidence among
such patients is estimated to be 10 to 15 times greater than that of the general population.

**Measles**

Measles is recognized to reduce tuberculin skin test reactivity [213, 214]. It is a commonly held belief that measles may also increase the risk of tuberculosis. Nevertheless, a critical review of available data suggests that there is little evidence to support the hypothesis that it does so [215].

**Gastrectomy**

Gastrectomy has been associated with high tuberculosis morbidity [216]. Although many studies claim a higher incidence among gastrectomized patients than in the general population, a major deficiency has been the lack of an appropriate control group. Thorn et al. partly overcame this deficiency in a study that found the incidence of tuberculosis among male gastrectomy patients to be five times greater than among men of the same age from the same area [217]. Whether gastrectomy itself is a risk factor, or whether it is secondarily associated with another risk factor

![Figure 63. Proportion with silicosis remaining free of tuberculosis among South African gold miners, by severity of silicosis. Figure reproduced with the permission of the American Thoracic Society / American Lung Association from [197].](image-url)
such as departure from ideal body weight, remains unclear, however. In the study by Thorn et al., for example, it was noted that gastrectomy patients weighing less than 85 per cent of ideal body weight were 14 times more likely to develop tuberculosis than those whose weight was normal for height [217].

**Jejunoileal bypass**

The limited available data suggest that jejunoileal bypass surgery for obesity might be a risk factor for tuberculosis [218]. Although a strikingly high prevalence and incidence have been reported among such patients [219, 220], the overall number of reported cases has been small.

**Corticosteroid treatment**

The relative importance of treatment with steroids as a risk factor for tuberculosis has been a controversial question. Lurie demonstrated in several experiments that glucocorticoids in large doses reduced resistance to a challenge with tubercle bacilli in rabbit models [221]. In a review of reports of tuberculosis complicating corticosteroid therapy, Horne concluded that the risk was small [222]. Similarly, Haanaes and Bergmann concluded from a large survey of tuberculosis patients in Norway that the risk of reactivation during prolonged therapy with corticosteroids at 10 mg or less per day, or with higher doses for short periods, must be low [223]. Schatz et al. found no tuberculosis among their asthmatic patients who were treated with systemic steroids [224]. Smyllie and Connolly prospectively studied 555 hospitalized patients who received corticosteroids compared with 499 who did not [225]. Assessment after follow-up of 1.5 to 7 years revealed only one tuberculosis case in the corticosteroid group and two cases in the control group. However, in the light of strong laboratory evidence [221, 226] and individual case reports that high doses of steroids may adversely affect the outcome of tuberculosis [227, 228], it is generally perceived as a risk factor requiring preventive intervention [229].

**Pregnancy**

The role of pregnancy in the development of tuberculosis has long been debated [230]. In his review, Snider came to the conclusion that there is no solid evidence that pregnancy has an adverse effect on tuberculosis [230]. However, there are indications that the post-partum period might increase the risk of progression to tuberculosis [231]. Out of 65 cases of tuberculosis occurring between the beginning of pregnancy and 6 months post-partum, 37 developed during the post-partum period, suggesting that the risk of progression to tuberculosis in the post-partum period might be double that during pregnancy.
Factors associated with the etiologic agent

Infecting dose effect

Experimental studies by Ratcliffe and Palladino in small mammals have shown that almost all tubercle bacilli inhaled as single organisms reached the alveolus and produced a tubercle [232]. It would thus appear that no dose effect would be expected in the risk of tuberculosis given that infection has occurred, i.e., disease might not depend on the number of inhaled tubercle bacilli. Nevertheless, a study in British Columbia and Saskatchewan, Canada, by Grzybowski et al. appears to show that the risk of disease given infection is larger if infection was caused by a sputum smear-positive than if it was caused by a sputum smear-negative source case [30]. This effect still holds after adjustment for race and sex (figure 64) [30, 132]. However, these data should be interpreted with caution. It seems that in this study, a tuberculin skin test was considered to be positive if the diameter of induration exceeded 5 mm [233]. At this diameter the specificity of the test was unlikely to have been 100 per cent.

Figure 64. Percentage of secondary cases with bacteriologically confirmed tuberculosis among infected contacts by type of contact and bacteriologic status of source case, British Columbia and Saskatchewan, Canada, 1966 – 1971; sm+ and sm- indicates sputum smear-positive and sputum smear-negative respectively; cul+ and cul- indicates sputum culture-positive and sputum culture-negative respectively. Adjusted for race and sex, calculated from [30] and published in [132].
Because the prevalence of tuberculous infection is expected to be higher in close than in casual contacts, the predictive value of 6 mm and more induration is lower in the latter than in the former group. Thus, the risk of tuberculosis among reactors would be expected to be lower in the latter. However, because cross-reactions due to environmental mycobacteria are likely to be of minor importance in British Columbia and Saskatchewan, if the Navy Recruit study findings in the neighboring States (Washington, Idaho, Montana) in the United States are taken as indicative [76], it remains unresolved whether non-specificity of the test can explain the entire difference.

**Strain virulence**

Different strains of tubercle bacilli exhibit different virulence in experimental animal models [234-239]. Recently, stronger evidence has emerged that strains that are more virulent in animal models might also be more virulent in humans [240], and it appears that certain virulent strains may have a particular propensity to cause tuberculous meningitis [241].

In guinea pigs, highly isoniazid-resistant, catalase-negative organisms have been shown to be of low virulence [242]. However, in a more recent experiment in mice, drug-resistant strains of tubercle bacilli exhibited a range of virulence [239]. Intravenous infection in mice with mycobacterial catalase-peroxidase protein \((katG)\)-deleted, isoniazid-resistant *M. tuberculosis* compared with mice infected with tubercle bacilli in whom \(katG\) was restored has shown that the \(katG\) restored catalase-peroxidase activity and enhanced tissue persistence [243]. These results indicate that \(katG\) contributes to the ability of *M. tuberculosis* to grow and survive within infected host tissue. Loss of the \(katG\) gene in another experiment led to loss of virulence of a *M. bovis* strain, but subsequent re-integration of a functional \(katG\) gene restored full virulence in guinea pigs [244].

In New York City, the implementation of effective control measures led to a 19 per cent decrease in tuberculosis cases from the period 1991–1992 to 1993–1994 [245]. However, in the same period the number of multidrug-resistant cases (resistant to at least isoniazid and rifampicin) fell by 44 per cent. This disparity continued into 1994 [246]. This is unexpected, as one would think that multidrug-resistant mutants would have a selective advantage to be transmitted (through prolonged transmission due to treatment failure) and thus cause proportionately more cases of secondary disease. Patients with multidrug-resistant strains will remain infectious for a longer time on average than will patients with fully susceptible organisms, as chemotherapy is likely to be less efficient in rapidly reducing transmissibility.
It seems to be confirmed that certain genetic mutations in tubercle bacilli which cause them to become isoniazid-resistant also reduce their virulence in experimental animals. However, because several different mutations may independently cause isoniazid resistance [247], it can not be concluded that all isoniazid-resistant strains are less virulent than fully susceptible strains.

**Infection with M. bovis**

As shown in figure 11, the association between tuberculous infection in humans and the prevalence of tuberculous herds is strong. However, the number of cases per 1,000 reactors decreased with increasing frequency of infected herds (figure 65) [50]. This indicates that the risk of disease given that infection with *M. bovis* has occurred is considerably lower than that given that infection with *M. tuberculosis* has occurred. Using mathematical modeling, Magnus [50] estimated differences in disease frequency following bovine and human strain infection (figure 66). The risk of tuberculosis following infection with *M. tuberculosis* is seven to 12 times greater than that following infection with *M. bovis*. These data thus demonstrate that bovine tuberculosis is a powerful determinant for the prevalence of infection, but has much less impact on morbidity subsequent to infection than infection with *M. tuberculosis* (and Koch was not all that wrong in judging that *M. bovis* was of lesser importance to humans than *M. tuberculosis*).

![Figure 65. Correlation between the prevalence of tuberculous herds and the risk of tuberculosis among human tuberculin reactors, Denmark. Figure reproduced with the permission of the World Health Organization from [50].](image)
Re-infection

One of the first comparative studies of BCG vaccination versus no vaccination was carried out in Norway in 1927 [248]. It was noted during the study that student nurses starting their work with a positive tuberculin skin test reaction had a much lower risk of developing tuberculosis than those with an initially negative reaction. From these findings it might be concluded that a pre-existing infection provided some protection against re-infection compared to acquisition of new infection. This is, however, not necessarily so, as those entering with infection pre-existing for an unknown duration are precisely those who had escaped disease, most likely because of better immunity. Thus, if infection occurs, those with pre-existing infection will have, on average, better defense mechanisms than those who have not been previously infected. Because nurses who were initially uninfected and were then vaccinated had a similarly low tuberculosis risk during follow-up as those with pre-existing infection, and much lower risk than that of uninfected student nurses who remained unvaccinated, one might conclude that a pre-existing infection does indeed provide some protection against re-infection. This is borne out by immunologic reasoning that in persons with pre-existing infection the cellular immune system is more rapidly prepared to respond to bacillary multiplication and dissemination than the immunologically unprepared system.

Protection against re-infection is, however, incomplete. There is plenty of anecdotal evidence bearing on this point. Disease due to re-infection has been

![Figure 66. Estimated morbidity rates for 20- to 39-year-old tuberculin reactors infected from human (and bovine) sources and bovine sources only. Data from [50].](image)
noted to occur in cases with a strain which was susceptible and from which they were cured. Some of these cured patients developed disease again with a resistant strain [249-253], a strain with a different phage type [254], or a different Restriction Fragment Length Polymorphism (RFLP) pattern [255-257] between a first and a second episode of tuberculosis. Vynnycky and colleagues have assessed the protection afforded by a primary infection against re-infection by means of mathematical modeling [258]. Using data from notifications in England and Wales, and taking age into consideration in their analyses, they suggest that the protection provided by pre-existing infection against re-infection for adults is approximately 40 per cent. This relatively low level of protection (compared to higher levels of protection observed in some BCG trials) might be attributable to the fact that tubercle bacilli are entirely removed by the host in a considerable proportion of cases [259]. With increasing time between complete removal of tubercle bacilli acquired during the first infection and the time of acquisition of the second, the immunologic memory may wane in a similar fashion to that imputed from the observations from long-term follow-up in BCG trials, where the protective effect decreased over time.

**Descriptive epidemiology: morbidity**

Tuberculosis incidence varies considerably in different populations and population segments. Most of these differences are likely to be attributable to differences in the underlying prevalence of infection with *M. tuberculosis*. Nevertheless, some of the differences might also be attributable to differences in the risk of developing tuberculosis given that infection has occurred. Very often, however, the underlying reasons for an increased incidence are difficult to disentangle. Nevertheless, precise descriptive epidemiology can be helpful in identifying groups and population segments that need particular attention for targeting scarce resources, irrespective of the etiologic reason for increased incidence.

**Secular trends in industrialized countries**

During this century, tuberculosis has rapidly decreased in industrialized countries. In the United States, a uniform reporting system was introduced in 1953, and since then tuberculosis has regularly decreased, with few exceptions, until 1985 (figure 67) [260-263]. In 1980, a large influx of Indochinese refugees only temporarily halted the steady decline in tuberculosis morbidity [264]. In 1985, the average 5.7 per cent annual decline was halted [265], and reversed from 1986 onwards [266].
After reaching a low in 1985, reported tuberculosis cases increased until 1992, and decreased subsequently through 1997 [263, 267]. This has been attributed, in part, to the spread of HIV, homelessness, increased drug abuse, and immigration from countries where tuberculosis is common; deterioration in the living conditions of and health care delivery to the poor have also been proposed as factors contributing to the increase [268, 269]. The increase in tuberculosis has been particularly conspicuous in New York City, where it reached its trough in 1978, and then increased steadily [269]. It is apparent that infection with HIV can not entirely explain this increase, which started quite some time before HIV was expected to have a noticeable impact.

It may be noted, however, that the only factor that has changed for the better since tuberculosis case rates started to decline again in the United States after 1992 is adequate funding for tuberculosis control. This seems to provide strong evidence that the resurgence of tuberculosis, although helped by HIV, homelessness, substance abuse, immigration, and other factors, was really only due to the lack of capacity to handle these problems during the 1980s.

Similarly to the United States, the crude number of tuberculosis cases in 14 European countries combined decreased regularly from 1974 through 1991, with an average annual decline of 5.4 per cent (figure 68) [270]. In almost all of the 14 countries evaluated in this study, tuberculosis declined over the period of observation (figure 69). However, the rate of decline differed in individual countries. Generally, in countries with higher initial rates, tuberculosis declined more rapidly than in those where the rates were initially lower. Figure 69 also shows the irregularity of reporting, particularly in two countries, highlighting the need to be careful in the interpretation of notification data. While the range of notifications by country was wider at the beginning of the observation period (7.8 to 76.3 per 100,000 population in 1974, an almost ten-fold difference), it was considerably narrower by the end of the period (6.6 to 19.7 per 100,000 population in 1990, a three-fold difference). In a number of these European countries, tuberculosis notification rates have not appreciably declined in recent years, and in several an increase has been observed.

Härö, in a comprehensive analysis of tuberculosis data from Finland, sheds some light on the development of tuberculosis over time [9, 271]. One indicator of progress in the control of the tuberculosis epidemic is the change in the age structure of diseased persons. Härö calculated the age-adjusted median age of tuberculosis cases over time (figure 3) [9, 271]. Until about 1975, female cases always tended to be younger than male cases, but subsequently the median age was lower among male than among female patients. A shift in the mean age or, perhaps a better measurement, the median age, towards older age groups points at a lessening of the problem in the society [272]. It indicates that transmission of tubercle bacilli is
**Figure 67.** Reported tuberculosis cases in the United States, 1953–1997. Data from [260] [261] [262, 263].

**Figure 68.** Trend in tuberculosis notifications in 14 Western European countries combined, 1974–1991. Figure reproduced with the permission of the World Health Organization from [270]
Figure 69. Trend in tuberculosis notifications in 14 Western European countries, 1974–1991. Data from [270].

Figure 70. Age-specific tuberculosis notification rates among males in Finland. Cross-sectional observations from 1954 to 1989. Data from [271].
decreasing and that an increasing proportion of cases emanates from the pool of persons infected many years in the past. The shift implies that cohorts with less and less infection are successively replacing cohorts born at the time when risk of infection was much higher.

Cross-sectional information is useful, but tends to hide the underlying epidemiology in successive cohorts. In 1930, Andvord proposed to study tuberculosis within generations, i.e., its development within birth cohorts [273-275]. It was a publication by Frost [276], however, which brought birth cohort analysis to the attention of epidemiologists [277].

The usual reporting systems for tuberculosis provide age-specific morbidity figures for the year the cases were reported (cross-sectional reports). Cases occurring in different age groups can also be viewed as cases occurring among people born in different calendar years (i.e., people belonging to different birth cohorts). If cross-sectional surveys are matching available age groups, i.e., if cross-sectional surveys are available for example every 10 years, and during each survey the age is known for 10-year age groups, then morbidity experience can be examined within each birth cohort. For example, surveys are available for 1975, 1985, and 1995, and people who became ill in these years were known to be aged on average 5 years (0 to 9 years of age), 15 years, 25 years, etc. In this example, persons with an average age of 25 years in 1995 belong to the birth cohort of 1970. In the 1985 survey, the 1970 birth cohort was 15 years old, and in the 1975 survey it was 5 years old. In addition to information about age effects, this analysis now also gives information about birth cohort effects: it allows the examination of morbidity (in this example, but it might be equally applied to infection or mortality) within each successive birth cohort. If sufficiently detailed information is available, a third component can be examined. In addition to the effects of age and birth cohort on morbidity, certain periods, such as times of war, might entail a particularly high risk of morbidity.

Härö analyzed respiratory tuberculosis morbidity cross-sectionally and among birth cohorts [9, 271]. In cross-sectional observations, age-specific notification rates increased, particularly in later years, to peak in the oldest age groups, as exemplified among males in figure 70. Because age-specific notification rates and notification years could be synchronized (by 5-year groupings), it was possible to analyze age-specific notification rates also within birth cohorts (figure 71). In each successive birth cohort, respiratory tuberculosis peaked in young adults. This phenomenon can be found when tuberculosis is in decline: the cross-sectionally observed peak among the elderly is merely a residual of the much higher morbidity experienced by the
cohort when it was young. A similar picture was found in Upstate New York, United States (figure 72) [278].

**Geographic differences in disease frequency**

International comparison of tuberculosis incidence is fraught with problems, because of differences in case definition [279, 280], notification criteria [281] and notification discipline [282]. One indicator for comparison for the relative differences in the magnitude of the problem can be obtained from countries with extensive immigration, such as Canada and the United States. As notifiable tuberculosis cases in these countries follow a uniform case definition, rates among different groups coming from various parts of the world can be compared. Enarson examined case notification rates among various population groups in Canada and found large differences between Canadian-born and foreign-born patients (figure 73) [283]. The risk ratio between the highest (Philippines) and the lowest risk group (patients immigrating from Belgium, the Netherlands, and Luxembourg) was about 30-fold.

**Figure 71.** Age-specific tuberculosis notification rates among males in Finland. Cross-sectional observation (dotted lines with circles) in 1954 and 1989 and analysis by birth cohorts (solid lines) born from 1892 through 1972. Data from [271].
Figure 72. Age-specific tuberculosis notification rates among females in Upstate New York, United States. Cross-sectional observation (dotted lines with circles) in 1940–1944 and 1970–1974 and analysis by birth cohorts (solid lines) born from 1900 through 1950. Figure reproduced with the permission of the American Journal of Epidemiology from [278].

Figure 73. Relative risk of tuberculosis notification among foreign-born compared to Canadian-born tuberculosis patients. Data from [283].
McKenna reported a similar analysis from the United States, showing a very similar uneven distribution in tuberculosis case notification rates by country of origin [284] (figure 74).

Despite the enormous difficulties, Murray et al. [285], Sudre et al. [286, 287], Raviglione et al. [288], and Dolin et al. [289, 290] have all made efforts to estimate the incidence of tuberculosis expected in the world. Given the often poor quality of tuberculosis notification data, estimates are not easy to make. The available detailed estimates by Dolin et al. are presented here (figure 75) [289]. The data indicate that the largest annual number of cases emanates from South-East Asia, accounting for almost half of the total cases in the world. The incidence rate, however, is estimated to be highest in Africa, and lowest in industrialized countries, with a ten-fold difference in the incidence between the extremes.

The estimates made by Sudre [286, 287] and Dolin [289, 290] suggest the occurrence of approximately 7.5 to 8 million new tuberculosis cases per year for the early 1990s.

**Differences in disease frequency in the community**

**Age**

The most conspicuous differences in disease frequency in the community are observed when comparing different age groups. Furthermore, the age-specific incidence of tuberculosis varies greatly over time within countries and concurrently

![Figure 74](image-url). Adjusted incidence rate ratios of tuberculosis among the foreign-born in the United States. Data from [284].
Figure 75. Estimated incidence of tuberculosis in the world in 1995. Data from [289].

Figure 76. Age-specific notification rates of infectious tuberculosis among females, Norway 1946 and 1972. Figure reproduced with the permission of the European Respiratory Society Journals Ltd. from [291].
between countries. As shown previously, the median age of tuberculosis patients has increased markedly in countries where the risk of infection declined rapidly and thus the infected population segments became increasingly older [271]. The example by Bjartveit from Norway, comparing age-specific notification rates of infectious tuberculosis among females, very clearly shows the large risk in young adults in 1946 which had completely disappeared by 1972 (figure 76) [291]. In 1972, tuberculosis notification rates had become very low compared to 1946, and the highest rates were now found in the oldest segments of the population.

In contrast, in many low income countries, tuberculosis notification rates currently still peak in young adults, as shown in the example of Senegal (figure 77) [292].

Sex

In virtually all countries, notification rates among males are higher than among females [293]. Disease notification is the result of multiple steps in a process which incorporates the following components: prevalence of infection, risk of disease given that infection has occurred, access of the patient to diagnosis, and notification discipline in the health care system. It has been demonstrated earlier (figures 29 – 31) that the age-specific prevalence of infection in adulthood is often greater among males than females, but to different degrees in various situations. Risk of progression from infection to disease also differs among females and males, and varies depending on age. The observation of differences is further complicated by gender issues, i.e., unequal access to diagnostic services for males and females with tuberculosis, which makes it difficult to disentangle the underlying epidemiologic differences. It is easier to determine the epidemiologic component in countries where males and females have presumably equal accessibility to health care. Thus, the two-to-one male-to-female ratio in tuberculosis incidence rates observed in many industrialized countries, such as the United States [294], appears best explained by the generally higher prevalence of infection observed among males than among females [295].

The male-to-female rate ratio may also differ over time as the age structure of tuberculosis changes, and may differ within sub-populations within the same country. This is exemplified by observations from Denmark [102]. In Copenhagen, males had consistently higher case rates from the early 1920s through the late 1960s, while in the rest of Denmark the ratio inverted over time (figures 78 and 79). Outside Copenhagen, males had consistently lower rates than females until the mid-1940s. Only in subsequent periods did the risk become similar, and actually became higher for males towards the end of the observation period. It is conceivable that this phenomenon is attributable to a combination of the fairly similar (but slightly lower)
Figure 77. Age-specific notification rates of sputum smear-positive tuberculosis among males and females, Senegal, 1997. Numbers above the notification rates among males indicate the age-specific relative risks of males compared to females. Figure reproduced from [292].

Figure 78. Incidence of pulmonary tuberculosis in Denmark, in Copenhagen and the rest of Denmark, among males (lines without symbols) and females (lines with circles), 1921–1957. Figure reproduced with the permission of the World Health Organization from [102].
prevalence of tuberculous infection among females compared to males (figure 29), but a higher risk of progression to disease among young females in the earlier part of the century. With an improving epidemiological situation, the age of tuberculosis patients increased to the point in age where the risk of progression from infection to disease inverted among the sexes and became larger for men than for women.

Socioeconomic status

Poverty has been strongly associated with the incidence of tuberculosis [296, 297]. Although these analyses are based on ecological studies with census tracts [296] or postal codes [297] as units for analysis, the differences are of such a magnitude that the strength of association lends credibility to a causal relation. A study from Upstate New York, United States, exemplifies the similar findings of various studies (figure 80) [278]. This phenomenon is at least partially attributable to the differences in the underlying prevalence of tuberculous infection, as shown by Kuemmerer and Comstock [112]. Clearly, low socioeconomic indicators tend to result in crowded living conditions, conditions that are conducive to increased transmission of tubercle bacilli should a case occur, resulting thus in a generally higher prevalence of tuberculous infection with subsequent increased incidence of disease.

Poverty may also reduce access to health care services [298], thus prolonging the period of infectiousness of tuberculosis patients, and further increasing the risk of infection among the contacts of such a patient.

Race and Ethnicity

Differences in risk among race/ethnic minorities have been studied in particular detail in the United States. Since this information has been recorded, they were noted to have a higher incidence of tuberculosis than the white non-Hispanic (majority) population [266]. The disparity between the white and non-white populations in that country has increased over the years, even long before HIV infection started to have an impact on the epidemiology of tuberculosis (figure 81). The decline of tuberculosis incidence in the United States has been greatest among the 5- to 14-year-old age group, greater among females than males in most age groups, and more pronounced among white than other adults (figure 82) [266]. Nevertheless, this information must be interpreted with caution, not least because of the increasing confounding over time of the non-white group with rapidly increasing Asian immigration.

Children whose disease is notified represent only the tip of the iceberg of all children who have disease, since many of them have few symptoms and go unnoticed. Nevertheless, the asymptomatic children form a pool from which future cases of tuberculosis will emerge. The increase in tuberculosis in young adult age groups in
Figure 79. Male-to-female tuberculosis rate ratio in Denmark (excluding Copenhagen), 1921–1957. Data from [102].

Figure 80. Incidence rates of tuberculosis by socioeconomic level, Upstate New York, United States, 1973. Figure reproduced with the permission of the American Journal of Epidemiology from [278].
minorities in the United States has a potentially serious impact on the future epidemiology in the country. Tuberculosis in children always indicates recent transmission of *M. tuberculosis*, and thus indicates a failure of the public health system to sufficiently contain the epidemic [299]. In 1992, it was reported that the number of tuberculosis cases among children in the United States increased by one third from 1987 to 1990 [300]. An analysis by race shows, however, that case rates among whites (including Hispanic whites) and blacks decreased in a similar fashion until about 1989. From then onwards rates increased sharply among whites and remained unchanged in blacks until 1992 (figure 83) [260, 261]. This could indicate that excess transmission during the recent re-emergence of tuberculosis in the United States did not affect minority children more strongly than white children.

**Migration**

The high incidence among South-East Asian refugees and other immigrants in the United States [264, 301] appears largely explicable by the high risk of tuberculous infection in their countries of origin [120], resulting in a large pool of infected persons from which cases arise.

Migration from high to low incidence countries continues to increase globally. This has an increasing impact on tuberculosis morbidity in several industrialized

![Figure 81](image-url)

**Figure 81.** Nonwhite-to-white tuberculosis notification case rate ratios, United States, 1953–1987. Figure reproduced with the permission of the American Medical Association from [266].
Figure 82. Average annual percentage decline in tuberculosis notification rates, by age, sex, and race, United States, 1963–1984. Figure reproduced with the permission of the American Medical Association from [266].

Figure 83. Tuberculosis case rates among white and black children, United States, 1981–1992. Data from [260, 261].
countries, as shown in the example from Sweden (figure 84, data courtesy of Victoria Romanus, Swedish Institute of Infectious Disease Control, written communication, February 4, 1997) [282]. While the number of reported cases among the Swedish-born population continuously decreased from 1984 to 1993, with only a minor increase in 1995, the number of tuberculosis cases among the foreign-born increased rapidly in the same observation period. As a result, the total number of cases has remained stable in the most recent past. Similar observations have been made in other countries in Europe [302]. From the European tuberculosis surveillance project, a total of 13 countries in the WHO Europe Region reported more than 100 cases of tuberculosis and provided information on country of birth of their tuberculosis patients [282]. Figure 85 shows that the proportion of foreign-born among all cases in these countries ranged from 0 per cent in the Czech Republic to 83 per cent in Israel in 1996.

As tuberculosis retreats in the indigenous population, cases among immigrants become both absolutely and relatively more pronounced. The epidemiology of tuberculosis is frequently quite different among foreigners compared with the indigenous population (figure 86) [303]. While the incidence of tuberculosis in the indigenous population is highest among the elderly, tuberculosis case rates among foreigners commonly peak in young age groups, reflecting the situation in their country of origin.

![Figure 84. Notified cases of tuberculosis in Sweden by country of birth, 1984–1995. Data courtesy: Romanus V, Swedish Institute of Infectious Disease Control and data from [282].](image)
**Figure 85.** Proportion of cases reported as foreign-born from 13 countries of the World Health Organization Europe region, 1996. Data from [282].

**Figure 86.** Reported tuberculosis cases among Swiss and foreign patients, by 5-year age groups, Switzerland, 1990. Reproduced with the permission of the Swiss Federal Office of Public Health from [303].
There is limited evidence from the Netherlands that tuberculosis among the foreign-born population measurably affects the epidemiology among the indigenous population [304]. The effect is, however, relatively small, and does not seem to affect the overall trend of tuberculosis morbidity once the balance has been tipped against the disease as it has in the indigenous population of most industrialized countries. One study in Montreal has shown that the prevalence of infection with *M. tuberculosis* among Canadian-born school children did not differ between areas with high and low proportions of foreign immigrants [305]. Furthermore, in the United States, the country with one of the most heterogeneous populations in the world, the epidemiology of tuberculosis has been very different in the majority population and the various racial/ethnic minorities over a very long time [266]. This suggests that a high tuberculosis incidence in one population segment does not necessarily affect that in another. That tuberculosis is not easily transmissible is evident from numerous studies among close and other contacts. They have consistently shown that those living in closest proximity are at highest risk of acquiring infection from an undiscovered infectious source [29-31]. Refugees, asylum seekers, and foreign workers are commonly set apart from the indigenous population at work and in their personal life. Prevailing circumstances often force them to live under crowded conditions, in environments that are conducive to transmission of tubercle bacilli. Because of multiple barriers, such people may have poorer access to health care facilities, and diagnosis of infectious tuberculosis may be unduly delayed.

Among foreign-born persons residing for prolonged periods in a low incidence country, the risk of tuberculosis seems to diminish over time, as was demonstrated in an elegant study from the United Kingdom [306]. The proportion of all cases among Asians in London who never returned to Asia over a 20-year period was highest in the time immediately following entry to the United Kingdom (figure 87). Those who re-visited Asia showed a stable distribution of cases over the entire period, yet, if analyzed according to time of re-entry into the United Kingdom, the distribution was similar to that of those who never returned to Asia. This analysis shows the importance of switching from a low to a high risk of infection environment (and vice versa), and that recently acquired infection during a visit to the country of origin played a major role.

While tuberculosis risk is highest immediately following immigration from a high to a low incidence country, the risk usually remains higher than that of the indigenous population for a prolonged time [284, 306, 307]. This is quite obviously attributable to the higher prevalence of pre-existing tuberculous infection in the immigrating population, prone to leading to reactivation.
Population density

As the age-specific prevalence of infection is frequently lower in rural than in urban areas, the incidence of tuberculosis is accordingly also expected to be lower. This was indeed shown in two studies from Denmark [308, 309]. The first study compared provincial towns with rural districts in 1950–1952 and found the rates consistently lower across all the age groups examined (figure 88) [308]. In the second study, conducted from 1960–1968, comparing the capital Copenhagen with rural districts showed that the differences had remained large (figure 89) [309].

Marital status

A study in Copenhagen has shown important differences in disease frequency by marital status (figure 90) [309, 310]. The highest incidence of pulmonary tuberculosis among men was found in those who were divorced, and the lowest in married men. Single and widowed men had an incidence between the two extremes. These findings might be attributable to different social behavior of men depending on their marital status.

Figure 87. Tuberculosis notifications among Asians in London who never visited Asia again after entry to the United Kingdom, and those who visited Asia again, by interval between entry and re-entry respectively. Data from [306].
marital status, resulting in lower or higher risk of infection and subsequent disease frequency. Nevertheless, it is conceivable that bereavement (widowers) may also adversely affect the immunologic functions, thus leading to a higher disease breakdown among those infected.

**Substance abuse**

Despite the long-standing notion of an association between alcohol consumption or other substance abuse and tuberculosis incidence, epidemiologic evidence of a causal association is not convincing. The proportion of patients with alcoholism in Canada was strongly associated with incidence of tuberculosis, but was most likely confounded by socioeconomic status [296]: the lower the socioeconomic group, the larger the proportion of alcoholics, both inversely associated with incidence of tuberculosis. Screening of substance abusers (of alcohol and other drugs) has repeatedly shown the excess risk of tuberculosis in these populations compared with the general population [311-314].

**Other risk groups**

Health care workers are an important population segment with an increased risk of tuberculosis [315, 316]. However, a large study in the United States showed that health care workers had a risk of disease no greater than expected based on the demographic composition of persons working in these occupations [317]. Because
Figure 89. Incidence of pulmonary tuberculosis among married men in Copenhagen compared to rural districts, Denmark, 1960–1968. Data from [309].

Figure 90. Incidence of pulmonary tuberculosis by marital status among men in Copenhagen, Denmark, 1960–1968. Data from [309].
health care workers will, on average, be more frequently exposed to tuberculosis patients than persons in most other occupations, an excess risk of tuberculosis in this group might be most readily identified in settings where tuberculosis remains a prevalent condition among patients attending health care services. Conversely, excess risk might be difficult to recognize in areas where tuberculosis has become a rare condition in the general population.

**Impact of HIV infection**

HIV may alter the epidemiology of tuberculosis in three different ways [318]:

- Endogenous reactivation of pre-existing infection with *M. tuberculosis* in persons who become infected with HIV.
- Progression from infection with *M. tuberculosis* to tuberculosis in persons with pre-existing HIV infection.
- Transmission of tubercle bacilli to the general population from tuberculosis patients who developed tuberculosis because of HIV infection.

A retrospective study in Florida [140] indirectly and a prospective study in New York City [141] directly demonstrated that the proportion of tuberculosis cases that developed by endogenous reactivation was of over-riding importance. That primary progression to tuberculosis and transmission of tubercle bacilli to susceptible individuals can become of major importance has been particularly well documented in numerous investigations of outbreaks of tuberculosis in health care facilities [319-326].

An outbreak in a hospital in Italy demonstrated all three mechanisms by which HIV may impact on tuberculosis (figure 91) [319]. An unrecognized case of tuberculosis in an HIV-infected patient, presumably the result of endogenous reactivation of latent, pre-existing infection with *M. tuberculosis*, led to secondary cases among HIV-infected patients in the same ward. This was most likely the result of direct and very rapid progression to tuberculosis from a tuberculous infection superimposed on HIV infection. Finally, there was transmission resulting in disease in a non-HIV-infected staff member.

From a public health perspective it is of particular concern if increased transmission occurs, because this will lead to an increase in the pool of infected persons from which future cases will arise. In one study, the trend of tuberculosis in children under 5 years of age born in the United States was taken as indicative of recent transmission [300]. Using this method, the risk of infection in the United States was
estimated to have increased by more than 10 per cent in the late 1980s and early 1990s. The increase in risk of infection in the general population is likely to be less, because children from ethnic/racial minorities are disproportionately represented among tuberculosis patients in the United States [146].

The impact of HIV infection on tuberculosis in children is expected throughout the world, but has been particularly well documented in sub-Saharan Africa [327-331]. While children are not an epidemiologically important source of transmission in the community [332], they are a sentinel for the course of the tuberculosis epidemic [299], and surveillance data suggest that excess transmission of tubercle bacilli might be occurring on a considerable scale in many countries particularly affected by the HIV pandemic.

**United States and Western Europe**

At the first International Conference on AIDS in 1985 in Atlanta, Georgia, United States, epidemiologists from New York City reported the association of the geographic distribution of increasing tuberculosis incidence and AIDS in the city [333]. A half year later, the United States CDC reported that ongoing surveillance suggested that tuberculosis cases during the first 39 weeks of 1985 failed to decline as was expected from previous years [265]. Subsequently, the CDC, in collaboration with State and local health authorities, initiated several studies, including in Florida [140, 154] and in New York City [141], to better determine the relation between AIDS and tuberculosis. It should be noted that at that time, the surveillance definition

![Figure 91. Secondary cases following an index case of tuberculosis in a ward for HIV-infected patients in Verona, Italy. The filled circle symbolizes the source case, hollow circles indicate the secondary cases. Data from [319].](image-url)
did not yet include tuberculosis as an AIDS-defining condition. This collaboration aimed to gain better knowledge about demographic and clinical characteristics of tuberculosis patients with and without AIDS. These descriptive, retrospective studies confirmed what had earlier been suggested in smaller studies [334]. The following points were noted [335]. Firstly, tuberculosis was common among patients with AIDS in populations with a recognized high prevalence of infection with tubercle bacilli. Secondly, tuberculosis often preceded conditions constituting the then-valid surveillance definition of AIDS. Thirdly, in patients with AIDS, clinical forms of tuberculosis often deviated from the accustomed manifestations of the disease in non-compromised hosts.

In the retrospective study in Florida, cases known to have developed AIDS were matched against the tuberculosis register in the State to determine the proportion of cases who had both tuberculosis and AIDS (figure 92) [140]. The differences were large and the frequency ranged from 2 per cent among non-Hispanic whites to 27 per cent among Haitians. The reason for this difference is clearly to be found in differences in the underlying prevalence of infection with \textit{M. tuberculosis}. Because tubercle bacilli are a necessary cause of the development of HIV-associated tuberculosis, little tuberculosis will be found in a population segment with a low prevalence of tuberculous infection. The frequency of HIV-associated tuberculosis thus critically depends on the size of the segment in a population with tuberculous infection, the prevalence of HIV infection, and the extent to which these two population segments overlap (figure 93). In the Florida study [140], not all Haitians with AIDS had been previously infected with tubercle bacilli, and not all of those who developed tuberculosis amongst them did so in the State of Florida and were thus amenable to case ascertainment. It may thus be assumed that the lifetime risk of a dually infected person to develop tuberculosis during the course of HIV infection is likely to exceed 30 per cent.

From the age-specific prevalence of infection (figures 43–45) it is apparent that HIV infection, which preferentially affects people in the 15- to 49-year-old age groups, will have quite a different impact on tuberculosis in sub-Saharan Africa than in the United States or Western Europe.

In the United States, it appears from data among blacks in New York City that HIV infection seemed to have a major impact on ethnic minorities, but much less on the white, non-Hispanic population [266]. Not all of this increase can be attributed to HIV infection, and other factors must be considered as well [268, 269, 336]. There is little doubt, however, that in the United States, the HIV epidemic has led to a deterioration of the tuberculosis epidemic in certain minority populations. Because
Figure 92. Tuberculosis among AIDS patients, by race/ethnicity, Florida, United States, 1981–1986. Data from [140].

Figure 93. Determinants for the frequency of HIV-associated tuberculosis in a community.
tuberculosis in these groups has continued to affect persons in their reproductive age to a large extent, transmission to their offspring is more likely than among the white population, where tuberculosis has long since become a disease of the elderly.

There is currently little evidence that HIV has seriously affected tuberculosis epidemiology in most of Western Europe [270]. Nevertheless, the differences between European countries are large, as is evidenced by the frequency with which extrapulmonary tuberculosis was diagnosed as an indicator disease for AIDS (figure 94) [337]. The highest proportions of AIDS cases with extrapulmonary tuberculosis were found in Spain and Portugal. This is not surprising, as these two countries reported the highest rates of tuberculosis among all Western European countries in 1996 [282]. Almost 30 per cent of AIDS patients in Spain had extrapulmonary tuberculosis as an indicator disease between 1988 and 1993 [314]. The crude incidence rate of tuberculosis in Catalonia (north-eastern Spain) increased by 50 per cent from 1987 to 1993, and 60 per cent of this increase was attributed directly to AIDS [338]. In other European countries, increases in tuberculosis notifications attributable to HIV infection have been much less conspicuous. Nevertheless, certain cities such as Paris, France [339], and Amsterdam, the Netherlands [340], have noted a pronounced effect. In contrast, the impact of HIV infection on increased notifications in England and Wales has been much less clear-cut [341], and its overall contribution to morbidity was judged to be modest in 1991 [342]. In a study from Edinburgh, Scotland, it was found that only five of over 1,000 cumulative HIV-infected patients had tuberculosis [343].

Sub-Saharan Africa

The tuberculosis epidemic in a large number of sub-Saharan African countries has been seriously affected by the HIV epidemic [143, 288, 344, 345]. It has been estimated that approximately one third, ranging from 0 to 72 per cent, of tuberculosis cases between 1985 and 1993 would not have occurred in sub-Saharan Africa if pre-1985 trends had continued [346].

In Burundi, tuberculosis notifications more than doubled from 1984 to 1991 [347]. In a hospital in Brazzaville, Republic of Congo, the prevalence of HIV infection among tuberculosis patients increased from 3 to 24 per cent just in the 4-year period between 1988 and 1992 [348]. In Côte d’Ivoire, an autopsy study in patients who died with HIV infection revealed that tuberculosis was the cause of death in 32 per cent [349]. In Malawi, the number of notified tuberculosis cases almost quadrupled from 1985 to 1995 [350]. Large increases in tuberculosis notifications attributed at least partially to HIV have also been reported from Kenya [351], Zambia [352], Zimbabwe [353], and other countries [354, 355].
Figure 94. Relative odds of extrapulmonary tuberculosis as an indicator disease for AIDS in Europe, by country (selected countries only), transmission group, and sex. The dotted vertical line indicates the unity for comparison, the hollow circles represent the referent for each variable and horizontal boxes indicate the 95% confidence intervals. Data from [337].

Figure 95. Tuberculosis case notifications by type of disease, Tanzania, 1984–1995. Figure reproduced from [356].
In Tanzania, tuberculosis case notifications increased virtually exponentially from 1984 to 1995 (figure 95) [356]. In addition to the increase in the case rates, it is of particular concern that the age distribution of sputum smear-positive cases also changed. From 1984 to 1995, the peak rate of sputum smear-positive cases shifted very clearly to a younger age group (figure 96) [356]. This is an epidemiologically ominous sign, as improvements in the epidemiological situation must invariably be associated with an increase in the median age of tuberculosis patients, while conversely a lowering of the median age would indicate a higher perpetuation of transmission to younger population segments.

The large impact of HIV on the tuberculosis situation in Tanzania (indicative of the situation in many other sub-Saharan countries) has been shown in a country-wide, representative HIV seroprevalence study among tuberculosis patients [357]. The results indicate that between 30 and 45 per cent of patients with tuberculosis had HIV infection. The risk of HIV infection is unevenly distributed in the country. This investigation showed that the likelihood of a tuberculosis patient being infected with HIV was greatest in the 25- to 34-year-old age group. Female tuberculosis patients were more likely to be infected with HIV. Extrapulmonary tuberculosis cases were much more likely to have HIV infection. BCG vaccination appeared to provide relative protection against tuberculosis among adults as long as they were not HIV-infected. A second round of the survey has shown a rapid progression of HIV infection even over such a short span as 3 years (figure 97, unpublished data Tanzania National Tuberculosis/Leprosy Programme, WHO, and IUATLD). The rapid increase is particularly conspicuous in rural areas.

Asia

In Asia, the impact of HIV on tuberculosis morbidity has manifested itself most obviously in northern Thailand. In a study that included systematic HIV testing of tuberculosis patients, a rapid increase in tuberculosis notifications was noted between 1990 and 1994 (figure 98) [358]. This increase was apparently solely attributable to HIV infection.

Little is yet known about HIV infection among tuberculosis patients in India. One study from North Arcot, Tamil Nadu State, showed a prevalence of HIV infection of only two per cent among the over 3,000 tuberculosis patients tested [359]. However, the potential for increases in tuberculosis cases is large: the first four patients presenting with HIV infection in Manipur, one of India’s north-eastern States, in 1991, all presented with tuberculosis [360].

The potential impact of HIV infection on tuberculosis in Vietnam is illustrated by a study in the respiratory disease center of a hospital in Ho Chi Minh City,
Figure 96. Age-specific notification rates of sputum smear-positive tuberculosis cases in Tanzania, 1984 compared with 1995. Data from [356].

where 70 per cent of symptomatic HIV-infected patients had tuberculosis [361]. The prevalence of HIV infection among injection drug users rose from 1 per cent in 1992 to 39 per cent in 1996 in Ho Chi Minh City. HIV infection has been identified in 42 of 53 provinces of the country [362]. Consequently, tuberculosis is expected to noticeably increase in the very near future.

Latin America

In 1987-1988, HIV infection among patients with pulmonary tuberculosis in one study in Brazil was below 1 per cent [363]. In another study in Rio de Janeiro, HIV infection among tuberculosis patients was documented to have increased to just above 5 per cent in 1989 [364]. In a more recent study, tuberculosis is recognized as the third most common manifestation of AIDS, and transmission of tuberculous infection among HIV-infected persons has been documented to an important extent in the prison system in a study among incarcerated women in São Paulo, Brazil [365].

In Santo Domingo, Dominican Republic, a study of children with tuberculosis showed a seroprevalence of HIV of 6 per cent between 1991 and 1994 [366].

Contribution of primary progressive disease, exogenous re-infection, and endogenous reactivation to morbidity

The three components contributing to tuberculosis morbidity are:

● progressive primary tuberculosis (progression from sub-clinical infection to disease within 5 years after acquisition of infection [137]);

● cases resulting from exogenous reinfection with \textit{M. tuberculosis} (tuberculosis resulting from a second infection, superimposed on the first);

● and cases resulting from endogenous reactivation of a first infection (after more than 5 years following the first infection) [367].

Sutherland and colleagues developed a model for the Netherlands for the period from 1951 to 1970 to fit estimates of the three component contributors to observed data during this period [367]. The example in figure 99 for males in two age groups, 15 to 19 years old and 60 to 64 years old, shows not only the huge differences estimated for different ages but also the large changes over time. Over this period, the risk of tuberculous infection in the Netherlands declined annually by almost 14 per cent on average. For any given calendar year, tuberculosis resulting from primary infection was the biggest contributor in the younger age group and the other two
Figure 98. Tuberculosis notifications in Chiang Rai, by HIV status, Thailand, 1990–1994. Data from [358].

Figure 99. Estimated percentage contribution to morbidity from tuberculosis resulting from recent infection (filled circles), endogenous reactivation (triangles), and from exogenous reactivation (hollow circles). Model fitted to data from the Netherlands for males aged 15 to 19 and 60 to 64 years old, respectively. Data from [367].
contributors played only a minor role. In contrast, in the older age group, primary tuberculosis was a minor contributor to morbidity over the entire observation period, while the relative contributions from exogenous reactivation and endogenous re-infection exchanged their places in an important way: the smaller the risk of infection became, the more important was the role of endogenous reactivation in this age group.

These findings hold most likely universally: the larger the decline in risk of infection, and the older the population, the more important is endogenous reactivation of a pre-existing infection as a source of new cases. This has generally been the situation in most industrialized countries. Knowledge of the relative contributions of these three components to overall tuberculosis morbidity has important programmatic repercussions. The larger the proportion of cases that is attributable to either recent infection or exogenous reinfection and the smaller the proportion of cases attributable to endogenous reinfection, the more it likely is that the intervention targeted at reducing the duration of infectiousness will be successful.

Molecular approaches to characterizing individual *M. tuberculosis* strains allow an actual determination of the role of disease resulting from recent and reinfection tuberculosis compared to disease resulting from endogenous reactivation. Using such techniques as RFLP typing of *M. tuberculosis* strains, it could be shown in certain areas of the United States that tuberculosis resulting from recently acquired infection has been noted to occur much more frequently than had previously been expected.

In New York City, 40 per cent of tuberculosis cases were attributable to recent transmission [368], and a third in San Francisco [369]. It is only on this basis that it can be understood that improved control measures in New York City had an immediate impact on the tuberculosis incidence [245].

**Tuberculosis attributable to *M. bovis***

In industrialized countries, tuberculosis caused by *M. bovis* is believed to have become infrequent. However, the laboratory diagnosis might be difficult [47], and underreporting, or better, misclassification of the organism as another species of the *M. tuberculosis* complex may thus be common. Kleeberg has reviewed the frequency of reports of *M. bovis* [370]. In the first decades of the twentieth century, bovine tubercle bacilli were frequently reported to be responsible for 10 to 30 per cent of cases of lymphatic tuberculosis, and for some 5 per cent of cases of pulmonary tuberculosis. By the 1960s, the proportion of cases had decreased to less than 1 to 2 per cent in most countries reporting.
Predictive epidemiology: prospects for the epidemic

Industrialized countries

The course of the tuberculosis epidemic in industrialized countries will be shaped by multiple factors. In the majority of countries of Western Europe, HIV infection is expected to exert only a small impact, as the population segments at particular risk of HIV infection are rapidly replaced by cohorts with virtually no tuberculous infection. While HIV infection may cause increases in tuberculosis in HIV-infected substance abusers, it is unlikely that this excess morbidity will affect morbidity to a great extent in the general, indigenous population of these countries.

Härö has projected tuberculosis incidence for Finland based on a historical analysis of tuberculosis morbidity within birth cohorts [271]. These projections predict a rapid decrease of tuberculosis in the indigenous population in Finland (figure 100). By the year 2015 to 2020, the expected incidence will have fallen below one case per million population.

Figure 100. Projected incidence of respiratory tuberculosis in Finland for males for the years 2000 and 2010. Observed values are age-specific reported tuberculosis cases within 5-year birth cohorts. Dotted lines indicate projections of incidence within the same birth cohorts. From these, the expected age-specific incidence is construed for the years 2000 and 2010. Data from [271].
The Finnish projections are likely to hold in general for the indigenous population of most of Western Europe, with a caveat, however: as has been shown, international migration increasingly impacts on tuberculosis notifications in industrialized countries. Thus, tuberculosis morbidity in many of these countries may continue to decrease, level off, or even increase, depending on the extent of migration and the origin of the immigrants. As the risk of infection from indigenous cases is rapidly decreasing, the contribution of transmission from cases among migrants will become relatively more important. For the Netherlands during the years from 1993 through 1995, it was estimated that slightly less than one fifth of all cases among the Dutch population were the result of recent transmission from a non-Dutch source [304]. The absolute magnitude of cases resulting from this type of transmission is relatively small, and is likely to remain so, but its relative contribution is expected to slow down the natural trend of self-elimination of tuberculosis in the countries concerned.

In multi-ethnic countries such as the United States it has been shown that laxity in control can rapidly lead to a resurgence of tuberculosis in ethnic/racial groups with a higher prevalence of tuberculous infection than the general population. This will be particularly important if several factors accumulate that favor transmission and progression from infection to disease. These experiences, especially those reported from New York City, also show that decisive control measures can clearly influence the course of the epidemic.

**Low income countries**

Tuberculosis and HIV infection are two conditions that have become intrinsically linked as the prevention of the former precariously depends on the integrity of the cellular immune system, while the latter destroys precisely that. Consequently it is likely that the tuberculosis epidemic will worsen considerably in the near future in many low income countries. In countries with a currently high prevalence of tuberculous infection in young age groups, but which are not seriously affected by the HIV pandemic, the situation remains precarious. The course of the epidemic in such countries will considerably depend on the success in implementing control strategies aiming at reducing the risk of infection in the community as swiftly as possible.

A model predicting the future size of the tuberculosis problem in sub-Saharan Africa suggests that under the most optimistic assumptions tuberculosis will increase by 50 to 60 per cent between 1980 and the year 2000 [371]. In this model, even scenarios with a ten-fold increase in sputum smear-positive cases are not entirely unlikely.
An analysis undertaken by the WHO in 1993 attempted to project the development of the tuberculosis incidence in various regions of the world. The projections considered the current state of control measures, population growth, the impact of HIV infection, and other factors [289]. The prospect for sub-Saharan Africa contrasts with that for highly industrialized countries (figure 101).

With current control efforts and without input of additional resources to control transmission from and expedient diagnosis of newly emerging infectious cases of tuberculosis, the global tuberculosis situation is likely to worsen in the coming years, as predicted in a model by Dye and collaborators from the WHO [126]. The threat is particularly acute for countries already seriously affected by the HIV pandemic and currently poor tuberculosis control programs.

Figure 101. World Health Organization forecasts of tuberculosis incidence, comparing Africa and highly industrialized countries, years 1990–2005. Data from [289].
Death from tuberculosis is the longest recorded indicator of the tuberculosis epidemic in industrialized countries. In contrast, information on deaths is rarely or never systematically collected in low income countries. Some precautions must be kept in mind when interpreting mortality or case fatality data. Tuberculosis mortality has been recorded for several centuries in some countries, long before the infectious nature of the disease and its causative organism were identified. The determination of tuberculosis as the cause of death, particularly in older data sources, is far from certain. In the second half of this century, chemotherapy has markedly reduced tuberculosis case fatality. In many industrialized countries, mortality has thus become a rather frail indicator of the extent of the problem. Furthermore, statistics in industrialized countries draw on information from death certificates, a source of multiple possible errors [372, 373].

Etiologic epidemiology: risk factors for death from tuberculosis (case fatality)

The risk of dying from tuberculosis is dependent on the site and type of the disease and the timeliness of diagnosis. In the pre-chemotherapy era, case fatality from tuberculous meningitis approached 100 per cent [374, 375]. On the other hand, tuberculosis of the peripheral lymph nodes is fairly benign in terms of case fatality [160], unless complicated by tuberculosis at other sites resulting from or in lymphohematogenous spread.

From an epidemiologic point of view, death from pulmonary tuberculosis is the most significant, and it accounts for the majority of deaths. The form of pulmonary tuberculosis is of great importance, as sputum smear-positive tuberculosis has a much higher fatality than sputum smear-negative tuberculosis. This is exemplified in a long-term follow-up of tuberculosis patients admitted to a Swiss sanatorium between 1912 and 1927, where case fatality approached 60 per cent in cases with
“open” tuberculosis compared to about 10 to 15 per cent in cases with “closed” tuberculosis (figure 102) [376]. Several studies from the pre-chemotherapy era testify to the high case fatality of sputum smear-positive tuberculosis [129, 377-382]. Three examples, from the United Kingdom [380], Sweden [129], and Denmark [382], show that about one third of patients with sputum smear-positive tuberculosis had died 1 year after diagnosis, and by 5 years half to two thirds of patients had died (figure 103).

The prognosis of tuberculosis has appreciably improved with the advent of chemotherapy. Nevertheless, case fatality from tuberculosis continues to remain unacceptably high for various reasons.

It is difficult to ascertain to what extent patients die from undiagnosed tuberculosis, because most of these cases never become known. Nevertheless, the relative importance of the problem has been addressed in several studies. In a 4-year period from 1985 to 1988, 5 per cent of all tuberculosis cases notified in the United States never received anti-tuberculosis treatment [383]. The proportion of cases diagnosed at death increased with age (figure 104), and was 15 to 20 per cent for miliary, meningeal, and peritoneal tuberculosis. This study did not distinguish between cases recognized at autopsy only and cases clinically suspected but with the definite diagnosis becoming available only after the patient’s death. From other studies, it

![Figure 102](image)

**Figure 102.** Long-term follow-up of the fate of pulmonary tuberculosis patients without specific treatment in a Swiss sanatorium, by initial bacteriologic findings. Data from [376].
Figure 103. Cumulative case fatality from untreated sputum smear-positive pulmonary tuberculosis in the United Kingdom 1928–1938 [380], Sweden 1910–1934 (15 to 29 years old) [129], and Denmark 1925–1929 (16 to 25 years old) [382].

Figure 104. Proportion of tuberculosis cases diagnosed after death, by age, United States, 1985–1988. Data from [383].
is known that both are relevant [143, 349, 384, 385]. Several autopsy studies have revealed that tuberculosis may be entirely missed and never suspected [385-388]. In three different studies, tuberculosis was the underlying or a contributing cause of death in up to 50 per cent of tuberculosis cases unrecognized until autopsy [384, 385, 389]. Case fatality has also been shown to vary according to the specialty of the diagnosing physician [390].

Failure to diagnose tuberculosis may result not only in the death of the patient, but may also — in persons with pulmonary tuberculosis — contribute to unrecognized transmission to family, friends, health care workers and other contacts. In a survey in Baltimore in the early 1960s, 44 per cent of cases first registered at death were considered to have been communicable [391]. A failure to diagnose pulmonary tuberculosis places those who have become infected by such cases beyond the reach of preventive measures. In industrialized countries, reports indicate that the overwhelming majority of patients reported or diagnosed at death were hospitalized prior to death [384, 387, 391]. This indicates that a window of opportunity for preventing these deaths was lost. The reasons for failing to make the diagnosis include omission of appropriate diagnostic measures for the confirmation of tuberculosis [373, 392, 393], and mistaking radiographic or clinical signs and symptoms of tuberculosis for those of malignancy, pneumonia or other conditions [387]. Recent studies of tuberculosis patients with HIV infection have shown that delayed diagnosis of tuberculosis resulting in death could have been avoidable if adequate numbers of specimens had been obtained, and if empirical chemotherapy of tuberculosis had been administered to symptomatic patients with chest radiographs suggestive of tuberculosis [394-396]. Once diagnosed, the quality of treatment also affects treatment outcome. A study from North Arcot, India, demonstrates this clearly [397]. Patients had an increasing probability of dying the less they took of the prescribed medication, a point accentuated with increasing age (figure 105).

**Descriptive epidemiology: mortality**

Data on mortality from tuberculosis have been collected for a long time in many industrialized countries [398-401]. Available data suggest that tuberculosis mortality in England began to increase in the sixteenth century, and reached a peak of mortality in London, probably before 1750 (figure 106) [399]. The tuberculosis epidemics in the capitals of Western Europe reached their peak in the first half of the nineteenth century, and those in Eastern Europe a few decades later. Although the accuracy of the diagnosis of tuberculosis was certainly limited, and data collection might have been irregular, the general pattern is obvious. It would perhaps justify the conclusion
Figure 105. Age-specific proportions of patients dying of pulmonary tuberculosis following rifampicin-containing treatment, 6 to 36 months after treatment start, by amount of drugs taken and age, North Arcot District, India. Data from [397].

Figure 106. Mortality from tuberculosis in London, Stockholm, and Hamburg, modeled from available data. Figure reproduced with the permission of the American Thoracic Society / American Lung Association from [399].
that the tuberculosis epidemic in Western Europe has spanned at least some 300 years, and has claimed, at its height, the lives of 1 per cent of the population per year.

Data from industrialized countries during the twentieth century show that tuberculosis mortality progressively declined, but each of the world wars resulted in a peak in mortality [402, 403]. Particularly instructive are the mortality data from Germany for the period from 1892 to 1940 (figure 107) [402]. Tuberculosis mortality increased steeply and early in the war, but at the end of World War I it fell rapidly and resumed its previous decline. Thus, the long-range trend of mortality remained unaffected by the epidemic during the war. It must therefore be concluded that risk of infection most likely continued to decline throughout the war, and that the increased morbidity that most likely occurred (but was not measured), was also associated with increased and accelerated case fatality that prevented excess transmission [122].

Usually, age-specific mortality data are displayed cross-sectionally as they are collected in a given year. Andvord from Norway first pointed out that cross-sectional mortality data provide an incomplete picture of tuberculosis epidemiology, and that an analysis by birth cohort was much more informative [273-275]. Analyzing tuberculosis mortality data from Massachusetts cross-sectionally, Frost found that the

![Figure 107. Tuberculosis mortality in Germany, 1892–1940. Figure reproduced with the permission of Georg Thieme Verlag from [402].](image-url)
death rate was relatively constant beyond the age of 30 years [276]. However, within each birth cohort, tuberculosis mortality peaked in infancy and between the ages of 20 and 30 years. The mortality data from respiratory tuberculosis from Switzerland are illustrative of the situation for industrialized countries where such information is available. For example, the cross-sectional data show that in each successive decade tuberculosis declined, but since 1901 it has increased with increasing age among males (figure 108) [404]. Among females, two peaks can be observed, one among young adults and the other in the elderly (figure 109). Clearly, the high rates seen in the elderly on cross-sectional examination of the data reflect only residuals of a much higher mortality experienced by the cohort when it was much younger.

Three major factors determine the characteristics of mortality. These include age-specific differences in mortality; differences in mortality in each cohort; and differences at particular periods or events, such as during a war.

Cross-sectional studies alone of mortality thus fail to tell the whole story. Age-specific differences in each cohort are apparent and indicate, as for example in Switzerland, that through the entire period of observation tuberculosis claimed its highest toll in young adults within each successive cohort. In the example from

**Figure 108.** Age-specific mortality from respiratory tuberculosis among males, 1901–1981, Switzerland. The dotted lines show the data analyzed cross-sectionally, the solid lines the analysis by birth cohort. Figure reproduced with the permission of Birkhäuser Publishing Ltd. from [404].
Switzerland each successive cohort experiences less mortality (figures 108 and 109) [404]. This way of looking at the data also permits one to look much further back into the past than is possible by examining the data only cross-sectionally. In this example, the first cross-sectional survey is available for 1901. Nevertheless, the cohort-contour approach suggests that tuberculosis mortality in Switzerland had started to decline much earlier.

The degree of parallelism observed between lines produced in successive cohorts is the essence of the period-contour approach. A distortion caused by deviation from parallelism indicates the importance of secular or period events that modify tuberculosis mortality. The model proposed by Collins [405] and his analysis of data from England and Wales suggest that the introduction of medical measures, although resulting in a marked period influence, could account for only 2 to 3 per cent of the total decline in mortality from respiratory tuberculosis. In other words, the trend in cohort effects suggested continuing declines in deaths from tuberculosis even in the absence of medical measures.

Estimates of the global death toll from tuberculosis have been given as 2.5 million per year [288], yet more recently WHO has revised these estimates, now set at 1.5 million deaths per year [406].

Figure 109. Age-specific mortality from respiratory tuberculosis among females, 1901–1981, Switzerland. The dotted lines show the data analyzed cross-sectionally, the solid lines the analysis by birth cohort. Figure reproduced with the permission of Birkhäuser Publishing Ltd. from [404].
Predictive epidemiology: expected death toll from tuberculosis

Because case fatality is so closely linked to the availability of timely and curative chemotherapy of clinically manifest, particularly sputum smear-positive, tuberculosis, the magnitude of tuberculosis mortality in the future will not so much depend on the epidemiology of tuberculosis as on the availability of effective treatment.

The WHO has provided estimates for the number of deaths attributable to tuberculosis for the years 1990 – 2000 [289]. These estimates are based on the assumption that implementation of intervention strategies stay the same as those estimated for 1990 throughout the period examined. According to these estimates, less than 1 percent of mortality occurs in the industrialized world. In Africa, the number of deaths is expected to increase disproportionately (figure 110).

Figure 110. World Health Organization estimates of the number of annual deaths due tuberculosis, 1990–2000. Data from [289].
A summary of the epidemiology of tuberculosis, looking at type of epidemiologic question in the tuberculosis model, is provided in table 3.

**Exposure**

Significant exposure depends on the presence of incident cases, but important modifiers determine the magnitude of exposure. These include most particularly duration of infectiousness and number of case-contact interactions per unit of time of infectiousness. These modifiers may vary greatly across populations, time, population density, climatic conditions, and other factors.

**Infection**

Tuberculosis is essentially transmitted by the airborne route. Milk contaminated with *M. bovis* can be an important source of infection in the community, but it appears that its contribution to overall tuberculosis morbidity is generally low. Tuberculosis is not as easily transmitted as other airborne communicable diseases. The number of bacilli expelled by a tuberculosis patient critically determines the number of persons that the patient is capable of infecting. Patients with sputum smears positive for acid-fast bacilli on direct sputum smear microscopy are the major sources of infection in the community.

The risk to a susceptible individual of becoming infected depends on the density of tubercle bacilli in inhaled air and the duration of exposure to that air. Ventilation is thus important to reduce the concentration of infectious droplet nuclei in the air. Infectious droplet nuclei may remain suspended in indoor air for a prolonged period of time retaining their potential of infectiousness, while outdoors, exposed to the ultraviolet rays of sunlight, rapid death of bacilli is the rule.

Theoretically, the incidence of infection is epidemiologically the most informative parameter, because it points to the extent of current transmission in the community.
Table 3. Summary of the epidemiology of tuberculosis.

<table>
<thead>
<tr>
<th>Type of epidemiology</th>
<th>Exposure</th>
<th>Infection</th>
</tr>
</thead>
</table>
| Etiologic            | - Incidence  
                       - Duration of infectiousness  
                       - Contacts / unit time | $P/V \cdot \Delta t$  
                       - $P$: Number of infectious particles expelled into shared air space – exogenous, source determinants (site of disease, aerosolization, bacterial load)  
                       - $V$: Volume of airspace shared  
                       - $\Delta t$: Duration of exposure |
| Descriptive          | ?? – Not amenable to quantification (no tool to measure available) |  
                       - Method: tuberculin test  
                       - Methodological issues – operating characteristics of tuberculin test  
                       - In low-income countries, half or more of productive population infected  
                       - In industrialized countries, half of the infected are elderly, 1/10 or less of productive population infected  
                       - Global prevalence of tuberculous infection: approximately 1/3 of the population |
| Predictive           | Depending on modifiable etiologic covariates |  
                       - Modeling risk of infection  
                       - Trends: rapidly declining in industrialized countries, slowly declining in low-income countries, probably rising in several HIV-affected low-income countries |
<table>
<thead>
<tr>
<th>Disease</th>
<th>Death</th>
</tr>
</thead>
</table>
| **Major risk factors (endogenous):**  
- Time since infection  
- HIV infection  
- Untreated or poorly treated previous tuberculosis  
- Other factors, e.g., age, gender, malnutrition, diabetes, silicosis, malignancies, genetic factors, etc. | **Major risk factors for case fatality:**  
- Major site of disease  
- Delay in diagnosis and treatment  
- Age |
| - Risk of disease highest in young adults  
- Long-term decline in morbidity in the indigenous population of industrialized countries, with minor disturbances  
- Increasing fraction of international migration-associated tuberculosis in industrialized countries  
- Morbidity only slowly declining in many low-income countries, exponential increases observed in some HIV-affected countries  
- Estimated 7.5 to 8 million new cases per year | - Mortality highest in young adults  
- Epidemic in industrialized countries spanning 300 years, coming to an end  
- Low-income countries: no information  
- Official estimates set at 1.5 million deaths / year, difficult to verify |
| - Gloomy prospects for countries with currently high prevalence of tuberculous infection in young population segments and concomitant HIV epidemic  
- Critical equilibrium in countries with high prevalence of tuberculous infection in young populations but low prevalence of HIV infection  
- Excellent for the indigenous population of most industrialized countries | - Reduced to insignificance in the indigenous population of industrialized countries  
- In low-income countries depending on development of morbidity and implementation of programs |
It is usually not feasible to measure infection incidence, and the derivation of the average annual risk of infection from a tuberculin prevalence survey, a proxy of this parameter, has become one of the most cherished tools in tuberculosis epidemiology. Unfortunately, tuberculin skin testing is fraught with problems of a technical nature, including selection of standardized tuberculin, the technique of administration, and reading of the test result. Even if all these barriers are overcome, it is often exceedingly difficult to interpret the results and to arrive at an estimate of the prevalence of infection. Sensitization to environmental mycobacteria and \textit{M. bovis} BCG results in cross-reactions with the standard tuberculin. The higher this sensitization and the lower the prevalence of true infection with tubercle bacilli, the more difficult it becomes to disentangle the truth from confounding factors.

Should the determination of infection prevalence be successful, then the calculation of the risk of infection from a single or even two sequential surveys provides only information on the extent of transmission somewhere in the past, determined by the age of those tested. Because of its “averaging” characteristic, the risk of infection may not capture short-term changes.

The knowledge of risk of infection cannot provide precise information on the expected case incidence. It can only express the extent to which such cases are capable of transmitting tubercle bacilli within the community. This is a function of the number of infectious cases, the duration of infectiousness, the characteristics of exposure, and the population exposed.

The determination of risk of infection has, nevertheless, been regularly used to compare the extent of the tuberculosis problem in various populations. It is, if technically interpretable, the only available mean to indicate the extent of transmission that has occurred, on average, over specified periods of time in the past. It is useful for global estimates of the level of the tuberculosis problem in a community and of trends over relatively long periods of time.

Infection is unevenly distributed in the world. In most industrialized countries, a very large proportion of those aged 65 years and over remains infected, while younger generations are increasingly free of infection, because the risk of infection has rapidly declined from very high to very low levels over the past decades. In contrast, in many countries in Asia, Africa, and Latin America, the risk of infection continues at much higher levels than is the case in industrialized countries, and the trend, rather than declining rapidly, has declined more slowly or not at all.

Current estimates suggest that about one third of the world’s population is infected with \textit{M. tuberculosis}. In industrialized countries, the bulk of infected persons is found among the elderly, while in most low income countries, the large majority of infected persons is in the economically most productive and reproductive age groups.
Improving the epidemiologic situation will require means that will reduce the risk of infection in the community. This will initiate a cohort effect, resulting in a reduction of the pool of infected persons by replacement of heavily infected cohorts with cohorts with increasingly less infection. Reducing the risk of infection entails both timely identification of infectious cases and curative treatment of those identified. The latter emphasis is of importance as inadequate treatment may prolong survival at the cost of the creation of patients with prolonged excretion of tubercle bacilli. Furthermore, bacilli in such patients are likely to be more frequently resistant to medications used in standard treatment protocols, thus leading again to secondary treatment failures among those who progress to clinically overt tuberculosis following infection from such a chronic excretor.

**Disease**

A multitude of risk factors for the progression from sub-clinical, latent infection with *M. tuberculosis* to overt disease has been identified. The three most important for all practical purposes are co-infection with HIV, recent infection with tubercle bacilli, and healed lesions from previous tuberculosis which was never treated. Globally, low body weight might be of considerable epidemiological importance because of its high prevalence, particularly in low income countries. Other factors, such as for example jejunoileal bypass surgery, are essentially epidemiological curiosities.

In industrialized countries, tuberculosis has markedly declined in the twentieth century and is on the brink of its elimination in the indigenous population of many of these countries. Nevertheless, increased international migration will further contribute to tuberculosis morbidity in many of these countries. Large differences in disease frequency can also be observed within industrialized countries, disproportionately affecting the poor and, as in the United States, ethnic/racial minorities. Among the latter, tuberculosis remains largely a disease of the young, while in most of Europe, among the non-Hispanic white population in the United States, and in non-aboriginal (but born in the country) populations of other affluent countries such as Australia, New Zealand and Canada, tuberculosis has become a disease of the elderly. In population segments in which a large proportion of young people continues to be infected with tubercle bacilli, the HIV epidemic is causing a resurgence of tuberculosis, be it in affluent or low income countries. In many industrialized countries, the proportion of cases that are foreign-born increases steadily.

The distribution of tuberculosis is very uneven throughout the world. Of the estimated 7.5 to 8 million estimated cases emerging globally each year, only
5 per cent occur in industrialized countries. The HIV pandemic is drastically worsening the tuberculosis situation in countries with a high prevalence of infection with *M. tuberculosis* among young people and a high incidence and prevalence of HIV infection. At this point in time, this is particularly the case in sub-Saharan African countries. HIV may alter the course of tuberculosis directly by leading to reactivation of latent infection or by direct progression from new infection to disease depending upon which comes first, HIV or *M. tuberculosis* infection. Even more ominous from an epidemiological point of view is the inherent danger that excess cases caused by these mechanisms are likely to result in a disruption of the encouraging decline in risk of infection with *M. tuberculosis* in the past.

**Death**

Tuberculosis case fatality is largely determined by site and type of disease and by appropriate and timely intervention. Untreated sputum smear-positive tuberculosis leads to death in about 30 to 40 per cent of cases within one year, and cumulatively kills about 50 to 70 per cent within 5 to 7 years. Unrecognized tuberculosis is not uncommon either in low income countries, or in industrialized countries. With decreasing frequency of the disease in industrialized countries, the diagnosis may indeed be missed more frequently, leading to unnecessary death and, in the case of infectious pulmonary tuberculosis, to unrecognized transmission of tubercle bacilli.

Tuberculosis mortality data indicate that the current tuberculosis epidemic in industrialized countries has spanned some several centuries, reaching its peak in different European countries between the end of the eighteenth and the middle of the nineteenth century. Tuberculosis mortality is no longer useful as an epidemiologic index, largely because it has become a rare event in most industrialized countries, and because it is usually not reliably recorded and reported in low income countries.

Currently, it is estimated that some 1.5 million people die each year from tuberculosis.

The extent to which tuberculosis continues to kill will largely depend on the extent to which modern intervention strategies become available to the low income countries of the world.
References


17. Wells WF. On air-borne infection. Study II. Droplets and droplet nuclei. Am J Hyg 1934;20:611-8


70. Seibert FB, Morley N. The relationship of the tuberculin proteins of different acid-fast bacilli to sensitization as indicated by their reactivity in sensitized animals. J Immunol 1933;24:149-56.


273. Andvord KF. Hvad kan vi lære ved å folge tuberkulosens gang fra generasjon til generasjon? (What can we learn by studying tuberculosis by generation?). Norsk Magasin for Laegevidenskaben 1930;91:642-60.


275. Andvord KF. What can be learned from following the development of tuberculosis from generation to generation. (Unpublished translation by Gerard Wijsmuller, undated typescript at the Annik Rouillon Documentation Centre of the IUATLD). Norsk Magasin for Laegevidenskaben 1930;91:642-60.


ISBN 2-9504238-8-4