CHAPTER 1

Tuberculosis Undefeated

There may be 12 to 25 million infectious cases of tuberculosis in the world. Confronted with such an astronomical computation of distress, it is clear that tuberculosis is very far from being defeated.

—*John Crofton (1960)*

In 2012, an estimated 8.6 million people developed TB and 1.3 million died from the disease. There were 12 million prevalent cases.

—*World Health Organization (2013a)*

"Tuberculosis Undefeated" was the title chosen by Crofton for his 1960 Marc Daniels lecture, given in memory of a colleague who made a vital contribution to the development of combination drug therapy against tuberculosis (Crofton 1960). Much has happened in the population biology of TB since 1960, but Crofton's title still reflects the state of TB in the world today, as reflected in WHO statistics (World Health Organization 2013a). Despite the development of highly efficacious drug therapy to prevent and cure TB, 8–10 million people developed some form of the disease in 2012, there were 11–14 million extant (prevalent) cases, and 1.2–1.5 million people died. The number of TB episodes each year has probably now fallen below an all-time high, reached in the decade 2000–2009, but TB remains the largest cause of death from a single, curable infectious agent.[1]

The challenge we face in the twenty-first century is to control—and ultimately eliminate—a pathogen that has inhabited human populations (including the earliest hominids) for tens of thousands and possibly millions of years. To begin to understand how elimination could be achieved, the purpose of this

---

[1] HIV/AIDS causes more deaths; HIV is treatable with antiretroviral drugs but is not yet curable.

opening chapter is to define the problem. The following sections describe the essential characteristics of the pathogen and the collection of illnesses it causes, its origins and distribution in human populations, the dominant trends through time, and the burden of disease today. This is an explicitly descriptive account, which prefigures the more analytical investigation of TB epidemiology and control in subsequent chapters.

MYCOBACTERIA

Everything about the biology of *Mycobacterium tuberculosis* happens slowly. Whereas other bacteria divide in minutes, *M. tuberculosis* takes hours. Most bacteria can be cultured in hours, but *M. tuberculosis* takes weeks. Most outbreaks of bacterial infections run over days or weeks, but an epidemic of *M. tuberculosis* takes years, decades or centuries.

M. tuberculosis is a small (measured in μm), rod-shaped, aerobic, mostly nonmotile, non-spore-forming bacterium. Following convention in bacteriology, mycobacteria are classified according to their reactions to standard stains used in microscopy. *M. tuberculosis* is bracketed with gram-positive bacteria, although its waxy cell wall, rich in fatty acids, is Gram neutral (Hinson et al. 1981). Nevertheless, once the poorly absorptive cell wall has been successfully impregnated with concentrated dyes (fuchsin), it withstands decolorization by acids—hence the use of an acid-fast stain in procedures such as Ziehl-Neelsen, which renders all mycobacteria bright red on a blue background when viewed under a light microscope. Mycobacterial species are distinguished by their appearance—*M. tuberculosis* is comparatively "rough" among its congeners—and by their response to biochemical tests (Ait-Khlaed and Enarson 2003).

The slow growth of mycobacteria in culture is somewhat less slow in liquid (days) than on solid media (weeks on the traditional Löwenstein-Jensen slopes). In diagnostic laboratories that rely on culture, the slow growth can cause significant diagnostic delays, with clinical and epidemiological consequences, including the loss of patients and continuing transmission before the start of treatment. In partial compensation, the variable growth rates of mycobacteria species are an asset in differential diagnosis.

Mycobacteria are widespread in the environment and commonly found in soil and water. Not all are obligate parasites. The *M. tuberculosis* complex (MTBC) of species contains the major mammalian pathogens, including *M. tuberculosis sensu stricto* and *M. africanum*. It also includes the mainly animal pathogens *M. bovis* (cattle), *M. caprae* (goats), *M. microti* (voles), and *M. pinnipedii* (seals, sea lions; Smith et al. 2006). As a cause of human TB, *M. bovis* is more commonly

associated with extrapulmonary than pulmonary disease (Durr et al. 2013). With improved management of infection in cattle, plus the pasteurization of milk, there are nowadays few human cases due to *M. bovis* (median < 3% worldwide). However, in localities where these basic control methods are not used, the percentage of human cases due to *M. bovis* can be much higher (up to 30%–40% in surveys; Muller, Durr et al. 2013). Two other important human mycobacteria in the same genus are the agents of leprosy (*M. leprae*) and Buruli ulcer (*M. ulcerans*). The other 100-plus known species—more surely await discovery—are called environmental mycobacteria or, less satisfactorily, nontuberculous mycobacteria (NTM, which can, in fact, cause lung tubercles), atypical mycobacteria, or mycobacteria other than tuberculosis (MOTT bacilli). Despite uneven sampling effort, it is clear that NTM are distributed worldwide and that species differ among continents and among countries within each continent (Hoefsloot et al. 2013).

M. tuberculosis and *M. africanum* are the principal causes of human lung disease (pulmonary TB), in addition to diseases of other organs and tissues (extrapulmonary TB), notably lymph glands, bone, brain and skin, but all organs are vulnerable. Disseminated disease, caused by the spread of bacteria away from the lung to other organs, is most likely in individuals with compromised or immature immune systems. Among these are people with HIV infection, diabetes, malignancies, infants, and people on therapies, including tumor necrosis factor inhibitors and corticosteroids.

Compared with other bacteria, *M. tuberculosis* is renowned for its lack of genetic variability, though the view that *M. tuberculosis* is essentially a clone has been overturned by ever more refined explorations of the genome (Borrell and Gagneux 2011; Gagneux 2013; Coll et al. 2014; Anderson and May 1991). Furthermore, *M. tuberculosis* has a remarkable capacity to generate drug-resistant mutations, which will appear in almost any patient that is treated with suboptimal doses of drugs (Chapter 5). Even with the limited genetic variability known so far, a suite of markers is proving to be immensely informative about *M. tuberculosis* biology. The toolbox for investigating genetic variation includes restriction fragment length polymorphisms (RFLP) of the *M. tuberculosis* insertion sequence IS6110, spoligotyping (based on unique, nonrepetitive DNA spacer sequences), variable number tandem repeats (VNTR), single nucleotide polymorphisms (SNPs), multilocus sequences (MLS) and, increasingly, whole genome sequencing (Tortoli 2011; Ford et al. 2012; Walker et al. 2013; Bryant et al. 2013; Merker et al. 2013; Kohl et al. 2014; Anderson 2014; Niemann and Supply 2014). Genetic markers are being used to study, among other things, who acquires infection from whom, whether recurrent episodes of TB are caused by reinfection or relapse, the population dynamics of benign, virulent, and drug-resistant strains, and the history and geography of *M. tuberculosis* dispersal.

HUMANS AS HOSTS

M. tuberculosis is a parasite of the human immune system. It attacks the host at the heart of its defenses, inducing transmission by leading the host to its own self-destruction.

The bacilli survive and replicate in phagocytic cells, including macrophages, neutrophils, monoctyes, and dendritic cells. Macrophages are the very cells that normally protect the host by engulfing foreign bodies, including microbes. A proportion of invading *M. tuberculosis* are killed immediately by an innate immune response. The survivors must establish an infection that will eventually allow transmission to other hosts by stimulating cells of the immune system to produce an adaptive response that is neither too weak nor too strong. Too little immunity allows the multiplication and dissemination of infection away from the lungs, causing widespread tissue and organ damage, eventually killing the host and limiting the chances for transmission. Too much immunity, innate or acquired, prevents bacilli from replicating or kills them prematurely before transmission is possible (van Crevel et al. 2002; Verrall et al. 2014; Ernst 2012; Cooper and Torrado 2012).

Transmission takes place when individuals with disease of the lung produce airborne droplets while coughing, sneezing, spitting, or just talking. The droplets typically stay airborne for a few hours, during which time they can be inhaled by another person.

The course of *M. tuberculosis* infection begins when inhaled droplets carrying *M. tuberculosis* lodge in the lung tissue, where they are captured by scavenging macrophages. Bacilli replicate slowly within each macrophage, bypassing the host cell's usual destructive mechanisms. Replication leads eventually to breakout and the colonization of more macrophages as they are recruited to the site of infection. In this initial phase, infected macrophages spread through the lymphatic system to the draining lymph nodes, but infection is otherwise contained in the lung.

With the stimulation of macrophages and other immune cells, the human host builds a cell-mediated immune response that stops bacterial replication after a period of several weeks. At this point, the presence of infection and the ensuing cell-mediated immune response are detectable by antigen stimulation tests. The Pirquet, Mantoux, or tuberculin skin test evokes an immune response in the skin via intradermal inoculation of *M. tuberculosis* purified protein derivative, causing erythema, edema, and induration on the skin surface. Alternatively, macrophages and T cells stimulated with antigens specific to *M. tuberculosis* will signal the presence of infection via the production of interferon gamma (the basis of interferon gamma release assays, IGRAs), among other cytokines (Grosset 2003; Pai et al. 2014).

Populations of activated macrophages and lymphocytes that congregate at the site of infection form granulomas, which appear as nodules or tubercles (Ehlers and Schaible 2012). Within the granulomas, the continuing effort to control the multiplication of bacilli causes the human host to destroy its own lung tissue. Cell death within the granulomas creates Emmentaler-like[2] lung cavities (caseous necrosis), typically in the middle and lower lung, a hallmark of pulmonary TB.

If the granuloma breaks down, the bacilli can continue to multiply outside as well as inside macrophages (Grosset 2003; Shakak et al. 2013). If the infection is able to proceed, replication leads to the creation of more cavities, and the lung is progressively destroyed in a process that may take months or years. The lung damage causes coughing, with the expulsion of bacilli to the outside air and the expectoration of sputum, with blood in the advanced stages of disease. For TB, the extent and severity of pulmonary disease is associated with infectiousness, and this alignment is used in mathematical modeling of *M. tuberculosis* transmission (Chapter 2).

The more effective the immune response, the slower the rate at which individuals progress to active disease, having acquired infection. Varying degrees of suppression by immune responses generate a spectrum of latent or subclinical states, ranging from the early containment of infection by an innate or an acquired immune response, through the containment of live but nonreplicating bacilli and bacterial replication at low levels partially contained by immunity (Barry et al. 2009; Young et al. 2009; Colangeli et al. 2014). Tuberculin skin testing to detect cell-mediated immune responses suggests that the human population carries a vast reservoir of live bacilli. The conventional wisdom is that, in long-term infections, the bacilli are in a quiescent or latent state, which can reactivate in any individual after years or decades of dormancy. However, as latency comes under closer scrutiny, the idea that this is a single state (Barry et al. 2009) or even that latency exists at all (Shakak et al. 2013) is under challenge (see also Chapter 2). A more contemporary view is that latent tuberculosis—infection in the absence of clinical symptoms—is a collection of conditions that form part of a heterogeneous response to *M. tuberculosis* invasion. Infection leads to the formation of distinct granulomatous lesions, with differential ability to support or suppress bacterial replication (Barry et al. 2009; Lin et al. 2013). Whatever the true nature of persistent infection, some people will develop TB months or years after first exposure. But the great majority of healthy people who have signs of exposure to infection and who presumably carry viable bacilli do not live long enough to develop TB disease.

---

[2] Although this cheesy metaphor is often invoked, the cavities caused by TB are rarely as smooth as Emmentaler holes, which are caused by another bacterium, *Propionibacterium freudenreichii*, releasing carbon dioxide (H.L. Rieder, personal communication, and Wikipedia).

M. TUBERCULOSIS AFFECTING THE
DISTRIBUTION OF OTHER DISEASES

M. tuberculosis is one of a set of pathogens interacting competitively or facultatively via the human immune system. The adaptive immune response described previously generates two antagonistic subclasses of T helper cells—Th1 and Th2—each with its own set of cytokine mediators. Microbial infections have the potential to influence the balance between Th1 and Th2 responses by altering cytokine profiles, with positive or negative consequences for the progression of disease. Bacterial infections probably have such a role in atopy, an allergic state producing mucosal inflammation characteristic of asthma and characterized by overreactive Th2 responses.

In simple terms, the Th1 response to *M. tuberculosis* infection is associated with granuloma formation and protection, whereas the Th2 response results in tissue-killing hypersensitivity and the progression of disease. The processes that determine the balance of the two responses also affect and are affected by the interaction between *M. tuberculosis* and other infectious agents, from viruses to helminths.

Because mycobacteria elicit strong Th1 responses, thereby shifting the Th1/Th2 balance away from Th2, *M. tuberculosis* infection could protect against asthma. One study of Japanese children found that strong tuberculin responses, probably attributable to *M. tuberculosis* exposure, were associated with less asthma, rhinoconjunctivitis, and eczema in later childhood (Shirakawa et al. 1997). In positive tuberculin responders the rate of current atopic symptoms was one-third the rate in negatives, and asthmatic symptoms were one-half to one-third as likely. On top of this, remission of atopy in children aged 7–12 years was six to nine times as likely in positive tuberculin responders. A study of South African children found an inverse association between *M. tuberculosis* infection and atopic rhinitis (Obihara et al. 2005). Other comparisons among countries have found that asthma tends to be more common where TB is not (von Mutius et al. 2000; Shirtcliffe 2002). The debatable implication of these results is that TB has been inhibiting the spread of asthma and other atopic disorders worldwide. Taking this one step further, a mycobacterial vaccine might be constructed to prevent atopy and asthma. BCG could already serve that purpose, though the evidence is ambiguous (Hopkin 2000). *M. tuberculosis* infection may also protect against leprosy, as does BCG (Karonga Prevention Trial Group 1996) and natural TB transmission could have contributed to the decline of leprosy in Europe (Lietman et al. 1997).

As the evidence for an immunological link between TB and asthma becomes more compelling, interactions between other infections have come under investigation. For instance, vigorous Th2 responses are seen in protective immune

reactions to helminth infections, and helminths could modulate atopic disease while compromising the immune response to BCG and *M. tuberculosis* (Hopkin 2000; Obihara et al. 2006; Ferreira et al. 2002; Franke et al. 2013).

The complexity of interactions between *M. tuberculosis* and individual immunity presents a set of fascinating and important problems in within-host population biology. Just one of the many relevant questions is how mycobacteria can survive for years within hosts in small but viable populations, neither being eliminated by the host nor being allowed to break out of granulomas to cause destructive and progressive disease. One theory proposes that an equilibrium can be maintained by a combination of intra- and extracellular populations of bacilli (Blaser and Kirschner 2007). As the ecology of bacilli, macrophages, T cells, and other cell species are further explored, a key question is whether the systems biology of within-host dynamics can offer insights into the between-host dynamics of TB in populations (Chapter 8; Young et al. 2008).

The synergistic and antagonistic interactions between bacterial, viral, and parasitic infections, mediated by immunity, are complex and unresolved. Nonetheless, the preceding examples at least raise the possibility that mycobacteria influence and are influenced by other infections to a far greater extent than hitherto appreciated, with potentially important effects at population level. Conceivably but speculatively, the manipulation of Th1 and Th2 responses has potential for TB immunotherapy.

INFECTIOUSNESS AND DISEASE IN HUMAN POPULATIONS

Replicated in many individuals, the transmission-infection-disease process yields the essential characteristics of tuberculosis in populations. These characteristics are outline here, as a preface to Chapters 2 and 3.

The key features are: a low rate at which cases transmit infection (typically < 10 infections/case/year; van Leth et al. 2008; Styblo 1991; Bourdin Trunz et al. 2006) and a low rate at which others acquire infection (nowadays mostly < 2%/year); durable subclinical or latent infection that creates a large reservoir of infection (commonly more than a third of people in highly endemic areas have detectable cell-mediated immune responses, indicating past exposure to infection, if not the continuing presence of live bacilli); a low proportion of infected people progressing to active disease; and a spectrum of more or less infectious disease states. Both subclinical infection and clinical disease can last from weeks to years.

In the absence of other predisposing conditions, 5%–25% of infected people develop progressive primary disease within 5 years of infection (Sutherland 1968; Comstock et al. 1974a; Sutherland 1976). After 5 years, the annual risk

of developing TB by the reactivation of latent infection is much lower (10–100 per 100,000 persons per year; Dye et al. 1998; Horsburgh et al. 2010; Shea et al. 2013). The risk of progressing from infection to active disease is relatively high in infancy, declines as the immune system matures (Shey et al. 2014), and is markedly lower in children 5–15 years old. The risk increases quickly again during adolescence, earlier in girls, and then more slowly for both sexes throughout adulthood (Comstock et al. 1974a; Sutherland et al. 1982; Vynnycky and Fine 1997b). The survival rate of subclinical infections is unknown, but, for some infected individuals, the risk of reactivation persists into old age. In short, a minority of people develop TB following infection; there is strong resistance to developing active disease.

The cell-mediated response to infection is a marker of an innate or acquired immunity that is only partially protective (Sutherland et al. 1982; Vynnycky and Fine 1997b; Dye et al. 1998; Andrews et al. 2012). Consequently, individuals who carry a latent infection and who continue to be exposed are at risk of TB following reinfection. Based on observations of strain phenotypes (notably drug resistance) and genotypes, numerous studies have demonstrated that TB patients can carry multiple strains, presumably from several infection events (Richardson et al. 2002; van Rie et al. 2005; Baldeviano-Vidalon et al. 2005; Warren et al. 2004) and that recurrent episodes of TB may be due either to the reactivation of old infections or to reinfection with new ones, especially among people infected with HIV (de Viedma et al. 2002; Van Rie et al. 1999; Crampin et al. 2010; Chiang and Riley 2005).

In the context of TB control, the partially protective immune response helps explain why developing an effective TB vaccine has been difficult—few manufactured vaccines against infectious agents are more protective than natural immunity. In the context of an age-old host-pathogen relationship, it is questionable how and why long-lasting latency (bacilli viable for decades) and partial immunity (maintaining the supply of infectible individuals) have evolved as *M. tuberculosis* life history strategies, a topic discussed further in Chapter 2.

In the most infectious cases, especially those with cavitary disease, bacilli can be seen under a microscope on stained sputum smears (60%–70% of pulmonary cases; Styblo 1991). These cases constitute the bulk of the "infector pool" (Canetti 1962), though, within that pool, patients that produce cough aerosols are the most infectious (Jones-Lopez et al. 2013). TB cases with negative sputum smears can be identified by more sensitive culture methods. Culture-positive but smear-negative cases are also capable of infecting others. Each case contributes relatively little to transmission, but whole populations of smear-negative TB cases have been found to account for 15%–40% of all transmission events (Behr et al. 1999; Hernandez-Garduno et al. 2004). That percentage will vary, of

course, according to the proportion of cases found to be smear negative by the microscopic procedures used in a specific population under study; however, it underlines the point that culture-positive, smear-negative cases do contribute to transmission. The proportion of TB cases that develops extrapulmonary disease varies from one population to another but is typically 10%–30%. Extrapulmonary TB is more common among women and children (particularly lymphatic TB) than adult men and in people with immunosuppressive conditions, notably HIV infection (Shafer and Edlin 1996; Rieder et al. 1990; Rieder 1999; Aaron et al. 2004).

Across a whole lifetime the risk of developing TB following infection depends on the rate of progression from infection to disease, the life expectancy from the time of infection, and the chance of reinfection. Acknowledging much variation, a rule of thumb is 10% on average, and it has been calculated at 12% for all forms of pulmonary disease in England and Wales during the second half of the twentieth century (Vynnycky and Fine 2000). The combination of low rates of infection and progression to active disease yields low incidence rates, which are rarely higher than 1%/person/year (1000 per 100,000 population) and more typically are around 0.1%/person/year (100 per 100,000 population).

Compared with other infections, such as influenza, these are low incidence rates. Thus TB is a major cause of ill health around the world, not because there are large numbers of cases but because a high proportion are fatal. More than half of all untreated smear-positive patients will die within 5 years (Berg 1939; Thompson 1943; Buhl and Nyboe 1967; Styblo 1991; Rieder 1999) and around 70% after 10 years (Tiemersma et al. 2011). Most of those who are still alive after 8 years will have quiescent TB—in a state of temporary self-cure, susceptible to relapse—and a few will become chronic excretors of bacilli. The proportion of untreated smear-negative cases that eventually die is lower because they suffer less extensive lung damage but is still of the order of 20% (Tiemersma et al. 2011; Krebs 1930). Even with good adherence to the most efficacious combinations of drugs, 5%–10% of all patients will die as a result of an episode of TB.

ORIGINS AND DISPERSAL OF *M. TUBERCULOSIS*

The present-day geographical distribution of genotypes is being used to reconstruct the origins and dispersal of human TB. When variable genetic markers are used to build mycobacterial family trees (phylogenies), the weight of evidence suggests that the common ancestor of extant *M. tuberculosis* strains was a human pathogen that originated in Africa (Brosch et al. 2002; Gutierrez et al. 2005; Comas et al. 2010; Smith et al. 2009; Comas et al. 2013). The
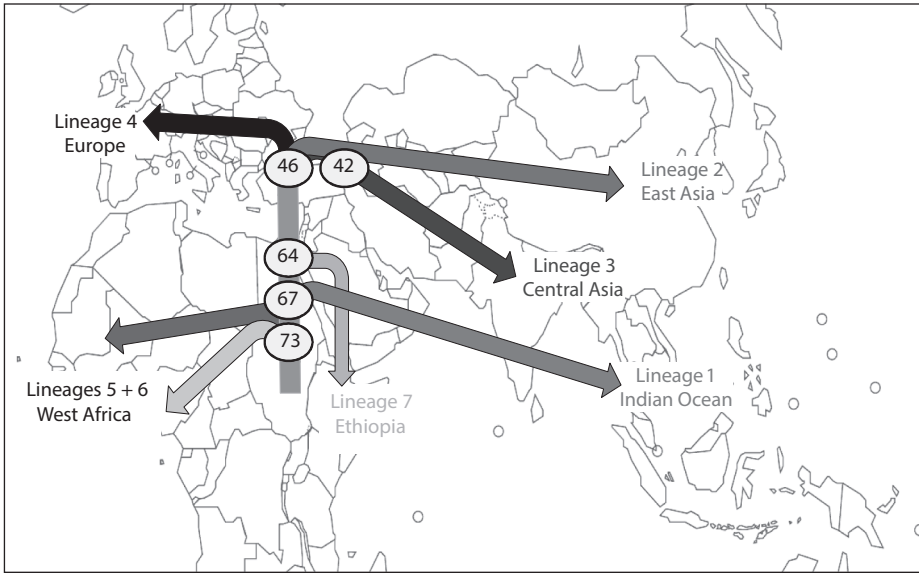
FIGURE 1.1. The "Out-of-Africa" hypothesis for TB, showing the phylogenetic branching points and geographical spread of seven lineages of the *M. tuberculosis* complex. At the branching points are marked the estimated date of the split (thousands of years ago). Adapted from Comas et al. (2013).

African continent supports the largest genetic diversity of *M. tuberculosis* and is the only continent on which all seven human-adapted lineages occur (Figure 1.1). Three of these seven lineages (5, 6, and 7) are unique to Africa. Two *M. africanum* lineages, 5 and 6, of the *M. tuberculosis* complex (de Jong et al. 2010) from West Africa are positioned close to the ancestral roots of phylogenetic trees, as compared with the probably ancient "smooth" variant *M. canettii*. *M. canettii* and other ancestral *M. tuberculosis* strains appear to be direct descendants of tubercle bacilli that existed before *M. tuberculosis sensu stricto* separated from *M. africanum, M. bovis*, and other members of the complex (Brosch et al. 2002). Lineage 7, found only in people originating from the Horn of Africa, apparently diverged a little later than lineages 5 and 6. Thus the common ancestor of the tubercle bacilli probably resembled *M. tuberculosis* or *M. canettii,* is likely to have been a pathogen of humans rather than of animals, coexisted with early hominids in East Africa, and has been coevolving with humans for hundreds of thousands or even millions of years (Brosch et al. 2002; Gutierrez et al. 2005).

The contemporary global distribution of *M. tuberculosis* strains can be linked to human migration out of Africa as well as to more recent movements

that followed the increases of human populations in Europe, India, China, and the Americas during the past few hundred years. Among various interpretations of these patterns (Filliol et al. 2006), Gagneux and colleagues (Comas et al. 2010; Hershberg et al. 2008; Comas et al. 2013) have proposed an "Out of Africa" phylogeography of the *M. tuberculosis* complex, which parallels that in humans. The earliest spread of modern humans carrying *M. tuberculosis* out of Africa was across the Indian Ocean to south-east Asia (lineage 1; Rasmussen et al. 2011). The more modern lineages 2 and 3 spread to south and east Asia and to Europe. Then, over the past few centuries of migration through trade and conquest, these three clades further spread to other areas of the world. The Euro-American strains (lineage 4) perhaps arrived on the American continent during the nineteenth-century exodus from overpopulated European cities.

   M. tuberculosis migrated not only out of Africa but also back again (Hershberg et al. 2008). Strains from India found in East Africa today can be explained in terms of the recent history of migration to this region from the Indian subcontinent. East Asian/Beijing strains (lineage 2) were probably imported to South Africa with slaves from South East Asia by Dutch colonialists and later by Chinese migrant workers employed in gold mines. Migrations account for the spread of modern Beijing, Haarlem, and other strains, not just to Africa but also to other parts of the world (Merker et al. 2015).

   However, phylogenetics does not give the complete picture of the origins and dispersal of *M. tuberculosis*. Archaeological evidence from skeletons (especially with the spinal lesions characteristic of Pott's disease) and from written records complements the inferences made from molecular epidemiology. The earliest evidence of destructive lesions of the spine is from Neolithic Europe (Italy, 5800 BC; Poland, 5000 BC) and the Middle East (Egypt, 4500 BC, Jordan, 3150–2200 BC; Roberts and Buikstra 2003, 2008; Stone et al. 2009). The first written description from China of glandular and pulmonary disease (with bloody sputum) is dated 2700 BC. Tuberculosis was certainly a disease of American peoples in pre-Colombian times, with skeletal evidence from AD 700 in South America and AD 900 in North America (Buikstra 1999; absent from the scheme in Figure 1.1). Unless *M. tuberculosis* had zoonotic origins in the Americas (Salo et al. 1994; Bos et al. 2014), it seems that TB colonized the New World with the earliest human migrants from Asia.

HUMAN TUBERCULOSIS SINCE 1600

Molecular genetics plus archaeology together present a compelling account of the evolutionary and cultural history of human tuberculosis. These are the only tools available to reconstruct events before the sixteenth century. But the

best way to track epidemics is to make direct observations on cases and deaths through space and time.

The rise of empiricism in sixteenth century England was a precursor to the Age of Reason. The first significant attempt to record deaths was made in the London Bills of Mortality, established in the plague year 1532. By the early 1600s the Bills were systematically reporting deaths by cause, albeit with uncertain accuracy, including the collection of conditions that would, by about 1840, be called "tuberculosis." In the 1600s, phthisis and consumption (lung and wasting disease) were the most important manifestations of TB. Among others were lupus vulgaris (skin), scrofula and the King's Evil (lymph nodes), and gibbus or Potts disease (spine).

John Graunt's 1632 analysis of the Bills found that 19% (1797/9535) of London deaths were from consumption, the largest cause of mortality after death in infancy (Graunt 1937). This percentage was consistent with later analyses of the Bills[3] in which, for example, consumption accounted for 17%–25% of deaths between 1675 and 1824; that is, one in four deaths were due to TB by the early nineteenth century (Landers 1993). Beyond London, other analyses showed how deaths from consumption varied in different parts of Britain and elsewhere (Young 1815). Dubos and Dubos (1987) emphasized the link between TB, urbanization, and industrialization, but TB was not exclusively or even predominantly urban everywhere. In Ireland in the 1840s, at least 7 out of every 10 consumption deaths occurred in the rural population, though the death rate per capita was higher in urban areas (Breathnach and Moynihan 2011).

Like other infectious diseases, these medical conditions linked to TB and the essentials of their epidemiology were known long before microbiology uncovered the cause. It was not until 1865 that Villemin demonstrated TB to be an infectious disease by inoculating laboratory rabbits with material from infected humans and cattle (Villemin 1868) and not until 1882 that Koch identified the bacterium (Koch 1882).

By the time the name tuberculosis was in common use, in the mid-nineteenth century, death rates from tuberculosis in England and Wales were already falling (McKeown 1976). The TB epidemic in England, Wales, and Scotland had risen and began to fall earlier than in Ireland and Scandinavian countries and much earlier than Japan and Puerto Rico (Figure 1.2A). The data from some European cities reinforced the view that TB began to decline in the United Kingdom (London) before continental Europe (Hamburg, Stockholm; Figure 1.2B; Grigg 1958b). The processes behind the rise and fall of TB epidemics—which elapse over hundreds of years—are investigated quantitatively in Chapter 2.

---

[3] Regrettably, the Bills reported causes of death only as proportions and not per capita.
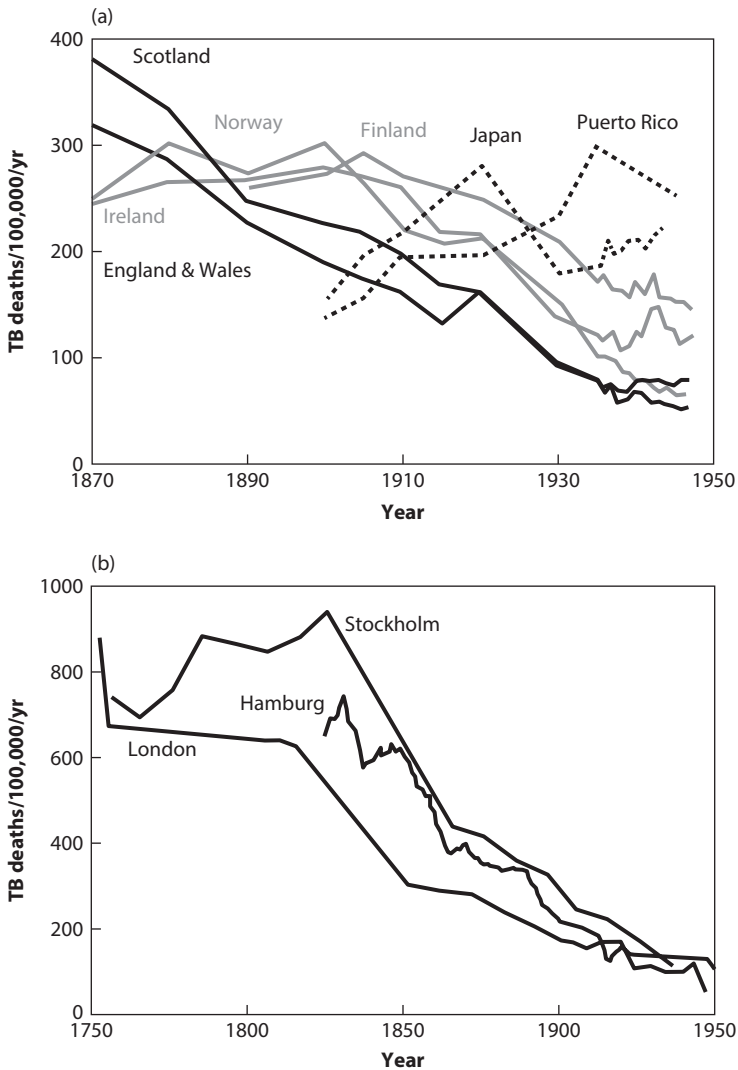
FIGURE 1.2. Tuberculosis epidemics, depicted by records of TB deaths, peaked at different times in different parts of the world.

    A. Countries. Epidemics peaked first in England and Wales and Scotland, followed by Ireland and Scandinavia. TB deaths reached their highest levels in other countries (Japan, Puerto Rico) only during the twentieth century. Data compiled by Dubos and Dubos (1987).

    B. Cities. Epidemics peaked first in London and then on continental Europe. Drawn from data in Grigg (1958).
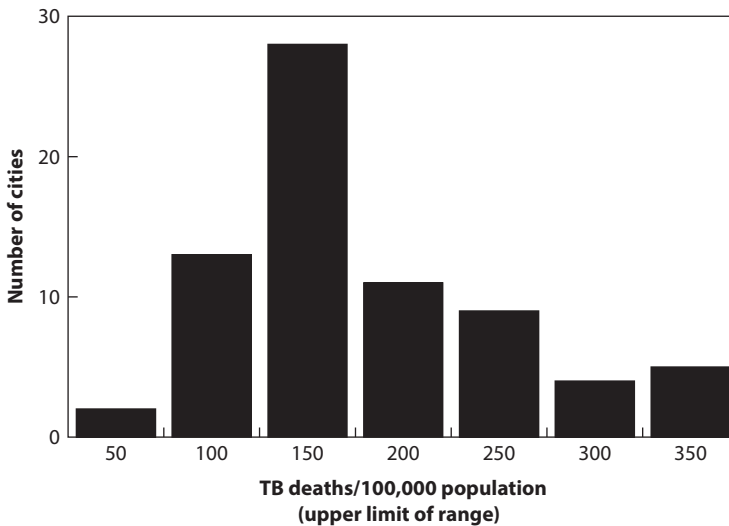
FIGURE 1.3. Frequency distribution of reported TB mortality rates in 72 cities worldwide in 1912. Data from Cliff et al (1998).

There is one further characteristic of TB made plain in the historical data from cities. By contrast with many other infectious diseases, the death rates are remarkably similar from one place to another. Among the reported death rates from TB in 72 cities around the world in 1912, 52 had rates in the range 50–200/100,000 population (Figure 1.3; compare also Figure 1.2; Cliff et al. 1998). This raises the question of why the mortality rates were so invariable. What population processes kept death rates within such a narrow band? Furthermore, once TB deaths had begun to decline in Europe and the United States, the rates of decline were similar throughout the first half of the twentieth century (although interrupted by World War I; Figure 1.4). Evidently common underlying processes were at work in Western Europe and North America. But was there a single dominant process that governed the rate of decline of TB? Or just a few, or many? This, too, is a question to be addressed quantitatively in Chapter 2.

TB became a notifiable disease in the UK in 1912, providing fuller epidemiological insights from case and death records. The UK series of cases and deaths (Figures 1.5), supported by tuberculin skin test surveys (Figure 1.6), tell a story that was replicated elsewhere in Western Europe: the rate of decline in incidence and deaths rates was 4%–5% per year up to 1950. With the advent of drug treatment in the 1940s—and especially after 1950—the decline in case incidence accelerated to around to 10% per year. The fall in the mortality rate
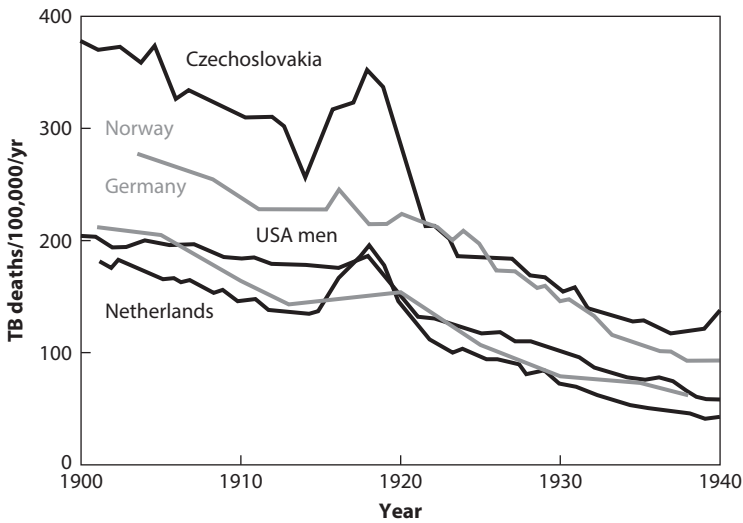
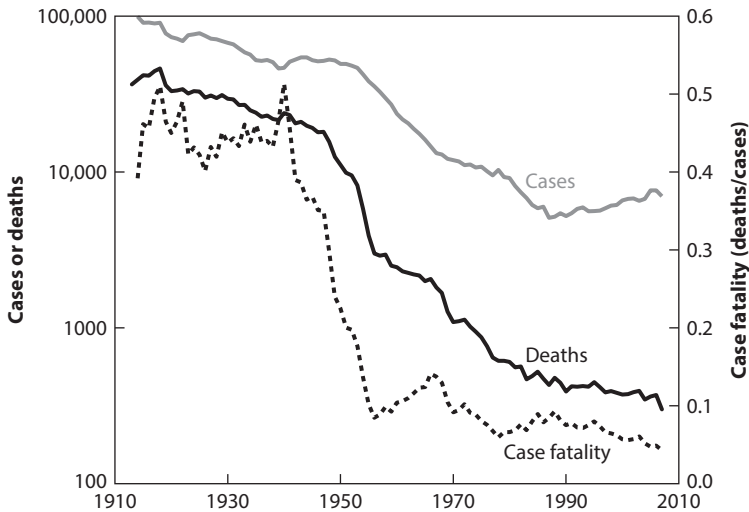FIGURE 1.4. Trends in TB deaths for selected countries, 1900–1940. Data from Styblo (1991).

FIGURE 1.5. Decline in TB cases and deaths in England and Wales from 1912. Case fatality remained at 40%–50% until TB drugs became available during the 1940s, fell sharply into the 1950s and fell more slowly thereafter. Data from UK Health Protection Agency.
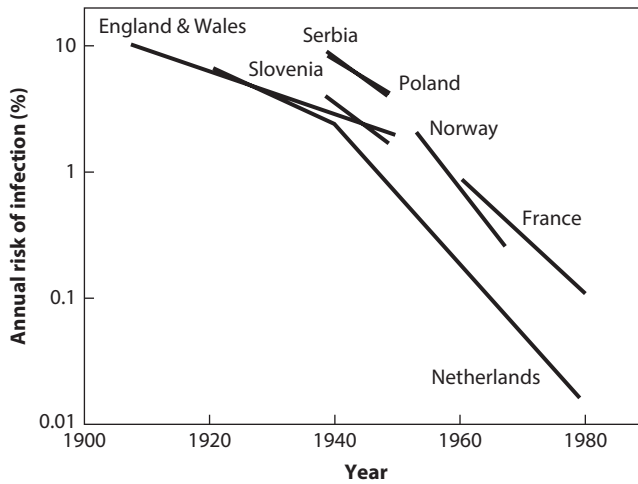
FIGURE 1.6. Trends in the annual risk of infection in European countries, 1900–1980. The rates of decline vary from 4%/year (England and Wales 1907–1950) to 13%/year (Netherlands 1939–1979). Data compiled by Rieder (1999).

was faster still; there were fewer deaths because there were fewer cases and because each incident case had a smaller chance of dying (reduced case fatality, Figure 1.5; Vynnycky and Fine 1997a; Styblo et al. 1969).

In the first half of the twentieth century, TB was closely observed during its slow decline in industrializing nations but remained largely invisible in the less-developed world. A few specific studies drew together, for example, the anecdotes that charted the spread of TB in sub-Saharan Africa. Although *M. tuberculosis* probably had its origins in Africa, the disease remained at low prevalence until European colonization and penetration beyond coastal areas (Roelsgaard et al. 1964). In a rare, documented example of the invasion of TB, Davies charted the rise in incidence as the mining industry developed in Rhodesia from the early 1900s up to 1964 (Figure 1.7; Davies 1966).

The principal diagnostic methods—tuberculin skin tests, sputum smear microscopy, bacterial culture, and chest X-ray—were brought together in systematic studies of TB epidemiology from the 1940s onward, notably in Africa (Roelsgaard et al. 1964) and India (Chakraborty 1996). However, the most abundant data are from tuberculin skin test surveys, which were used to measure and compare the annual risk of infection (ARI, the incidence of infection per person per year) around the world. ARI data from the late 1940s to the early 1980s were compiled by Cauthen et al. (1988, reprinted in 2002) and later summarized
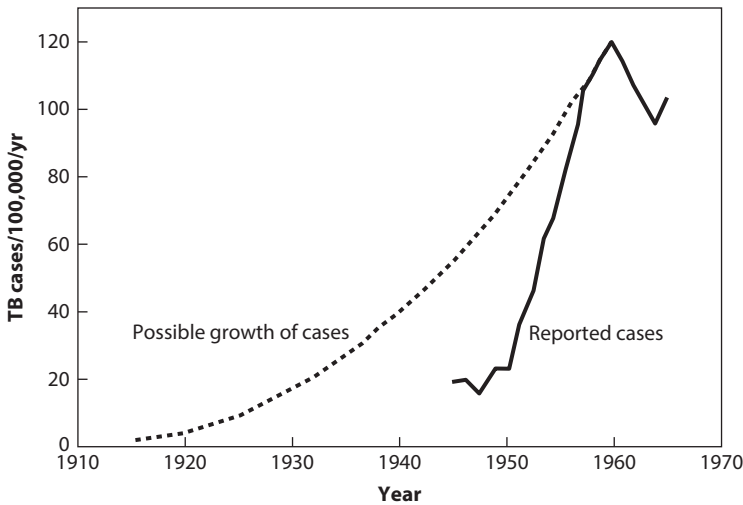
FIGURE 1.7. A rare, documented example of the growth of a TB epidemic: The postulated and observed increase in TB cases per 100,000 population in Rhodesia (now Zimbabwe) as reported by Davies (1966).

by Kochi (1991). In that assessment ARI was falling at more than 10%/year in industrialized countries, 5%–10%/year in North Africa and the Middle East, and more slowly or not at all in East and South Asia (Kochi 1991). There are some differences between then and now—for instance, the accelerated decline of TB in East Asia, especially China—but much that is broadly the same.

The cheap and rapid mapping of *M. tuberculosis* genetic variation provides another powerful tool for epidemiological investigation on large and small spatial scales. In a study of historical epidemiology, DNA-VNTR data showed that the population of Quebec in Canada was the likely source of *M. tuberculosis* infecting Aboriginal people when they made contact during the heyday of the fur trade, 1710–1870 (Pepperell et al. 2011). Genetic markers have been used to track contemporary outbreaks of isoniazid resistant strains in the United Kingdom (Maguire et al. 2011) and Sweden (Sandegren et al. 2011). Coupled with traditional outbreak investigations, whole genome sequencing (WGS) revealed that the circulation of two distinct lineages of *M. tuberculosis* in British Colombia, Canada, centered on crack-cocaine users (Gardy et al. 2011). WGS studies of TB outbreaks in Birmingham, Leicester, and Oxford in the United Kingdom have identified superspreaders, distinguished imported from locally acquired infections, and allowed inferences about the direction of transmission between cases (Walker et al. 2013, 2014). WGS also helped to distinguish TB

arising from reinfection and relapse among patients enrolled in the REMoxTB trial (of 4-month treatment regimens including moxifloxacin) carried out in Malaysia, South Africa, and Thailand (Bryant et al. 2013).

TB AND THE "GRAND CONVERGENCE" IN HEALTH?

Since the 1980s, the evaluation of TB burden and trends has been made by a combination of routine case and death notifications, population-based surveys of infection and disease prevalence and mortality, and deaths through vital registration (Chapter 2; Dye et al. 1999; Corbett et al. 2003b; World Health Organization 2013a). Every country in the world has a national TB control program—because every country still has TB—and TB is now monitored more closely than ever before.

Although the methods used by national TB control programs have been standardized and steadily improved from the 1990s onward, the data are imperfect and, therefore, need careful interpretation. A suite of estimation techniques has been devised for that purpose (Dye et al. 2008a; World Health Organization 2013a). Based on all the evidence, about one-third of people worldwide have been exposed to infection with *M. tuberculosis*, of which an unknown fraction are still be carrying viable infections (Sudre et al. 1991; Dye et al. 1999). The number of people exposed to infection is an approximate estimate of a quantity that will change slowly over decades.

As stated above, number of people who develop TB disease is far fewer than the number who acquire infections. There were an estimated 8.6 million new TB cases in 2012 (122 per 100,000 population), of which 2.9 million were women. WHO calculated that 530,000 children under 15 years of age developed TB, which is likely to be more accurate than an independent study that put the figure at approximately 1 million (Jenkins et al. 2014; Sismanidis et al. 2014).

The WHO African region (mainly sub-Saharan Africa) had by far the highest annual incidence rate (255/100 000 population), but the most populous countries of Asia carry the largest numbers of cases. The WHO South-East Asia and Western Pacific Regions together accounted for 58% of the global total. In those two regions, India, China, Indonesia, Bangladesh, and Pakistan accounted for about half (51%) of all new TB cases arising in 2012.

Among the 1.26 million TB deaths in 2012, the largest numbers were in South-East Asia (501,000) and Africa (480,000). An estimated 320,000 deaths were among people infected with HIV. The ratio of deaths to cases in South-East Asia gives an estimated case fatality rate of 16%, similar to the global average. In Africa, however, case fatality is higher (22%) because a relatively high

proportion of TB cases are infected with HIV (35%, compared with the global average of 13%). These estimates place TB as the second most important cause of death from a single infectious agent, behind HIV (approximately 2 million deaths in 2012), and the largest cause of death from a curable infectious disease.

Judging from trends in case notifications, the global TB epidemic is on the threshold of decline—but a slow decline (Figures 1.8 and 1.9). The incidence rate *per capita* was growing during the 1990s but stabilized during the decade 2000–2010 (peak 2003) and then began to fall. However, because the populations of the countries most heavily affected by TB are still growing, the total number of new TB cases arising each year continued increasing until the later years of the decade (peak 2008; Figure 1.9C).

This rather static picture of the global epidemic conceals much variation in the dynamics of TB among regions (Figures 1.8 and 1.9). While the burden of TB is carried predominantly by Asian countries, Africa and Eastern Europe have determined global trends. The countries of sub-Saharan Africa and the former Soviet Union showed the most striking increases in case load during the 1990s (Figures 1.8 and 1.9) due, respectively, to the spread of HIV in Africa (Corbett et al. 2006) and to the collapse of health and health care in the Soviet Union (Coker et al. 2005; Toungoussova et al. 2006). These rises offset the slow decline in case numbers in other parts of the world: West and Central Europe (4%/year), the Region of the Americas (2%/year), and the Eastern Mediterranean (< 1%/year), South-East Asia (<2%/year) and Western Pacific (>2%/year) regions. China (3%/year) is probably doing better than India (1%/year). High-income countries (mostly members of the Organization Economic Cooperation and Development, OECD) are typically seeing fewer cases among nationals each year but steady or rising numbers of cases among immigrants. In 2011, the majority of TB cases were foreign-born or foreign citizens in 12 of 28 European Union countries (Chapter 3; European Centre for Disease Control and Prevention/ World Health Organization Regional Office for Europe 2011).

Just as there is diversity among regions, there are also important differences among countries within regions. Over the past decade, few countries have achieved the 5%–10% annual rates of decline comparable to those seen in postwar (1950–1980) Western Europe (Figure 1.10). However, the countries with the fastest reductions in case-notification rates tend to be those that already have lower incidence rates per capita (Figure 1.11A). OECD countries with rates of decline faster than 5%/year include Austria, Germany, Japan, Slovenia, Slovakia, Switzerland, and the United States. Outside the OECD, only Costa Rica, Ecuador, Honduras, and Puerto Rico in the Americas, Estonia and Latvia in the European Union, and wealthy island nations such as French Polynesia, Maldives, and New Caledonia have rates of decline more than 5%/year. As
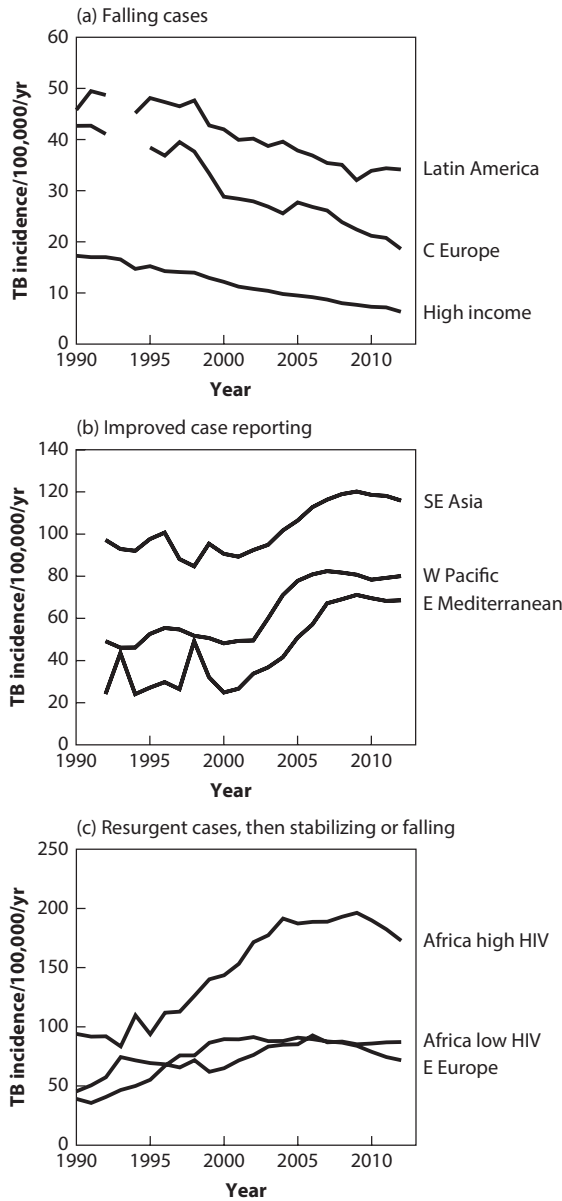
FIGURE 1.8. Trends in annual TB case reports (expressed as rates per 100,000 population) in nine epidemiologically distinct regions, 1990–2012.

A. Three regions showing moderate rates of decline of 2%–4%/year.

B. Three regions where TB is stable or in slow decline, but improvements in reporting have pushed up case notifications since 2000.

C. Three regions where TB incidence increased markedly during the 1990s but has since stabilized. Data from WHO (World Health Organization 2013a).
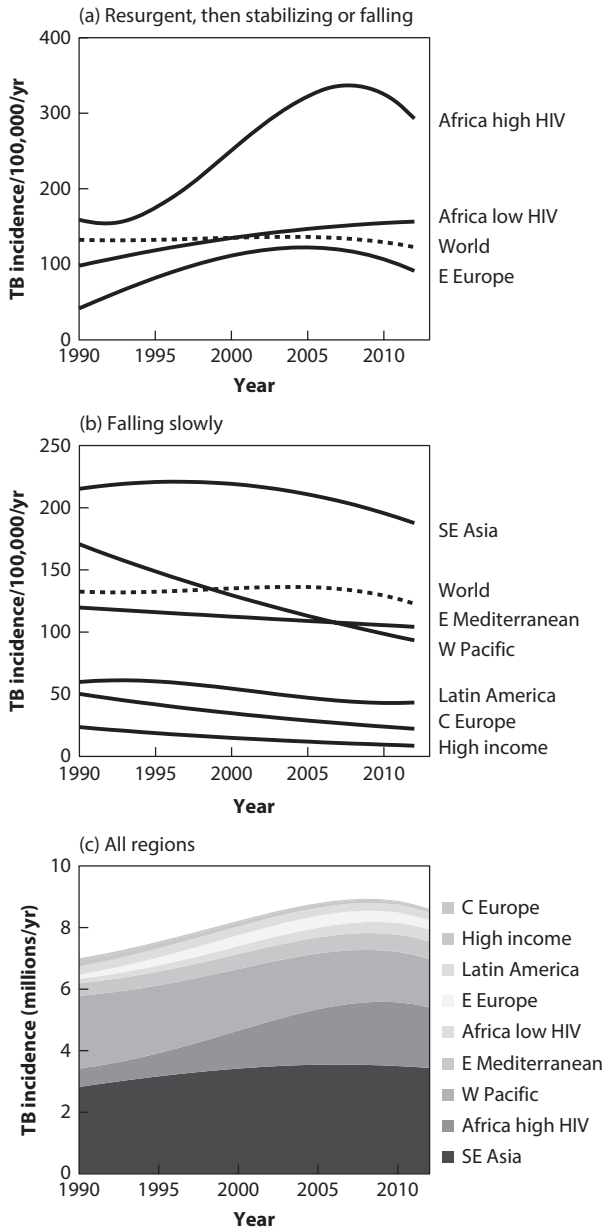
FIGURE 1.9. Estimated TB case incidence rates (A, B) and case numbers (C) for nine regions of the world.

  A. Regions with resurgent TB, then stabilizing (Africa, low HIV) or falling.

  B. Regions in which TB incidence rates have been falling more or less slowly.

  C. All regions summing to the global total.

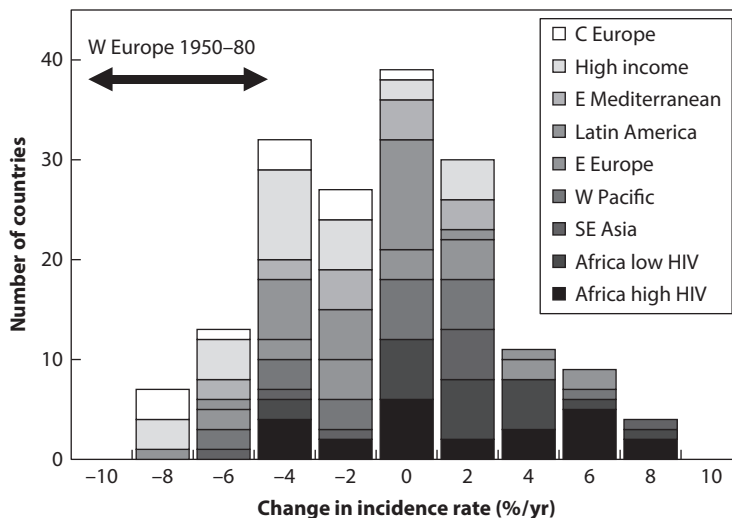Data from WHO (World Health Organization 2013a).

FIGURE 1.10. Trends in TB incidence rates, based on case reports for 172 countries
in 9 regions, 1996–2012. Over this period very few countries have achieved the
rates of decline in TB seen in Western Europe between 1950 and 1980. Calculated
from WHO data (World Health Organization 2013a).

international policymakers set a course toward a "grand convergence" in health
in a generation (Jamison et al. 2013), there is a weak tendency towards diver-
gence, rather than convergence, in TB trends between high- and low-incidence
countries (Figure 1.11A). With respect to mortality, there is no obvious trend
toward convergence or divergence, although low-incidence countries also tend
to have low case fatality rates (Figure 1.11B). To achieve convergence across
countries, toward uniformly low incidence and mortality rates, is a challenge
for the discovery, development, and delivery of better solutions for TB control
(Chapters 4–7).

Time trends in TB cases, both falling and rising, are also reflected in the dis-
tribution of cases with age (Figure 1.12). As transmission and incidence decline
and as a higher proportion of cases arise from older rather than recent infections,
the burden of disease shifts to older people. The transition is illustrated very well
by India, China, and Japan, which are, respectively, at early, mid, and late stages
of this epidemic transition (Figure 1.12A–C). However, this pattern of change
can be radically perturbed by other major events. In particular, the incidence
rate of TB has remained high among young adults because of the spread of HIV/
AIDS (South Africa, Figure 1.12D), because of TB's resurgence following the
breakdown of health and health services in former Soviet countries (Russia,

(a) Trends in case incidence (notifications)

Trend in case notification rate (%/yr)

Estimated case incidence (per 100,000/yr)

(b) Case fatality

Case fatality (proportion)
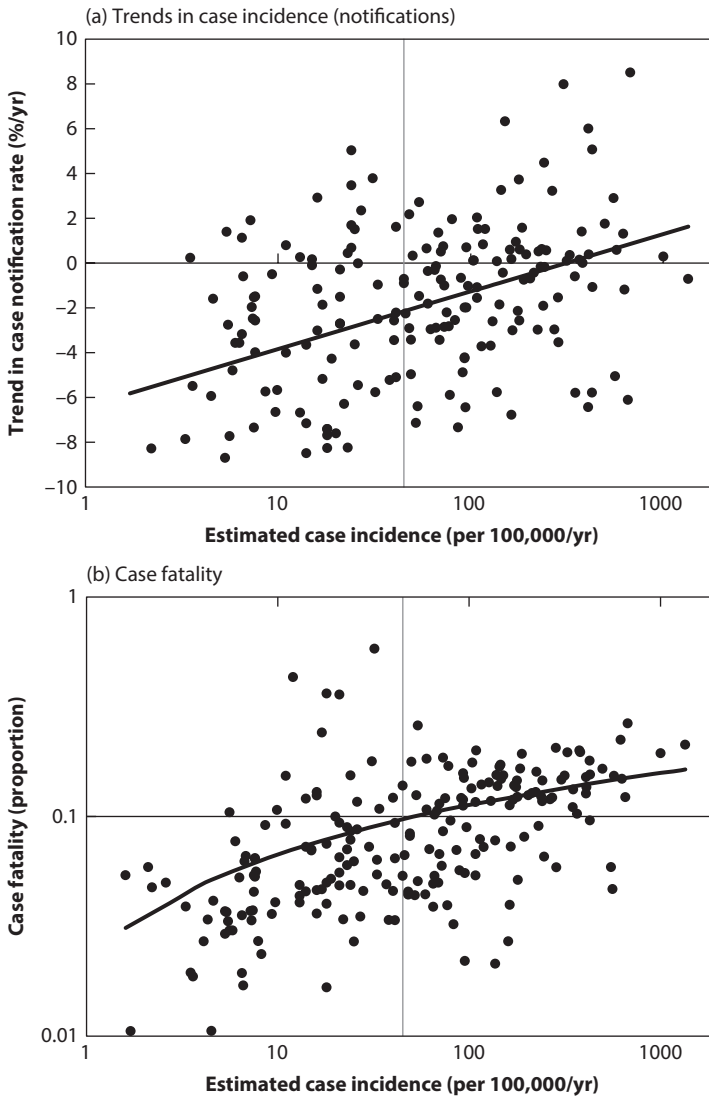
Estimated case incidence (per 100,000/yr)

FIGURE 1.11.

A. Trends in case notification rates, 2003–2012, in relation to estimated case incidence rates in 2012 for 188 countries. The rate of decline tends to be faster (more negative) where estimated incidence is lower ($r = 0.33$), suggesting that trends in incidence are divergent rather than convergent.

B. Case fatality in relation to estimated case incidence rates in 2012 for 209 countries. Case fatality tends to be lower where estimated incidence is lower ($r = 0.40$), presenting a further challenge for convergence. Lines in A and B are log-linear regressions. Data from World Health Organization (2013a).
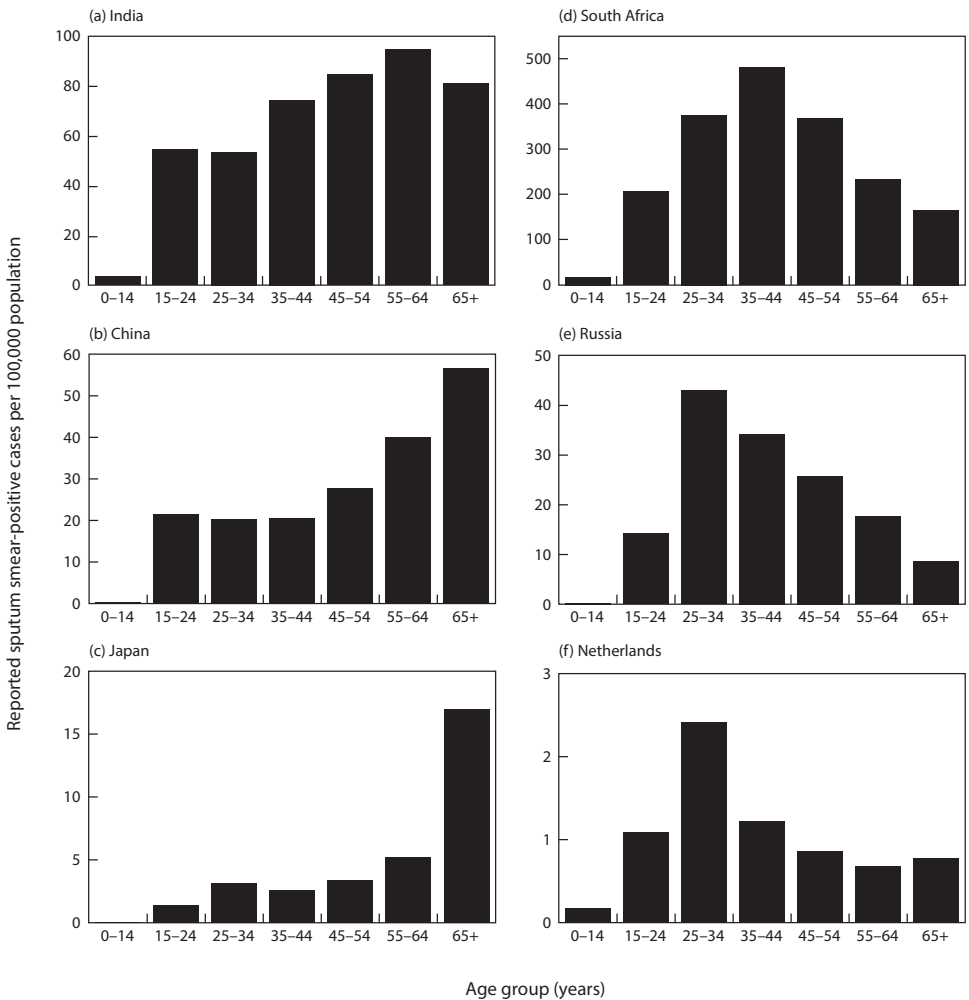
FIGURE 1.12. Reported sputum smear-positive TB cases per 100,000 population in 2012 (*y*-axis) in relation to age (*x*-axis). As TB transmission and incidence decline, the burden of disease shifts from younger to older people. (A) India, (B) China and (C) Japan are, respectively, at early, mid, and late stages of this epidemic transition. Perturbations to this general pattern of change have been caused by, for example, (D) the spread of HIV/AIDS in South Africa, (E) the resurgence of TB in post-Soviet Russia, and (F) immigration from high-incidence countries to the Netherlands. Calculated from WHO data (World Health Organization 2013a).

Figure 1.12E), and because of immigration from high-incidence to low-incidence countries (Netherlands, Figure 1.12F).

National and regional trends in TB deaths reflect the trends in cases. But deaths can change more quickly than cases because mortality depends greatly on drug treatment practices, whereas incidence depends on transmission and the underlying reservoir of infection. Despite the lack of medically certified death records in most highly endemic countries, it is clear that TB deaths increased in Africa and in ex-Soviet countries during the 1990s but were already falling in most other parts of the world. The spread of HIV, mainly in Africa, pushed the risk of death among TB patients (case fatality rate) up to a peak of 20% around year 2000 (Chapter 6). Since the turn of the century, improvements in treatment for both HIV-infected and uninfected cases brought the fatality rate back down to 16% in 2010, and TB deaths are now falling faster than TB cases (World Health Organization 2013a).

More than 50 years after Crofton's lecture, the rise of tuberculosis might at last have been halted. But in terms of case incidence, TB is barely in retreat and certainly very far from defeat. To get a better understanding of why this is so, Chapter 2 presents TB epidemiology from the perspective of population biology, bringing together the key concepts in a quantitative framework that will be used for further investigation in Chapters 3–7.