CHAPTER 2

Long-term Trends in Population Health

That the people's health has improved importantly over the 300 years since the early 18th century is well-known, but how has the pace of improvement varied over time, across health conditions and between European countries? In this chapter we will give a broad-brush picture of these changes and variations, in order to have firm ground under our feet when we review the more detailed disease-specific trends.

Changes in Over-all Population Health

The best way to start our analysis is by looking at trends in mortality. This may sound illogical: how can counting the dead inform us about the health of the living? However, the reason is simple: death usually follows disease or another health condition, and mortality rates therefore often tell us something about the frequency of disease. Furthermore, mortality data are the only health data going back to the 18th century, and are also available for a far larger number of countries than morbidity data.

Declining Mortality

The increase in life expectancy that we saw in Figure 1 is, of course, due to a decline in the mortality rate. When life expectancy at birth was below 40 years, as it was in the 18th century, the annual death rate was around 30 per 1000. This is equal to saying that each year 3 out of every 100 European citizens died. Now that life expectancy has nearly doubled, the annual death rate is around 10 per 1000. (Because of population aging, the real decline in the risk of death has actually been much greater than the two-thirds suggested by this simple comparison of crude mortality rates.) Figure 2 shows that this decline of the death rate has actually been very bumpy.

Over the three centuries for which data are available, the European experience is punctuated by enormous spikes of mortality, which have only recently diminished in amplitude. Previous studies have suggested that the amplitude of these spikes gradually diminished during the 18th and 19th centuries, but this is not immediately obvious from our data which cover an unusually large number of countries. Previous studies focused on the trend of mortality in
countries like Sweden, France and England, and in these countries the swings in mortality did indeed diminish in an early stage, as can be seen for Sweden which has been highlighted in Figure 2.¹

Yet, for other countries mortality fluctuations during the 19th century are not necessarily less than during the 18th century, and some dramatic upswings of mortality even occurred in the first half of the 20th century. When all countries are taken together there was a gradual decline of the average amplitude of mortality fluctuations, but this decline reversed in the first half of the 20th century.

¹ Alfred Perrenoud, “The Attenuation of Mortality Crises and the Decline of Mortality,” in The Decline of Mortality in Europe, ed. Roger S. Schofield, David Reger, and Alain Bideau (Oxford: Clarendon Press, 1991).
century, and it was only in the second half of the 20th century that the trends really calmed down.\(^2\)

What were the causes of these mortality spikes? The highest peak in Figure 2 is for Ukraine in 1933, when more than 8% of the population died in a famine. Many other spikes were also caused by large famines, such as the 1758 peak in Iceland, the 1773 peak in Sweden, the 1868 peak in Finland, the 1933 peak in Russia, and the 1942 peak in Greece. The more widely known famine of the 1840s (due to the ‘Potato Blight’, see Chapter 4) is invisible. This is partly because it is hidden below other, more dramatic mortality elevations, partly because there are no national mortality data from the 19th century for Ireland, where this famine caused a very strong increase in mortality rates.\(^3\)

Large-scale epidemics are another important cause of mortality spikes. Smallpox epidemics caused cycles of mortality fluctuations on a time-scale of 5 – 15 years in many European countries. The 1918/19 influenza pandemic can also clearly be seen in Figure 2. Of all the epidemics in the 19th century, cholera is by far the most publicized, but the excess mortality involved is generally less than that of other mortality ‘crises’. The main exception is the 1873 cholera epidemic in Hungary which caused a huge mortality peak.\(^4\)

The spikes caused by famines and epidemics are more prominently visible in Figure 2 than spikes in mortality caused by wars, but this is partly due to interruption of data collection in war-time. Nevertheless, the Franco-Prussian War (1870–71) coincided with a visible spike in France’s mortality trend, and so did World War I (1914–1918) and World War II (1939–1945) in the mortality trends of the countries involved that kept their vital registration systems running.

As can be seen in Figure 2, not only have mortality spikes gradually diminished in amplitude, but ‘normal’ levels of mortality have declined as well. Around 1870, European countries’ annual death rate in ‘normal’ years still ranged between 20 and 35 per 1000. This rapidly declined to between 10 and 20

---

\(^2\) The average coefficient of variation (i.e., the average of all available countries' standard deviation of their yearly mortality rates, as a fraction of their average mortality rate over the same period) was 22% in 1720–1749, 15% in 1750–1799, 12% in 1800–1849, 11% in 1850–1899, 15% in 1900–1949, and 7% in 1950–1999.

\(^3\) Cormac Ó Gráda, “Ireland,” in *Famine in European History*, ed. Guido Alfani and Cormac Ó Gráda (Cambridge etc.: Cambridge University Press, 2017).

\(^4\) Some of the highest spikes occurred in Iceland: a country with less than 100,000 inhabitants in which infection with diseases like small-pox and measles could not sustain, so that epidemics imported from abroad would subside, sometimes for decades, and then flare up again when the virus returned to infect a largely non-immunized population.
in 1930, and even further to between 7 and 13 in 1960. Many countries in Northern and Western Europe passed below the 15 per 1000 level in the 1900s or 1910s, whereas most countries in Southern, Central-eastern, South-eastern and Eastern Europe reached this low level only in the 1930s or 1940s.5 Because, during the ‘demographic transition’, mortality decline preceded fertility decline, European population numbers increased enormously. In the beginning of the 19th century, Europe had around 200 million inhabitants, and this number almost tripled to around 550 million in 1950, despite massive emigration and a series of bloody wars.

**Young and Old, Men and Women**

Mortality has declined at all ages, indicating that health improvements have occurred throughout the lifespan. The largest percentage declines in mortality occurred among infants and children, but even among the elderly reductions in mortality have been substantial. When we take all European countries together, and calculate average percentage declines of mortality between 1870 and 2010, we find the following astonishing figures.

Among men, mortality declined by 98% among infants, by 99% among children (ages 1 to 15 years), by 88% among young people (ages 15 to 45 years), by 67% among middle-aged people (ages 45 to 65 years), and by 55% among elderly people (ages 65 to 85 years). These declines are almost incredible, particularly in the youngest age-groups. Among women, the corresponding figures are 98%, 99%, 94%, 81%, and 71%, indicating that with the exception of infants and children the mortality declines among women have been even greater than among men.6

As always, these over-all averages hide important variations (see Suppl. Figure 2). Among infants and children, mortality decline first accelerated, to reach its maximum speed in the period 1930–1960, and then slowed down again. This is probably because after 1960 mortality had reached such low levels, that it became increasingly difficult to lower it further. Among the elderly, however, mortality decline accelerated over time, and reached its maximum speed in the last period (1990–2010), suggesting that the end of mortality decline is not yet in sight.

5 After the early 1960s, population ageing caused the crude mortality rate to stabilize, and to even increase somewhat. Turkey’s mortality rates in this period, which stand out as very low in Figure 2, are explained by its very young population.

6 These percentage declines are based on data in the Human Mortality Database and the Human Lifetable Database, and were calculated from simple arithmetic averages of the age-adjusted mortality rates (in each of the age-groups mentioned) for 13 countries in 1870, and 25 countries in 2010.
For young and middle-aged women, the time-pattern of mortality decline is like that seen for infants and children: acceleration followed by deceleration. For young and middle-aged men the time-pattern is different, with a clear dip in the speed of mortality decline between 1960 and 1990. This caused a temporary interruption of the secular increase of male life expectancy at birth, which we already saw in Figure 1. It was due to a rise of ischaemic heart disease and other ‘diseases of affluence’ that we will discuss in more detail below.

Generally speaking, these time-patterns can be observed in all European countries, but the dips in the speed of mortality decline among young and middle-aged men in the 1960–1990 period have been much more prominent in some countries than in others. In Northern, Western, and Southern European countries this was just a deceleration of mortality decline, whereas in Central-eastern, South-eastern and Eastern Europe mortality among adult men actually increased between 1960 and 1990.

A striking acceleration of mortality decline in recent decades can also be seen among the oldest-old, i.e., those above 85 years of age. As a result of declining mortality remaining life expectancy at the age of 85 has recently increased rapidly. Until the early 1950s, remaining life expectancy from the age of 85 hovered between 3 and 4 years among men and between 3.5 and 4.5 years among women, but then it started to increase, gaining speed along the way (see Suppl. Figure 3).

In 2016, the highest average life expectancy from age 85 among women in Europe, 7.7 years, was found in France. France is also the country with the oldest living European (Plate 3), and the country where Jeanne Calment (1875–1997) lived. She holds the world record with a documented age-at-death of 122 years and 165 days.7

These trends in mortality at advanced ages are followed closely, not only by demographers but also by gerontologists, economists and many others. It was once thought that average life expectancy would not increase beyond age 85, and that age 85 even represented something like the average maximum lifespan of the human species. If this were true – so the reasoning went – not only

---

7 On Jeanne Calment, see Jean-Marie Robine and Michel Allard, “The Oldest Human,” Science 279, no. 5358 (1998): 1831–831. Recently raised doubts on the true age of Jeanne Calment have been refuted by Jean-Marie Robine et al., “The Real Facts Supporting Jeanne Calment as the Oldest Ever Human,” Journals of Gerontology: Series A 74, no. S1 (2019): S13–S20. In Europe, several small areas with an unusual concentration of centenarians have been found (e.g., on Sardinia). These are usually isolated communities that have retained a traditional life style, but that also benefit from modern hygiene, health care and living standards; see Michel Poulain, Anne Herm, and Gianni Pes, “Blue Zones: Areas of Exceptional Longevity around the World,” Vienna Yearbook of Population Research 11 (2013): 87–108.
would a ‘compression of mortality’ around this age of 85 occur, but a ‘compression of morbidity’ in a few years before death would also be inevitable.8

However, this is not what happened. The almost-doubling of remaining life expectancy from age 85, and more generally the rapid declines of mortality among the oldest-old, have belied the idea that human life expectancy is approaching a maximum. And even if there would be such a maximum, it is unlikely to be near: average life expectancy at birth among women in France, Spain, Italy, and Switzerland already exceeds 85 years. There is very little tendency for a ‘compression of mortality’ in a narrow age-range around the

---

8 This was famously argued in 1980 in James F. Fries, “Aging, Natural Death, and the Compression of Morbidity,” *New England Journal of Medicine* 303 (1980): 130–35.
average age-at-death. Whether or not a ‘compression of morbidity’ has occurred will be discussed below.\(^9\)

**Regional and Social Inequalities**

Below the surface of national levels and trends in mortality, there have always been substantial regional and social variations within countries, to the extent that some observers have claimed that for the history of mortality one cannot rely on the “misleading mean.” While that may be an exaggeration, it is certainly true that studies of subnational variations are important for our understanding of the causes of mortality decline.

In the long run, all geographical areas and all social groups within European countries have benefitted from health improvements, but timing and speed usually differed between areas and groups. This has not only created opportunities for identifying determinants of population health change, but has also often led to widening health gaps. French demographers Jacques Vallin and France Meslé have even suggested to use a divergence of health indicators between regions and social groups as an indicator of new epidemiologic transitions.\(^10\)

One of the most illuminating examples of these within-country differences are the disparities in health and health improvement between socioeconomic groups, as defined by level of education, occupational class, or income. The massive literature on this topic, which has grown explosively since these ‘health inequalities’ were ‘re-discovered’ in the early 1980s, indeed shows that

---

9 Some studies have found evidence for compression of mortality (e.g., John R. Wilmoth and Shiro Horiuchi, “Rectangularization Revisited: Variability of Age at Death within Human Populations,” *Demography* 36, no. 4 (1999): 475–95; Wilma J. Nusselder and Johan P. Mackenbach, “Rectangularization of the Survival Curve in the Netherlands, 1950–1992,” *Gerontologist* 36, no. 6 (1996): 773–82), but this has not been the general tendency in high-income countries since 1980 (Colin D. Mathers et al., “Causes of International Increases in Older Age Life Expectancy,” *Lancet* 385, no. 9967 (2015): 540–48). On “broken limits to life expectancy,” see Jim Oeppen and James W. Vaupel, “Broken Limits to Life Expectancy,” *Science* 296 (2002): 1029–31.

10 The expression “misleading mean” is from a paper by S. Ryan Johansson and Alice B. Kasakoff, “Mortality History and the Misleading Mean,” *Historical Methods* 33, no. 2 (2000): 56–8. Using divergence as a signal of new transitions was proposed in Jacques Vallin and France Meslé, “Convergences and Divergences in Mortality: A New Approach of Health Transition,” *Demographic Research* 2 (2004): 11–44. Regional variations in trends in mortality or other health indicators within European countries cannot be summarized in terms of one or a few generalized tendencies; see, e.g., Sebastian Klüsener et al., “Spatial Inequalities in Infant Survival at an Early Stage of the Longevity Revolution,” *Demographic Research* 30 (2014): 1849–864.
inequalities in mortality and other health indicators by socioeconomic position tend to widen during over-all health improvement.\textsuperscript{11}

Historical evidence suggests that socioeconomic inequalities in health are not a recent phenomenon. Inequalities in mortality between persons with higher and lower social positions were probably already present in the 17th and 18th centuries, at least in European urban areas. Yet, recent studies found that inequalities in infant mortality only emerged, when health started to improve at the end of the 19th century. This indicates that a higher socioeconomic position (i.e., more money or knowledge) only started to make a difference after the causes of high mortality had been discovered, and after effective interventions had been developed.\textsuperscript{12}

During the first half of the 20th century, only few countries kept registers of socioeconomic inequalities in mortality. In England and Wales, inequalities in infant mortality declined substantially between the 1920s and 1970s, suggesting that, after a temporary widening at the end of the 19th century, inequalities in infant mortality narrowed again. However, over the same period inequalities in adult mortality remained largely stable in absolute terms, and even increased in relative terms.\textsuperscript{13}

We are much better informed about trends in health inequalities over the last half-century. In most European countries, mortality has steadily declined among both lower and higher socioeconomic groups. Nevertheless, because

\textsuperscript{11} It was for some time believed that health inequalities had disappeared after World War II, due to higher living standards, social security and universal access to health care. In the early 1980s, the Black report in Britain (Douglas Black et al., \textit{Inequalities in Health (the Black Report)} (London Department of Health and Social Services, 1980; repr., Penguin Books, 1982)) and similar reports in other European countries shattered this illusion. For an analysis of socioeconomic inequalities in health in Europe, see Johan P. Mackenbach, \textit{Health Inequalities: Persistence and Change in European Welfare States} (Oxford etc.: Oxford University Press, 2019).

\textsuperscript{12} The empirical evidence on health inequalities before the 1970s is very fragmented; see Tommy Bengtsson and Frans van Poppel, “Socioeconomic Inequalities in Death from Past to Present: An Introduction,” \textit{Explorations in Economic History} 48, no. 3 (2011): 343–56. For studies finding emergence (or widening) of inequalities in mortality at the end of the 19th century in Sweden and England, see Tommy Bengtsson and Martin Dribe, “The Late Emergence of Socioeconomic Mortality Differentials,” \textit{Explorations in Economic History} 48, no. 3 (2011): 389–400; Robert Woods, \textit{The Demography of Victorian England and Wales} (Cambridge etc.: Cambridge University Press, 2000).

\textsuperscript{13} For analyses of long-term trends of inequalities in mortality in England and Wales, see Elsie R. Pamuk, “Social Class Inequality in Mortality from 1921 to 1972 in England and Wales,” \textit{Population Studies (Cambridge)} 39, no. 1 (1985): 17–33; Elsie R. Pamuk, “Social-Class Inequality in Infant Mortality in England and Wales from 1921 to 1980,” \textit{European Journal of Population/Revue Européenne de Démographie} 4, no. 1 (1988): 1–21.
percentage declines were usually faster among the higher educated and the higher occupational classes, relative inequalities have generally increased considerably, even in the most advanced welfare states. Inequalities in life expectancy between people with primary and tertiary education now amount to between 5 and 10 years in North-western and Southern Europe, and more than 10 years in Central-eastern and Eastern Europe.\textsuperscript{14}

A widening of relative inequalities in mortality is generally seen whenever mortality declines, regardless of country or cause of death. The same applies to several other health indicators, such as the incidence of cancer or injury. This results from the fact that higher socioeconomic groups have less difficulty following health advice (e.g., stop smoking), more often participate in prevention programs (e.g., cancer screening), and have better treatment outcomes (e.g., higher survival after myocardial infarction). Apparently, a higher socioeconomic position provides the individual with ‘flexible resources’, such as knowledge, money, power and social connections, which can be used to avoid disease or to minimize the consequences of disease.\textsuperscript{15}

While the trend description above applies to North-western and Southern Europe, the story is rather different for Central-Eastern and Eastern Europe. Even under communism there were socioeconomic inequalities in health, but after the fall of the Berlin Wall these health inequalities exploded. Whereas higher socioeconomic groups soon took advantage of liberalized economic conditions and new access to medical technologies from the West, with declining mortality as a result, mortality among the less well-off rose. This dramatic widening of the gap in mortality was only reversed when, in the early 2000s, mortality in lower socioeconomic groups finally started to decline.\textsuperscript{16}

\textsuperscript{14} For an overview, see Johan P. Mackenbach et al., “Socioeconomic Inequalities in Health in 22 European Countries,” \textit{New England Journal of Medicine} 358, no. 23 (2008): 2468–481. For the distinction between relative and absolute inequalities, see Johan P. Mackenbach et al., “Changes in Mortality Inequalities over Two Decades,” \textit{British Medical Journal} 353 (2016): i1732.

\textsuperscript{15} This paragraph is based on the ‘fundamental causes’ theory of health inequalities; see Bruce G. Link and Jo Phelan, “Social Conditions as Fundamental Causes of Disease,” \textit{Journal of Health and Social Behavior} Spec No (1995): 80–94. This theory finds considerable support in the European experience; see Johan P. Mackenbach et al., “Fundamental Causes of Inequalities in Mortality,” \textit{Sociology of Health and Illness} 39, no. 7 (2017): 1117–133.

\textsuperscript{16} On recent trends in mortality inequalities in Europe, see Johan P. Mackenbach et al., “Recent Trends in Health Inequalities in 27 European Countries,” \textit{Proceedings of the National Academy of Sciences} 115, no. 25 (2018): 6440–445.
Rising Height

Partly inspired by controversies about the role of living standards in the rise of life expectancy, many studies have been made of trends in body height – Europeans have grown considerably taller since the 19th century. American economist and Nobel laureate Robert Fogel (1926–2013) has used the positive correlation between people's adult height and their risk of mortality to argue that improvements in nutrition (as reflected in increasing body height) made an important contribution to increases in life expectancy.  

However, things are less simple than that. Height reflects growth, and growth reflects nutrition, but growth also reflects health during the first two decades of life, and is negatively affected by the extra nutritional demands during episodes of infectious diseases. Trends and variations in height thus capture the effect on growth of both nutrition and disease. This is reflected in the ambiguous term 'biological standard of living', which is commonly used in the economic-historical literature to denote what height actually measures.

For many European countries, data on average height of military recruits are available since the 19th century or even earlier (see Suppl. Figure 4). (In the case of compulsory medical service, these recruits will be men of ca. 20 years old.) During the 19th century height of military recruits fluctuated in a pattern suggesting that the early phase of industrialization was sometimes associated with a temporary decline in the 'biological standard of living'. For example, in England there was a decline in height from the 1820s to the 1840s, in which the average recruit lost a few centimetres.

---

17 Robert W. Fogel, The Escape from Hunger and Premature Death, 1700–2100 (Cambridge etc.: Cambridge University Press, 2004).
18 For an analysis along these lines, see Roderick Floud et al., The Changing Body (Cambridge etc.: Cambridge University Press, 2011). Fogel has also argued that better health contributed to accelerated economic growth; see Robert W. Fogel et al., "Secular Changes in American and British Stature and Nutrition," Journal of Interdisciplinary History 14, no. 2 (1983): 445–81.
19 Time-series data on height show that, at least in England, the early phases of industrialization had negative effects, partly through the side-effects of rapid urbanization; see Roderick Floud and Bernard Harris, "Health, Height and Welfare: Britain 1700–1980," in Health and Welfare During Industrialization, ed. R.H. Steckel and R. Floud (Chicago & London: Chicago University Press, 1997); John Komlos, "The Secular Trend in the Biological Standard of Living in the United Kingdom, 1730-1860," Economic History Review 46, no. 1 (1993): 115–44. This was, however, not observed in some other European countries, such as Sweden and the Netherlands; see Lars G. Sandberg and Richard H. Steckel, "Was Industrialization Hazardous to Your Health? Not in Sweden!" in Health and Welfare During Industrialization, ed. Richard H. Steckel and R. Floud (Chicago & London: University of Chicago Press, 1997); J.W. Drukker and Vincent Tassenaar, "Paradoxes of Modernization and Material Well-Being in the Netherlands During the Nineteenth Century," in Health and Welfare
Starting in the late 19th century, height of recruits gradually increased by a staggering 10 – 15 centimetres, from around 165 cm to between 175 and 180 cm in most countries. Detailed analyses have shown that declines in infant and childhood mortality are generally followed by increases in height in the same birth cohorts. This suggests that increases in height were indeed not only due to improved nutrition, but also to reduced exposure to infection.\textsuperscript{20}

As in the case of the increase in life expectancy, European countries also differed in the timing of the increase in height, roughly following the pattern in which their national incomes started to grow. Height of recruits in the Nordic countries was already relatively high in the 19th century, perhaps because of proximity to protein in the form of cattle and milk. They also passed the 170 cm level relatively early: this already happened in recruits born around the turn of the 20th century. The same early rise is seen in recruits from, among others, Great Britain, the Netherlands and Switzerland. Southern, Central-eastern, South-eastern and Eastern European countries generally passed this 170 cm threshold much later: here, this happened in recruits born in the 1930s, 1940s or 1950s. Portugal was the last (together with Turkey) to pass the 170 cm level, by recruits born in the 1970s.\textsuperscript{21}

Many European countries abolished compulsory military service in the 1990s, after the collapse of the Soviet Union and the end of the Cold War, so that time-series data of recruit heights stopped abruptly. By then, most countries had passed the 175 cm level. In a few (Sweden, Denmark, the Netherlands, Germany), the average recruit was even taller than 180 cm. Fortunately, studies of growth of children have continued, showing that the secular trend towards

---

\textit{During Industrialization}, ed. R.H. Steckel and R. Floud (Chicago & London: Chicago University Press, 1997). In later phases the effects of industrialization on living standards and health became largely positive; see Joerg Baten and Matthias Blum, "Growing Tall but Unequal," \textit{Economic History of Developing Regions} 27, no. Suppl. 1 (2012): S66–S85.

\textsuperscript{20} See, e.g., Timothy J. Hatton, "How Have Europeans Grown So Tall? ," \textit{Oxford Economic Papers} 66, no. 2 (2014): 349–72; Ida M. Schmidt, M.H. Jørgensen, and Kim Fleischer Michaelsen, "Height of Conscripts in Europe: Is Postneonatal Mortality a Predictor? ," \textit{Annals of Human Biology} 22, no. 1 (1995): 57–67.

\textsuperscript{21} For the association between growth in national income and rise in height, see, e.g., Timothy J. Hatton and Bernice E. Bray, "Long Run Trends in the Heights of European Men, 19th–20th Centuries," \textit{Economics & Human Biology} 8, no. 3 (2010): 405–13; Baten and Blum, "Growing Tall but Unequal." For the relationship between height and access to protein, in the form of cattle, meat and milk, see Joerg Baten and Matthias Blum, "Why Are You Tall While Others Are Short? ," \textit{European Review of Economic History} 18, no. 2 (2014): 144–65. Proximity to protein may also partly explain the extraordinary heights of Dutch recruits (Suppl. Figure 4). For height trends in Portugal, see Yvonne Stolz, Joerg Baten, and Jaime Reis, "Portuguese Living Standards, 1720–1980, in European Comparison," \textit{Economic History Review} 66, no. 2 (2013): 545–78. For height trends in Russia, see Chapter 7, note 60.
greater height is still on-going – but not in the world’s tallest nation, the Netherlands, where the trend seems to have stopped in the early 2000s.\textsuperscript{22}

**More Years in Good Health, More Years in Bad Health?**

Data on long-term trends in morbidity, instead of mortality, are scarce before the 1970s. It is only around that time that several European countries started to conduct nation-wide surveys of (self-reported) chronic disease, disability and general health status. The most important earlier source of information is sickness insurance claims of workers in Britain during the 19th century. This is, admittedly, a narrow basis for making claims about long-term trends in morbidity, but for lack of something better we will briefly summarize the findings of these studies.

The source of information for these studies are the records kept by British ‘Friendly Societies’. Before the modern welfare state developed, these mutual societies provided income replacement and medical attendance during sickness to their members, who were mostly men in manual occupations. Some studies have found a rise of sickness insurance claims coinciding with the decline of mortality in the last decades of the 19th century. This has been attributed to the fact that “[i]f more people survive, then individuals who would have died live longer, although their health remains poor.” Yet, the findings of these studies are far from consistent, so it is difficult to reach firm conclusions.\textsuperscript{23}

Were declines in mortality accompanied by a rise in morbidity or not? It is easy to imagine that they were, if mortality declines were mostly based on

\begin{scriptsize}
\textsuperscript{22} Yvonne Schönbeck et al., “The World’s Tallest Nation Has Stopped Growing Taller,” *Pediatric Research* 73, no. 3 (2013): 371–77. In the Nordic countries, the secular trend is still continuing; see Anton Holmgren et al., “Nordic Populations Are Still Getting Taller—Secular Changes in Height from the 20th to 21st Century,” *Acta Paediatrica* 108 (2019): 1311–320.

\textsuperscript{23} The quote is from George Alter and James C. Riley, “Frailty, Sickness, and Death: Models of Morbidity and Mortality in Historical Populations,” *Population Studies* 43, no. 1 (1989): 25–45, p. 43. Some of the main studies have been done by American historian James Riley; see James C. Riley, *Sickness, Recovery, and Death* (London: MacMillan Press, 1989). The purported rise of morbidity seen in these studies has been criticized for the sensitivity of morbidity rates to ‘cultural inflation’, due to rising health expectations and institutional pressures (S. Ryan Johansson, “The Health Transition: The Cultural Inflation of Morbidity During the Decline of Mortality,” *Health Transition Review* 1, no. 1 (1991): 39–68). However, an empirical study of changing ‘thresholds’ for declaring oneself unable to work did not find much evidence for this (Martin Gorsky et al., “The ‘Cultural Inflation of Morbidity’ During the English Mortality Decline,” *Social Science & Medicine* 73, no. 12 (2011): 1775–783). Other studies than Riley’s have not found much change in age-specific incidence or duration of sickness episodes between the 1870s and 1940s (Bernard Harris et al., “Long-Term Changes in Sickness and Health,” *Economic History Review* 65, no. 2 (2012): 719–45).
\end{scriptsize}
increased survival for sick individuals, as suggested in the quote in the previous paragraph. But it is equally easy to imagine that they were not, if mortality declines were mostly based on declines in the incidence of disease, or if increased survival of sick individuals was due to factors that also increased their recovery rates. In all likelihood, all these mechanisms played a role, but in unknown quantities.

This lack of precise insight in trends in morbidity even applies to more recent time-periods. Despite the increased availability of survey data, and despite attempts at harmonization of survey questions across European countries, the picture that arises from studies of trends in morbidity is far from consistent. A recent comparative study found a decline in severe disability among elderly people in four European countries (Denmark, Finland, Italy, the Netherlands) and an increase in two (Belgium and Sweden) over the last two decades of the 20th century. Why some countries should have experienced a decline and others an increase is a complete mystery.

Combining survey data with data on life expectancy, it has become possible to calculate measures of ‘healthy life expectancy’, or ‘disease-free life expectancy’, or ‘disability-free life expectancy’. Time-trends can only be calculated over the period for which survey data are available, and for an illustration of what this looks like we have prepared Figure 3. This shows trends in the Netherlands, one of the few European countries with survey data reaching back to the early 1980s (and life expectancy data reaching back to the 1850s).

Life expectancy without moderate or severe disability has increased in parallel with total life expectancy, implying that years lived with disability have remained more or less constant during the last three decades. Apparently, the effect of increased survival of people with diseases causing disability has been offset by a decreased risk of these people becoming moderately or severely disabled. Both are probably the result of more effective medical care.

---

24 The prevalence of different levels of disability represents an equilibrium between ‘inflow’ (incidence) and ‘outflow’ (death or recovery). Improvements in medical care reduce the risk of death among disabled individuals, and at the same time shift severely disabled persons to lower levels of disability, so that declines in mortality may go together with declining levels of severe disability and increasing levels of mild disability (Kenneth G. Manton, “Changing Concepts of Morbidity and Mortality in the Elderly Population,” *Milbank Memorial Fund Quarterly Health and Society* 60, no. 2 (1982): 183–244).

25 Gaetan Lafontaine and Gaëlle Balestat, *Trends in Severe Disability among Elderly People*, vol. 26, *OECD Health Working Papers*, (Paris: Organisation for Economic Cooperation and Development, 2007). One possibility is that these differences are spurious and result from variations in reporting.
At the same time, life expectancy without chronic disease has declined, and the number of years lived with chronic disease has increased, as a result of earlier detection of chronic diseases. This may have been a necessary condition for the increased benefits of medical care just mentioned – as we will see in later chapters, early detection and treatment of cardiovascular disease and cancer has been an important ingredient of recent advances in medical care.

Nevertheless, when we take a longer time-perspective, it is clear that whatever measure of ‘healthy life expectancy’ we use, life expectancy in good health is higher in the Netherlands today than total life expectancy 150 years ago, which must mean that the number of years spent in good health has increased.
enormously. Similar results, but over a shorter period, have been found for other European countries.²⁶

That Europeans currently live more years in good health, but at the cost of also living more years in less good health, is confirmed by findings of the Global Burden of Disease study. This study has estimated trends in life expectancy, healthy life-years and years lost to disability since 1990. Of the total number of life-years gained between 1990 and 2017, 26% were spent in poor health, and 74% in good health. Whether the good news in these results outweighs the bad news is difficult to say, even for the specialists who are familiar with these calculations.²⁷

Changes in Disease Patterns

Shifting Causes of Death

The remarkable mortality declines that we saw in the previous section were accompanied by profound changes in the share of specific causes of death. Although some countries started to register causes of death before the 20th century, the cause-of-death classifications that were applied at the time differed between countries and are now considered obsolete. Only in a few cases (e.g., tuberculosis and maternal mortality) will it be possible to present pre-20th century cause-of-death data in this book.

Trends in mortality by cause of death show that the increase of life expectancy during the 20th and early 21st centuries was accompanied by important changes in the percentage of deaths attributed to specific diseases. When life expectancy rose from 50 to 70 years, the share of specific infectious diseases (such as tuberculosis and dysentery) and respiratory diseases (such as pneumonia and influenza) declined, whereas that of cardiovascular diseases, cancers and injuries rose (see Suppl. Figure 5).

²⁶ For an European overview for the early 2000s, see Jean-Marie Robine and Emmanuelle Cambois, “Healthy Life Expectancy in Europe,” Population & Sociétés, no. 499 (2013): 1–4. For the interpretation of trends in total and disability-free life expectancy, see, e.g., Lois M. Verbrugge, “Longer Life but Worsening Health?,” Milbank Memorial Fund Quarterly. Health and Society 62, no. 3 (1984): 475–519; Eileen M. Crimmins, Yasuhiko Saito, and Dominique Inggeri, “Changes in Life Expectancy and Disability-Free Life Expectancy in the United States,” Population and Development Review 15, no. 2 (1989): 235–67.

²⁷ The figures quoted are for high-income countries; see GBD 2017 Collaborators, “Global, Regional, and National Disability-Adjusted Life-Years (DALYs) for 359 Diseases and Injuries and Healthy Life Expectancy (HALE),” Lancet 392, no. 10159 (2018): 1859–922.
After life expectancy had reached a level of 70 years, however, further rises were accompanied by reversals of some of the disease-specific trends. The share of infectious and respiratory diseases started to rise again, instead of continuing to fall, and the share of cardiovascular diseases and injuries started to fall, instead of continuing to rise. Beyond a life expectancy of 70, the share of cancers continued to rise sharply, whereas the share of mental and neurological diseases (including diseases like dementia and Parkinson’s disease) also rose sharply.28

Most of the earlier changes are well-known, and were already incorporated in Omran’s theory of the epidemiologic transition. However, the more recent reversals have not yet fully reached the collective consciousness of demographers and epidemiologists. They point to more radical changes in the health profile of long-lived populations than hitherto recognized.29

The rising share of infectious and respiratory diseases that accompanies rises of life expectancy above the level of 70 years, is probably due to the fact that conditions like sepsis and pneumonia are the only remaining gateways to death among very old people, whose lives have been extended by reducing mortality from cardiovascular diseases. Furthermore, the place of cardiovascular diseases appears to have been partly taken by mental and neurological diseases, particularly dementias, which are now more explicitly recognized as a possible cause of death.30

The recently declining share of cardiovascular diseases is highly remarkable. Some studies have suggested that recent rises in life expectancy were achieved by shifting the average age-at-death of people suffering from chronic diseases, without changing the distribution of deaths by disease. Yet, this idea is incorrect: the average age-at-death has indeed changed, but so has the

---

28 The cause-specific mortality data from which the shares in Suppl. Figure 5 were calculated, were age-standardized so that the effect of changes in age-distribution of the population, which have occurred in parallel to, and partly as a consequence of, rising life expectancy, do not distort the analysis. Levels of life expectancy cited are for men.

29 An exception must be made for French demographers Jacques Vallin and France Meslé, who have made several studies of recent trends in old-age mortality (France Meslé and Jacques Vallin, “Diverging Trends in Female Old-Age Mortality,” Population and Development Review 32, no. 1 (2006): 123–45).

30 For the rise of infectious diseases as a cause-of-death among elderly people, see Aline Désesquelles et al., “After the Epidemiologic Transition,” International Journal of Public Health 60, no. 8 (2015): 961–67. For the rising share of dementias as a cause of death, see Chapter 6 and Johan P. Mackenbach, Marina Karanikolos, and Caspar W.N. Looman, “The Rise of Mortality from Mental and Neurological Diseases in Europe, 1979–2009,” BMC Public Health 14, no. 1 (2014): 840.
proportion of all deaths attributed to cardiovascular disease, which has declined from around 50% to around 30%.\textsuperscript{31}

These changes have occurred throughout Europe, but with important differences in timing. The share of infectious diseases in total mortality was already below 10% in many countries in Northern and Western Europe in the 1930s, whereas countries in Southern, Central-eastern, South-eastern and Eastern Europe often reached that point only in the 1950s. The share of cardiovascular diseases dropped below 40% in the 1990s in many countries in Northern, Western and Southern Europe, but has still not reached that low level in most of Central-eastern, South-eastern and Eastern Europe.

\textit{Shifts in the Burden of Disease}

Even today, we know much more about mortality and its composition by cause-of-death, than about how the burden of morbidity in a population is built up. The Global Burden of Disease study has recently published the results of a heroic attempt to reconstruct the burden of disease in all countries of the world. Filling up the holes in the data necessitated lots of assumptions and imputations, but the results are of great interest. They do not only give a picture of the present, but may even provide a glimpse of what the historical trends in Europe might have been. This is because the study includes low-income countries in other parts of the world, which give an indication of what the situation in Europe might have been before mortality declined to the very low levels of today.

Taking all diseases together, the total disease burden is only some 10% lower in Europe than in the low-income countries of Sub-Saharan Africa and South Asia. This resembles our earlier conclusion that Europe's increase in life expectancy has not been accompanied by an equally strong rise in 'healthy life expectancy'. The 10% lower disease burden in Europe is entirely due to a lower burden of 'communicable, maternal, neonatal, and nutritional diseases'. By contrast, the burden of 'non-communicable diseases' is somewhat higher in Europe than in low-income countries, as is the burden of 'injuries'.\textsuperscript{32}

This suggests that the prevalence of disease has changed much less than one might expect on the basis of the strong declines in mortality. It is the composition of morbidity that has changed: communicable diseases and other

\textsuperscript{31} For the idea that the renewed rise of life expectancy was due to ‘delayed degenerative diseases’, see S. Jay Olshansky and A. Brian Ault, “The Fourth Stage of the Epidemiologic Transition: The Age of Delayed Degenerative Diseases,” \textit{Milbank Quarterly} 64, no. 3 (1986): 355–91.

\textsuperscript{32} See GBD 2017 Collaborators, “Global, Regional, and National Incidence, Prevalence, and Years Lived”, esp. Figure 4.
health problems of the past have declined, but cardiovascular diseases, cancer and other non-communicable diseases, plus the sequelae of injuries, have to a large extent taken their place.

When we stop to think about what exactly happened to these diseases, this is not difficult to understand. The incidence of a disease is the number of new cases arising in the population over a particular period of time, whereas the prevalence of a disease is the number of existing cases present in the population at one point in time. The long-term decline of infectious diseases was mainly a matter of declining incidence, which led to a lower prevalence of disease. On the other hand, the rise of non-communicable diseases was not only due to rising incidence, but also to declining case fatality and thus to a longer duration. The combination of more new cases occurring and longer duration of disease led to a substantially higher prevalence.33

Let's look in more detail at some of the underlying trends, starting with infectious diseases. During the 20th century, many European countries have introduced compulsory notification of infectious diseases. This usually included tuberculosis, syphilis, diphtheria, measles and many other diseases. Although notifications were often compulsory for a limited period of time only, the general picture is clear enough: incidence of many infectious diseases has declined enormously.

For example, in England and Wales notifications declined for respiratory tuberculosis (since the early 20th century), scarlet fever (late 1930s), diphtheria (early 1940s), whooping cough (mid-1950s), poliomyelitis (late 1950s), and measles (late 1960s). At the other side of Europe, in the Soviet Union, the incidence of many infectious diseases also declined, albeit often much later than in Western Europe. For example, notifications declined for smallpox (since the early 1920s), malaria (late 1940s), diphtheria (late 1950s), whooping cough (late 1950s), poliomyelitis (late 1950s) and measles (late 1960s). These declines in incidence have dramatically reduced the disease burden from infectious diseases.34

---

33 The relationship between prevalence and incidence is given by the following formula: \( P = I \times D \), in which \( P \) = prevalence, \( I \) = incidence, and \( D \) = the duration of disease. This was first noted in Brian MacMahon and Thomas F. Pugh, *Epidemiology: Principles and Methods* (Boston: Little, Brown & Company, 1970), p. 65.
34 The history of infectious disease notification in Britain, which started in the late 19th century already, often against fierce resistance of physicians, employers, teachers and others who had to carry out the necessary procedures, can be found in Graham Mooney, *Intrusive Interventions* (Rochester: Rochester University Press, 2015). Long-term trends in England & Wales are shown in Spence Galbraith and Anna McCormack, “Infection in England and Wales, 1838–1993,” in *The Health of Adult Britain 1841–1994. Volume 2*, ed. J. Charlton and M. Murphy (London: The Stationery Office, 1997). Long-term trends in the Soviet Union are shown in S.N. Zatravkin et al., *Заболеваемость Инфекционными*
Let’s now look at the data for an important non-communicable disease, cancer. Cancer registries with which trends in incidence can be followed, have only relatively recently been established. In Europe, Denmark was the first country with a cancer registry (1940s), followed by some of the other Nordic countries (1950s). Coverage of more than 10 European countries was only achieved in the 1980s. In most countries, measurement of the average survival rates of cancer patients started even later.

Nevertheless, there can be no doubt that incidence rates of cancer have risen substantially. This can already be inferred from the rise of the mortality rates from cancer since the beginning of the 20th century, but is also clear from the rise of cancer incidence since whenever cancer registrations were first established (Figure 4). Among Danish women, cancer incidence increased continuously since the early 1940s. Rising trends in total cancer (all sites combined) have been found all over Europe, as a result of rising incidence of a majority of specific cancers. Most of these increases have continued until the present day, although some previously rising cancers (such as lung cancer among men) have recently started to fall.35

At the same time, cancer survival rates have gradually increased. The Nordic countries were the first to start registering the survival of patients with cancer, and data from Norway and Finland show that between the early 1950s and the early 1970s five-year survival of many cancers already improved a little, and that survival of some cancers improved a lot. For all cancers combined, five-year survival increased from 30% to 40% over this period; cancers for which survival improved most were thyroid cancer, Hodgkin’s disease, corpus uteri cancer, malignant melanoma, and testicular cancer, as a result of early advances in cancer treatment.36

35 Long-term cancer incidence trends in the Nordic countries from the 1940s (Denmark) or 1950s (other countries) onwards have been documented in T. Hakulinen et al., “Trends in Cancer Incidence in the Nordic Countries,” Acta Pathologica, Microbiologica, et Immunologica Scandinavica 94 no. S288 (1986): 1–151. For recent trends, see Henrike E. Karim-Kos et al., “Recent Trends of Cancer in Europe,” European Journal of Cancer 44, no. 10 (2008): 1345–389.

36 On Finland, see Paul W. Dickman et al., “Survival of Cancer Patients in Finland 1955–1994,” Acta Oncologica 38, no. 12 (1999): 1–103. On Norway, see Cancer Registry of Norway, Survival of Cancer Patients – Cases Diagnosed in Norway 1968–1975 (Oslo: Cancer Registry of Norway, 1980).
These increases in survival have continued, sometimes accelerated by new advances in treatment. Many more registries now monitor the survival of cancer patients, and European-wide studies show that between the mid-1990s and mid-2000s average five-year survival for all cancers combined increased from 45% to 50%. Survival is, however, considerably better in Northern, Western and Southern Europe than in Central-eastern, South-eastern and Eastern Europe. Progress is slow, and these changes may partly reflect earlier detection of cancer (which artificially increases cancer survival). Nevertheless, the combination of increasing cancer incidence and increasing cancer survival has fuelled a strong rise of cancer prevalence.\textsuperscript{37}

\footnotesize\textsuperscript{37} Milena Sant et al., “EUROCARE-4. Survival of Cancer Patients Diagnosed in 1995–1999. Results and Commentary,” \textit{European Journal of Cancer} 45, no. 6 (2009): 931–91; Paolo Baili et al., “Age and Case Mix-Standardised Survival for All Cancer Patients in Europe 1999–2007,”
As we will see in Chapter 6, increases in survival have also occurred for ischaemic heart disease and stroke. Although incidence of these diseases has declined much more than the incidence of cancer, the prevalence of cardiovascular diseases has risen as well.

**Diseases Rise, Diseases Fall**

These changes in the broader composition of mortality and morbidity are the net result of what happened to many specific diseases. Many diseases have manifested themselves in a striking pattern of ‘rise-and-fall’, but with a total lack of simultaneity. Whereas some ‘rises-and-falls’ occurred over many millennia, others took only two decades to unfold. Whereas some diseases peaked in the 18th century, others did so in the 19th or 20th century. Yet others are still rising and will hopefully decline in the future.38

This will be documented more extensively in Part 2 of this book, of which Table 1 presents a ‘sneak preview’. When I started these analyses, I knew that ischaemic heart disease and road traffic injuries had first risen and then declined, because the peak had occurred during my life-time and had been discussed in research papers published in epidemiological journals. Yet, when a sufficiently long time-scale is applied, a pattern of ‘rise-and-fall’ appears to be the general rule. Ischaemic heart disease and road traffic injuries have risen over more than a half-century before peaking in the second half of the 20th century. Appendicitis and AIDS took a much shorter period to rise to their peak in the 1930s and 1990s, respectively. Tuberculosis and syphilis took several centuries before they reached their peak in the 19th century.

The rise of some other diseases, such as dysentery and other intestinal infections, probably started with the advent of agriculture in the Neolithic period (around 6,500 BCE in Europe). The same applies to famine, which was probably less common among hunter-gatherers than among agriculturalists who

---

38 That human diseases tend to occur in a pattern of ‘rise-and-fall’ was also noted by René Dubos (e.g., René J. Dubos, *Man Adapting* (New Haven and London: Yale University Press, 1965), Chapter 9) whose views we will discuss in more detail in Chapter 3. The notion of ‘rise-and-fall’ has been used in connection with many specific diseases; see, e.g., Reuel A. Stallones, “The Rise and Fall of Ischemic Heart Disease,” *Scientific American* 243, no. 5 (1980): 53–9; David J. Barker, “Rise and Fall of Western Diseases,” *Nature* 338, no. 6214 (1989): 371–27. As this book shows it is a truly generalized phenomenon.
were dependent on the growth of a limited number of crops. To my surprise, on a time-scale of millennia even war-related violence has developed in a pattern of ‘rise-and-fall’: rising until the costs of war finally became prohibitive.39

| Group | Health problem       | Rise and fall? | Start of rise\(a\) | Start of fall\(b\) |
|-------|----------------------|----------------|---------------------|-------------------|
| Health problems of pre-industrial societies | | | | |
| War   | Rise and fall        | (before 4300 BCE) | 16th century\(c\) |
| Homicide | Fall only       | N/A            | 16th-17th century |
| Famine | Rise and fall      | (6500 BCE\(d\)) | 18th century |
| Plague | Rise and fall     | 1347           | 17th century |
| Smallpox | Rise and fall   | 6th century | 18th century |
| Typhus | Rise and fall     | Late 15th century | 17th century? |
| Malaria | Rise and fall | 16th century | 18th century |
| Health problems of industrializing societies | | | | |
| Cholera | Rise and fall | 1829-1837 | 1846-1860 |
| Intestinal infections\(f\) | Rise and fall | (6500 BCE\(d\)) | Mid-19th century |
| Tuberculosis | Rise and fall | 18th century | Mid-19th century |
| Syphilis | Rise and fall | Late 15th century | Early 20th century |
| Childhood infections\(g\) | Rise and fall | 18th century | Late 19th century |
| Pneumonia | Fall only       | N/A            | Early 20th century |
| Influenza | Rise and fall | 16th century | 1918-1919 |
| Puerperal fever | Rise and fall | 18th century | Mid-19th century |
| Infant mortality | Fall only | N/A         | Late 18th – late 19th century |
| Health problems of industrializing societies | | | | |
| Still-births | Fall only | N/A         | 1940s |
| Pellagra\(h\) | Rise and fall | 18th century | Late 19th century |
| Rickets | Rise and fall | 17th century | Late 19th century |
| Goitre | Fall only         | N/A            | 1920s |
| Peptic ulcer | Rise and fall | Late 19th century | 1930s-1940s |
| Appendicitis | Rise and fall | Late 19th century | 1930s-1940s |
| Pneumoconiosis | Rise and fall | 19th century | Early 20th century |
| Mesothelioma | Rise only | 1970s | N/A |
| Stomach cancer | Rise and fall | 19th century? | Early 20th century |

39 See Chapters 4–6 for these and many other examples of ‘rise-and-fall’. It has not always been possible to capture the complete ‘rise-and-fall’ trajectory in the figures, because most of these rely on mortality data by cause of death, and because the rise of many diseases occurred before the introduction of national cause-of-death registration. In these cases, the evidence for an earlier ‘rise’ comes from the literature.
Industrialization and urbanization also went together with a rise of many diseases, ranging from rickets to tuberculosis, and from diphtheria to pneumoconiosis (occupational lung disease). These diseases have since largely disappeared, but have partly been replaced by the health conditions with which we are now familiar, such as ischaemic heart disease, cancer and road traffic injury. This generalized pattern of ‘rise-and-fall’ is not only fascinating,
but also holds important clues for the explanation of long-term trends in population health to which we will come below.

Epidemiologic Transition 2.0

* A Theory in Need of Repair

Omran’s idea — briefly summarized in Chapter 1 — was that total mortality had declined because, as a result of socioeconomic development, mortality from infectious diseases had declined. After that had occurred, total mortality stabilized because, as a side-effect of the same socioeconomic development, other causes of death had increased and replaced infectious diseases. However, because the new diseases caused death at higher ages (i.e., among the middle-aged and elderly, instead of among babies and children), the exchange led to higher life expectancy. According to Omran, this all happened somewhere between 1870 and 1970, albeit with important differences in timing and pace between countries, e.g., later but faster in Eastern Europe than in Western Europe. At first sight, this may seem a reasonable summary, but upon closer inspection it is far from accurate. There are three important problems.

A first problem concerns Omran’s *periodization*. In Northern and Western Europe, the reduction of infectious disease mortality started much earlier than in the last decades of the 19th century. The decline of plague, smallpox and typhus, to mention just a few examples, had already started at least a century earlier. More generally, a clear attenuation of the great fluctuations in mortality, due to devastating epidemics and famine, already occurred in the 18th century, at least in Northern and Western Europe, and represented an important advance in population health that cannot be ignored.

In addition to a problem with the beginning, there is a problem with the end of Omran’s epidemiologic transition. Coincidentally, just when Omran published his theory, a renewed decline of mortality started in many high-income countries. A reversal in the trend of ischaemic heart disease was one of its main drivers, but declines in other causes of death such as road traffic injuries also played a role. This renewed decline has sometimes been referred to as a fourth stage of the epidemiologic transition. This has been labelled the “age of delayed degenerative diseases” (because it was mistakenly thought that this did not entail an elimination of mortality from heart disease, but only a delay to older ages) or a “hybristic stage” (because trends were increasingly influenced by individual behaviours and life-styles).

---

40 Omran defined the end of the epidemiologic transition as the point in time when mortality rates stabilised after their decline. However, when crude mortality rates are analysed,
Another surprise was that it soon appeared that infectious diseases had not been eliminated, but on the contrary seemed to return with a vengeance. After 1970, new infectious diseases emerged (such as legionnaire’s disease and AIDS), some old diseases re-emerged as a result of lapses in infectious disease control (such as tuberculosis and whooping cough), and some other infectious diseases became more difficult to control because of antibiotic resistance (such as hospital-acquired staphylococcal infections). Adding to the confusion, it was proposed to either label this a fifth stage of the epidemiologic transition, or to consider this an entirely new epidemiologic transition.41

Although a certain rise of mortality from infectious diseases has been noted in many European countries, it is not obvious that this is an important enough phenomenon to warrant the addition of another stage of the epidemiologic transition, let alone to distinguish an entirely new transition. On the other hand, the decline of cardiovascular diseases certainly was important, and qualitatively different from what had happened earlier. This suggests that distinguishing a new stage to accommodate this change may indeed be appropriate.

It does not seem wise, however, to label this a new transition, because by definition a ‘transition’ is an intermediate stage between two more or less stable states. This can only be accurately identified and defined after the event, and at present we do not yet know where we are heading. Which causes of death will ultimately replace the causes of death that are being pushed back?

---

41 The proposal to label this a ‘fifth stage’ (of ‘emerging and re-emerging infectious diseases’) comes from S. Jay Olshansky et al., “Emerging Infectious Diseases: The Fifth Stage of the Epidemiologic Transition?,” World Health Statistics Quarterly 51, no. 2–4 (1998): 207–17. The proposal to label this a ‘third epidemiologic transition’ comes from Ronald Barrett et al., “Emerging and Re-Emerging Infectious Diseases: The Third Epidemiologic Transition,” Annual Review of Anthropology 27, no. 1 (1998): 247–71. For readers who are losing count: in the latter conception the first epidemiologic transition involved the rise of infectious diseases that accompanied the Neolithic agricultural revolution, while the second one is synonymous with Omran's epidemiologic transition. The idea of a ‘second epidemiologic transition’ was more fully developed in Molly K. Zuckerman, ed., Modern Environments and Human Health (Hoboken: John Wiley & Sons, 2014). In later updates of his ‘epidemiologic transition theory’ Omran actually embraced some of the extensions (Omran, “Epidemiologic Transition Theory Revisited”).
We have seen above that the decreasing share of cardiovascular diseases in total mortality is accompanied by an increasing share of neoplasms and of mental and neurological disorders, but neoplasms have been on the rise for long, and the share of mental and neurological diseases is still very low.\footnote{We will see in Chapter 6 that the rise of mortality from Mental and neurological diseases is mainly due to rising mortality from dementia, and that it is uncertain whether this rise of mortality from dementia is real.}

A second problem relates to Omran’s \textit{labelling} of groups of diseases. Enormous shifts have occurred in both the absolute rates and relative importance of specific causes of death. But how should these shifts be characterized?

Omran characterized the shift as one from “pandemics” of infectious diseases to “degenerative and man-made diseases.” However, although cholera and influenza were pandemic diseases, most of the other infectious diseases that were responsible for declining mortality, such as tuberculosis, pneumonia, and dysentery, were endemic diseases. Furthermore, we now know that some of Omran’s ‘degenerative’ diseases (such as stomach and cervical cancer) have partly infectious origins as well.

Nor is the vaguely moralistic heading of ‘degenerative and man-made diseases’ an adequate label for the causes of death that have replaced the infectious diseases. Although cardiovascular diseases and neoplasms are diseases of middle and old age, we now know that they are not primarily caused by age-related biological processes of ‘degeneration’, but are predominantly ‘exogенно’ caused, just like the infectious diseases that dominated the cause-of-death pattern in previous centuries. Also, while many of these exogenous causes, such as smoking and excessive alcohol consumption and occupational exposures, are ‘man-made’, so were the living conditions and habits which promoted the transmission of infectious agents.

In theory, it may be possible to solve these terminological problems by finding other umbrella terms, but in practice this has proven to be difficult, and no solution is entirely satisfactory. What this really demonstrates is that the long-term improvement in population health that the epidemiologic transition tries to capture, was a heterogeneous phenomenon. It consisted of sequentially ‘falling’ diseases, with an emphasis on declines for some infectious diseases in one period, an emphasis on declines for some other infectious diseases in another period, and an emphasis on declines for cardiovascular diseases in still another period, but without clear-cut differences by type of disease, however defined.\footnote{For some further elaboration, see Johan P. Mackenbach, “The Epidemiologic Transition Theory,” \textit{Journal of Epidemiology and Community Health} \textbf{48}, no. 4 (1994): 329–32.}
The third problem is that Omran largely ignored *other health outcomes than mortality*. There is an implicit assumption in many of his interpretations that changes in mortality equal changes in incidence of the underlying diseases. However, some of the changes in mortality, even before the early 1970s, were due to declines in case fatality, as in the case of the treatment of pneumonia and other infectious diseases by antibiotics. Over time, trends in morbidity (and health expectancy) have become at least as important for evaluating population health as trends in mortality (and life expectancy), but they are not captured by the epidemiologic transition theory.

These and related omissions have led to proposals to replace the concept of the ‘epidemiologic transition’ by that of a ‘health transition’. This concept includes changes in morbidity in addition to mortality, but also society’s responses to health problems such as the expansion of health care services. While that may risk mixing up the dependent variable (i.e., changes in population health) with the independent variable (i.e., what society does to improve population health), it is certainly important to look beyond life expectancy. As shown above, data on disability-free life expectancy suggest that life expectancy without severe disability has increased in parallel with total life expectancy, implying that years lived with severe disability have remained more or less constant, whereas years lived with any disability and particularly with chronic illness have increased during the last stage of mortality decline.44

Taking these problems into account, I propose the following modification of the epidemiologic transition theory: ‘epidemiologic transition 2.0’.

**How: Characterizing Change**

As we have seen above, long-term trends in population health have been the result of many superimposed disease-specific trends, and the latter are more often than not characterized by a distinct pattern of ‘rise-and-fall’. One way to understand this endless succession of ‘rises-and-falls’ is to regard them as an accompaniment of socioeconomic development. As a result of the drive for socioeconomic improvement, and the changes in behaviour associated with this drive, mankind is continuously confronted with new disease risks. However, because higher levels of development also bring higher living standards and a greater desire and capacity to control disease, after some time these disease risks decline again (but often to be replaced by new disease risks).

This suggests that, to the extent that we need a distinction in different stages of population health change, it may be useful to relate disease trends to stages

44 The concept of the ‘health transition’ was proposed in Julio Frenk et al., “Elements for a Theory of the Health Transition,” *Health Transition Review* 1, no. 1 (1991): 21–38.
of socioeconomic development. This is what I have already done, without explicitly mentioning it in the text, in Table 1. In this table I grouped diseases into three large groups, depending on the period in which their mortality rates peaked in North-western Europe: ‘health problems of pre-industrial societies’, which generally peaked before 1800; ‘health problems of industrializing societies’, which peaked between 1800 and 1950; and ‘health problems of affluent societies’, which peaked after 1950. Let me explain.

In the first group I brought together a number of diseases (like smallpox and typhus) and other causes of ‘crisis mortality’ (war and famine) which caused many of the large peaks in mortality that were mostly gone – at least in North-western Europe – before 1800. These can be regarded as ‘health problems of pre-industrial societies’, not only because they peaked before most North-western European countries began to seriously industrialize, but also because their causes were linked to the material hardship that was prevalent before modern economic growth took off. (Homicide has been included in this group as well, because it has been in secular decline since the 15th century.)

In the second group I brought together the diseases that were associated with the large-scale industrialization and urbanization of the 19th century. Their rise had often started before the 19th century, but continued or accelerated to reach a peak before the middle of the 20th century, the subsequent decline being partly a result of the success of control efforts. This group includes well-known diseases like tuberculosis and diphtheria, but also infant mortality which, in many countries, had a temporary peak in the late 19th century as a consequence of urbanization. It also includes a few nutrient deficiency diseases, which rose as a result of dependence on a more monotonous diet, as well as occupational lung disease, which increased as a result of the mechanization of grinding, and peaked in the 1920s.

In the third group I brought together the diseases that peaked after 1950, or that have risen but not yet peaked at all. These diseases have been labelled ‘diseases of affluence’, but this is the trickiest label of the three. During the 20th century, when scientists tried to make sense of the rise of these diseases, various labels have been proposed. In addition to Omran’s ‘degenerative and man-made diseases’, these include ‘diseases of affluence’, ‘diseases of civilization’ and ‘Western diseases’, as well as more neutral terms like ‘non-communicable diseases of affluence’, ‘diseases of civilization’ and ‘Western diseases’, as well as more neutral terms like ‘non-communicable

---

45 The inclusion of war in this group could be debated, because war-related deaths only peaked in the middle of the 20th century. However, the frequency of war peaked in the 16th century.

46 Stomach cancer and haemorrhagic stroke can also be included in this group, because mortality declined since the early 20th century. However, for practical reasons trends of these diseases will be discussed in Chapter 6.
diseases’ and ‘chronic diseases’. None of these terms is satisfactory, and all are inaccurate – even a simple distinction between communicable and non-communicable diseases no longer holds.47

I have decided to retain the term ‘diseases of affluence’ because the rise of most of these diseases was facilitated by rising living standards, or by what mankind did to achieve these living standards. This applies to cardiovascular diseases, diabetes mellitus, most cancers, and many causes of injury. I also included dementia and depression in this group, although dementia (and other mental and neurological diseases) may turn out to be characteristic of a future stage of health development, and although it is unclear whether depression is more frequent in affluent societies than it was in the past. Finally, this group includes a ‘ghost from the past’, i.e., the infectious disease AIDS. This spread as a result of international travel and liberalization of sexual mores, conditioned by the affluence of the second half of the 20th century.

When: Staging Change
This distinction of three categories of diseases also suggests a distinction of three stages of population health change over the past three centuries. This distinction is necessarily somewhat arbitrary, because diseases’ peak-years are spread out rather evenly over the three centuries since 1700, without natural dividing-lines. This emerges very clearly from Table 1: the last column of this table shows a virtually continuous succession of diseases starting to fall. Also, the peak-years differed between European regions; for example, whereas typhus may have peaked in the 17th century in Sweden, it probably peaked in the 20th century in Russia. We should thus not assume a universal timing of these changes. Nevertheless, three possible stages would be the following:

47 For example, as will be discussed in Chapters 5 and 6, peptic ulcer and stomach cancer are now considered to be essentially infectious diseases. It has been speculated that microorganisms are also involved in the causation of cardiovascular diseases, which suggests a deeper link between different stages of the epidemiologic transition (Alex Mercer, Infections, Chronic Disease, and the Epidemiological Transition (Cambridge: Boydell & Brewer, 2014)). For the term ‘diseases of affluence’ see Thomas McKeown, The Origins of Human Disease (Oxford: Basil Blackwell, 1988), for ‘diseases of civilisation’ see Henry E. Sigerist, Civilization and Disease (New York: Cornell University Press, 1943), and for ‘Western diseases’ see Hubert C. Trowell and Denis P. Burkitt, Western Diseases, Their Emergence and Prevention (Cambridge (Mass): Harvard University Press, 1981). All these terms are easy to criticize, but difficult to replace by a satisfactory alternative. Many of the so-called ‘diseases of affluence’ have fallen in later stages of the rise in living standards; the term ‘diseases of civilisation’ refers to a Euro-centred view of human civilization; and the idea of ‘Western diseases’ is obsolete because most of these disease are now global.
In a first stage the frequency and amplitude of annual fluctuations in mortality started to diminish, due to a reduction in the frequency and/or severity of famine and epidemics of plague, smallpox, typhus and malaria. In Northwestern Europe, these declines started before 1800, but some of these problems continued to cause spikes in mortality in the 19th and even 20th century before disappearing completely. This stage thus does not have a natural end, and if an end is necessary it seems best to take the start of the second stage as the end of the first. Before 1800, mortality in non-crisis years may or may not have started to slowly decline as well, but life expectancy at birth remained low at a level not much higher than 30 years.48

In a second stage a range of health problems characteristic of industrializing societies, and often causally linked to industrialization and urbanization, started to decline. Among the diseases studied in this book, this applies to various infectious diseases (cholera and other diarrheal diseases, tuberculosis, syphilis, childhood infections, pneumonia,...), maternal and infant mortality, peptic ulcer and appendicitis, nutrient deficiency diseases, and occupational lung disease. This decline was mostly the result of decreasing disease incidence, but during this stage decreasing case fatality also started to play a role, particularly towards the end. Like the first stage, this second stage also does not have a natural end, as declines from these conditions have continued even after the third stage had started.

In Northwestern Europe the first signs of declining mortality from some infectious diseases (e.g., tuberculosis) could already be noticed in the first half of the 19th century, but declines really got under way in the last decades of the 19th century. At that point, the modest increases in life expectancy that had already occurred, suddenly accelerated, mainly as a result of the start of a rapid decline of infant mortality. In this second stage, the proportion of total mortality due to infectious diseases declined from more than 40% to less than 10%. Life expectancy increased from little more than 30 years in the first half of the 19th century to around 65 years (men) and 70 years (women) in the 1950s.

In a third stage a number of diseases that had risen during the second stage, probably as a side-effect of increasing prosperity, started to decline: ischaemic heart disease, stroke, a few cancers, and road traffic injuries and suicide. Like in previous stages, this was still partly the result of declines in incidence, but

---

48 Perrenoud, “The Attenuation of Mortality Crises” puts the end of the age of large mortality crises in the late 1600s for England, Italy and Geneva, in the early 1700s for France, in the early 1800s for Sweden, Norway and Denmark, and in the late 1800s for Finland. As noted above, however, large mortality fluctuations continued to occur in the 20th century (see Figure 2).
increasingly declines in case fatality also played an important role. In this stage, the share of cardiovascular mortality in total mortality declined from a peak between 40 and 70% to much lower values, whereas the share of cancer continued to increase, and the share of mental/neurological diseases now started to increase.

One other characteristic of this third stage is that mortality decline among the elderly, including the oldest old, accelerated, contributing to a continued increase of life expectancy at birth to levels sometimes exceeding 80 years among men and (more often) women in the first two decades of the 21st century. In this stage, we can also observe what happens to health expectancy: both years-in-good-health and years-in-bad-health go up, but it may well be that years-in-very-bad health have remained stable or even declined somewhat.

Of these three stages, the second more or less corresponds with Omran’s epidemiologic transition. There are still good reasons to see this as a very special stage. This was a stage with ultra-rapid increases in life expectancy, in which disease-specific declines strongly dominated disease-specific increases, as a result of unprecedented improvements in living conditions and a series of important breakthroughs in the control of infectious diseases.

Where: Locating Change

In the preceding section we gave an approximate timing of the three stages for North-western Europe. However, there were large differences in timing of these changes between countries (see Suppl. Table 4).

It is challenging to assess differences in timing of the first stage of mortality decline, because so few countries have data going back into the 18th century. It is only at the end of the 19th century that some countries in all European regions were collecting data permitting the calculation of crude mortality rates. At that point in time, annual fluctuations in mortality were still substantial in South-eastern and Eastern Europe, indicating that these countries were late in going through this first stage.

If we take decline of infant mortality as a marker of the second stage, there are again clear differences between European regions. By the beginning of the 20th century, countries in Northern and Western Europe were distinctly ahead of Southern, Central-eastern, South-eastern and Eastern Europe, where levels below 10% would often only be reached after World War II.

In a similar way, when we use the proportion of all deaths due to cardiovascular disease as a marker of progression of the third stage, we find that in the beginning of the 21st century Northern and Western Europe are far ahead of Central-eastern, South-eastern and Eastern Europe. Interestingly, Southern
Europe has now shed off its disadvantage, and has joined the front-runners in Northern and Western Europe.\textsuperscript{49}

Although all European regions have experienced a strong increase in life expectancy since the late 19th century, and although some convergence of life expectancy has occurred during the 20th century, it is amazing to see how certain patterns of variation have persisted throughout this period. Broadly speaking, the Nordic countries are still ahead of many other European countries, and South-eastern and Eastern Europe still lag behind. This suggests that some of the causes of these patterns of variation in population health have also persisted over more than a century, and perhaps longer.

At the same time, not all regions have kept their relative position. This applies most clearly to Southern or Mediterranean Europe, which lagged behind North-western Europe until the mid-20th century, but now has some of the longest-lived populations in the world. As we will see in later chapters, this is due, on the one hand, to a process of catching-up with North-western Europe (e.g., in eliminating communicable diseases). On the other hand, these countries have not experienced a similar rise as North-western Europe in cardiovascular disease mortality during the second stage of mortality decline.

\textsuperscript{49} This pattern is confirmed when we look at differences in the year in which European countries passed certain thresholds for female life expectancy (Suppl. Table 4): 50 years (as a marker of progression of the first stage), 65 years (as a marker of progression of the second stage), and 80 years (as a marker of progression of the third stage). Northern and Western Europe were ahead of the rest in all three stages, but in the third stage they were joined by Southern Europe.