The Determinants of Mortality

David Cutler, Angus Deaton and Adriana Lleras-Muney

The pleasures of life are worth nothing if one is not alive to experience them. Through the twentieth century in the United States and other high-income countries, growth in real incomes was accompanied by a historically unprecedented decline in mortality rates that caused life expectancy at birth to grow by nearly 30 years.

The value of reductions in mortality risk can be roughly estimated from (admittedly heroic extrapolations of) differential wages in the labor market corresponding to differentials in the risk of death across occupations. Applying this methodology, Nordhaus (2002, p. 35) has calculated that “to a first approximation, the economic value of increases in longevity in the last hundred years is about as large as the value of measured growth in nonhealth goods and services.” Falling mortality has also usually meant better health for the living, so that people are also living better, healthier, and longer lives than did their forebears. Murphy and Topel (2005), who measure both the value of mortality decline and the benefits of better health for the living, estimate that, between 1970 and 2000, the annual value of increased longevity was about half of conventionally measured national income.

Improvements in life expectancy in the United States have been matched by similar improvements in other rich countries. Indeed, there has been a rapid
convergence of older adult mortality rates since 1970 in rich countries, particularly among men (Deaton, 2004, Figure 7). Outside of the rich countries, average health is strongly correlated with income. As shown in Figure 1, the current version of a graph first drawn by Preston (1975) in which countries are represented by circles and the size of the circle is proportional to population, life-expectancy is profoundly lower for countries with lower levels of per capita income.

In the years just after World War II, life expectancy gaps between countries were falling across the world. Poor countries enjoyed rapid increases in life-expectancy in the 1950s, 1960s and 1970s, with the gains in some cases exceeding an additional year of life expectancy per year. The HIV/AIDS epidemic and the transition in Russia and eastern Europe have changed that situation. The best estimates of life-expectancy in some sub-Saharan African countries are lower now than they were in 1950. Life expectancy in Russia fell by nearly seven years over the 1990s. However, at least up to the 1990s, compound welfare measures that incorporate both health and income show both much greater inequality at any point in time and much greater international convergence than do income measures alone (Becker, Philipson and Soares, 2005).

There is also a positive relationship between income and health within countries—low-income people live shorter lives than high-income people in a given country. Americans in the bottom 5 percent of the income distribution in 1980 had a life-expectancy at all ages that was about 25 percent lower than the corresponding life-expectancies of those in the top 5 percent of the income distribution (Rogot, Sorlie, Johnson and Schmitt, 1992). These “health inequalities,” also known as

**Figure 1**
The Preston Curve: Life Expectancy versus GDP Per Capita

*Source:* Reproduced from Deaton (2003, Figure 1).

*Note:* Circles are proportional to population.
gradients, are part of a wider pattern that relates mortality to measures of socio-economic status. For example, American blacks had a life expectancy in 2002 that was 5.4 years less than that of American whites. In England and Wales in 1997–2001, male manual workers could expect to live 8.4 years less than professionals, a gap that has increased since the early 1970s.

The decline in mortality over time, differences in mortality across countries and differences in mortality across groups within countries are phenomena worthy of serious attention by economists and others. We will first lay out the historical decline in mortality, then move to mortality differences between rich and poor countries, and then discuss differences in mortality within countries. A good theory of mortality should explain all of the facts we will outline. No such theory exists at present, but at the end of the paper we will sketch a tentative synthesis.

Determinants of the Historical Decline in Mortality

For most of human history, life was properly described in the famous phrase of Thomas Hobbes as “nasty, brutish, and short.” From the dawn of *Homo sapiens* perhaps 100,000 years ago until the first agricultural revolution in roughly 10,000 BCE, world population was about four million people. Life expectancy at birth for our hunter–gatherer ancestors was perhaps 25 years. There had been little, if any, progress by the Roman Empire, and even in 1700, life expectancy at birth in England—after the Netherlands, the richest country in the world at the time—was only 37 years (Wrigley and Schofield, 1981).

In the eighteenth century, mortality began to decline. In England and Wales (which we refer to as “England” for convenience), the decline started around the middle of the eighteenth century. By 1820, life expectancy at birth in England was about 41 years, up six years over the previous century. Between 1820 and 1870, the period of greatest industrialization, life expectancy remained stable at about 41 years. Since 1870, mortality has fallen relatively continuously as well as more rapidly than in the first phase of mortality decline. Life expectancy in England climbed to 50 years in the first decade of the twentieth century, and is about 77 years today. A similar transition, with some moderate differences in timing, took place in all developed countries. Mortality reduction in France was broadly similar to that in England. In the United States, the mortality reduction appears to start around 1790, with a similar overall pattern. Life expectancy at birth in the United States rose from 47 years in 1900 to 78 years today.

The reduction in mortality was not uniform by age. The vast bulk of the historical reduction in mortality occurred at younger ages. Figure 2 shows trends in life expectancy by age in England since 1840. Between 1841 and 1950, life expectancy at birth increased by 30 years, while life expectancy at age ten increased by only half that amount. The decline in infectious disease explains this disparate age pattern. In 1848, 60 percent of deaths in England were from infectious disease. Between then and 1971, infectious disease mortality declined by 95 percent. Since
infants and children are the most vulnerable to infections, their mortality rates were most affected by the decline in infections. The sources of the reduction in infectious disease mortality have been extensively debated in the demographic community. We discuss the relevant factors in (possible) order of historical importance and conclude with some open issues.

**Improved Nutrition**

Agricultural yields increased significantly during the eighteenth century. Better-fed people resist most bacterial (although not viral) disease better, and recover more rapidly and more often. The British physician and demographer Thomas McKeown was the first person to argue for the importance of nutrition in improved health, writing several seminal papers on the topic which culminated in his widely read 1976 book. McKeown argued by residual analysis: neither personal health care nor public health appeared to have had much impact prior to the 1900s, when most of the mortality decline had already occurred. In a famous example, McKeown showed that mortality from tuberculosis fell by 80 percent before there was any effective treatment for the disease. The same is true for other infectious diseases as well. However, many analysts found unconvincing both McKeown’s dismissal of public health, as well as the argument by elimination that nutrition was the crucial factor (for example, Szreter, 1988; Guha, 1994).

Direct evidence on the role of nutrition in improved health and mortality reduction comes from the work of Robert Fogel, in a series of papers summarized

---

**Figure 2**

*Expected Age at Death, England and Wales*

![Figure 2](image_url)

*Source:* Data for 1751–1841 are from Wrigley and Schofield (1981, Table 7.15, p. 230); data from 1841 are from the Human Mortality Database.
in Fogel (1997) and in his 2004 book. Fogel begins by showing the enormous increase in caloric intake after the middle of the eighteenth century, measured both directly from agricultural output and diary surveys, and indirectly through changes in adult height. Between the middle of the eighteenth century and today, for example, caloric intake per person increased by more than a third, and heights in most of Europe increased by ten centimeters or more (Fogel, 1994, Table 1). Mortality is U-shaped in the body mass index (weight divided by height squared), and declines with height given the body mass index. Fogel (1997) uses these relationships to argue that nearly all of the reduction in mortality from the late eighteenth century to the late nineteenth century can be attributed to improved nutrition, as well as half of the mortality improvement in the century after that.

But the evidence on calorie availability has not convinced everyone. One line of argument is that the increase in life expectancy in England from 1750 to 1820 had nothing to do with increased income per head, but was just one of the fluctuations in mortality that characterized pre-industrial Europe. Indeed, Wrigley and Schofield (1981) estimate that life expectancy in 1600 was the same as in 1820, with 1750 being the low point of a two-century swing. Steckel (2004) argues on the basis of skeletal remains that people were taller (and presumably better nourished) in early medieval times. If we accept this argument and date the modern decrease in mortality from 1870, when it began in earnest in England and several other European countries, the link between economic growth and mortality becomes tenuous, because the timing of the beginnings of modern growth is far more dispersed across countries than is the onset of the modern mortality decline (Easterlin, 2004). Also, as we shall see, improvements in public health offer a more coherent explanation for mortality declines after 1870.

Another concern with the nutritional story is that, from the sixteenth to the eighteenth centuries, English aristocrats had no life expectancy advantage over the rest of the population, despite presumably better nutrition. Moreover, mortality was not lower in well-fed populations of the same period, such as in the United States (Livi-Bacci, 1991).

Further, there are powerful two-way interactions between disease and nutrition. Children who are frequently malnourished often continually suffer from poorly-controlled infectious disease. Diseases such as diarrhea prevent food intake from nourishing the body; children who suffer repeated episodes of diarrhea may be able to digest less than 80 percent of what they consume (Dasgupta and Ray, 1990). As a result, some argue that it was disease, not nutrition, that was the primary exogenous influence, and that disease burdens changed most strongly as a result of public health intervention.

Public Health

The argument for the role of public health in reduced mortality is made most prominently by Samuel Preston (1975, 1980, 1996). If economic growth were the sole reason for improved health, countries would move along the “Preston curve” shown in Figure 1, but the curve itself would remain fixed. However, even at a given
level of income, people live substantially longer today than they did in the past. For example, China in 2000 has the income level of the United States in the 1880s, but has the life expectancy of the United States in 1970—about 72 years. Preston estimates that only about 15 percent of the increase in life expectancy between the 1930s and 1960s is a result of increases in income alone (Preston, 1975). While income was certainly mismeasured historically, creating uncertainty in this estimate, there is no recorded income in 1930 that would have predicted the life expectancy observed in many countries in 1960.

Public health improvements are an obvious explanation for this shift. Macro public health involves big public works projects: filtering and chlorinating water supplies, building sanitation systems, draining swamps, pasteurizing milk and undertaking mass vaccination campaigns. Micro public health involves changes made by individuals but encouraged by the public sector, including boiling bottles and milk, protecting food from insects, washing hands, ventilating rooms and keeping children’s vaccinations up to date. Macro public health was always present to some extent. Even in the Middle Ages, it was known that people living in areas where bubonic plague was rampant should be quarantined. (Unfortunately, rats were not quarantined.) Benjamin Latrobe built a water system in Philadelphia early in the nineteenth century, at least partly to reduce the disease burden. In 1854 John Snow compared cholera fatalities between households supplied by two different water companies, one of which was recycling human waste, and one of which was not. He thus demonstrated that cholera was water-borne and that its spread could be halted by uncontaminated water supplies. But big public health did not fully come into its own until the acceptance of the germ theory of disease in the 1880s and 1890s, which led to a wave of new public health initiatives and the conveyance of safe health practices to individuals.

The dramatic reduction in water and food-borne diseases after that time—typhoid, cholera, dysentery, and non-respiratory tuberculosis—highlights the role of public health. From a mortality rate of 214 per 100,000 in 1848–54, these diseases were virtually eliminated in the United States by 1970. By one estimate, water purification alone can explain half of the mortality reduction in the United States in the first third of the twentieth century (Cutler and Miller, 2005).

Urbanization

If rising living standards were good for health, urbanization was not, at least initially. The preponderance of the evidence suggests that the lack of improvement in mortality between 1820 and 1870 was due in large part to the greater spread of disease in newly enlarged cities. Nutrition may or may not be the culprit here; debate about whether nutrition and real wages were rising or falling in the middle of the nineteenth century continues to rage. But the effect of unsanitary conditions was larger and the spread of disease was easier in bigger, more crowded cities.
Vaccination

Prior to the twentieth century, there was little effective medical treatment for infectious disease. Over the course of the twentieth century, however, the role of medical advances increased in importance.

The first important medical interventions were vaccinations. Variolation against smallpox, practiced in China as early as the tenth century, was an early form of immunization whereby matter from the scabs of previous victims was introduced into the bodies of healthy people. Variolation was introduced to Europe from Turkey and to the American colonies by African slaves in the early eighteenth century. George Washington variolated his entire army. Vaccination was introduced by Edward Jenner at the end of the eighteenth century, but wide-scale research on vaccines depended on the germ theory of disease and did not occur until a century later. Since the late nineteenth century, there have been a number of new vaccines, including those for rabies (1885), plague (1897), diphtheria (1923), pertussis (1926), tuberculosis (1927), tetanus (1927), yellow fever (1935), polio (1955 and 1962), measles (1964), mumps (1967), rubella (1970), and hepatitis B (1981).

The morbidity consequences of these diseases were high, but the best available historical data suggest that, in the now-rich countries, direct mortality from these diseases was relatively rare immediately prior to introduction of these vaccines, except for tuberculosis. As many as half a million people contracted measles in the United States just before the vaccine was developed, for example, but measles directly accounted for fewer than 1,000 deaths. Exclusive of tuberculosis, reductions in these causes of death account for only 3 percent of the total mortality reduction. The reduction in tuberculosis mortality is another 10 percent, but in the United States, unlike most other countries, the tuberculosis vaccine has never been routinely used, so none of the reduction was due to vaccination. These conclusions parallel those of McKeown; the BCG vaccine for tuberculosis was widely used in Britain, but without any evidence of an effect on trend mortality. Apart from polio, the same is true for the introduction of other vaccines. Of course, the indirect consequences of eliminating infectious diseases may be greater; people with measles may succumb more readily to other diseases, for example. Evidence suggests there are indirect mortality effects for some water-born diseases (Cutler and Miller, 2005), but the extent of such indirect effects in the disease environment as a whole is not known.

Medical Treatments

Quantitatively more important for mortality was the development of new therapeutics for people with disease. Figure 3 shows mortality for infectious diseases and cardiovascular disease. Infectious disease declined greatly in the first half of the century, while cardiovascular disease mortality reductions were particularly important after 1960.

Antibiotics, developed in the 1930s and 1940s, were the first of the new wave of medical therapies. Sulfa drugs and penicillin were the wonder drugs of their era. By 1960, mortality from infectious diseases had declined to its current level.
More intensive medical interventions date in importance largely from the post–World War II era, and are associated with a different cause of mortality. Since 1960, cardiovascular disease mortality has declined by over 50 percent, and cardiovascular disease mortality reductions account for 70 percent of the seven-year increase in life expectancy between 1960 and 2000. Cutler (2004) matches the results of clinical trials to actual mortality declines, and attributes the bulk of the decline in cardiovascular disease mortality—as much as two-thirds of the reduction—to medical advance. Beyond medical advance, the major factor in reduced cardiovascular disease mortality is the reduction in smoking. Smoking rates in the United States have fallen to half their level at the time of the Surgeon General’s 1964 report on the harms of smoking. Continued public health campaigns against tobacco use have been an important part of this decline.

An additional 19 percent of the increase in life expectancy since 1960 is a result of reduced infant mortality. Cutler (2004) attributes a large share of continued infant mortality reductions to improved neonatal medical care for low birth-weight infants. The remainder of the decline in mortality since the 1960s includes reduced mortality from external causes, primarily motor vehicle accidents, reduced mortality from pneumonia/influenza, and a slight decrease in cancer mortality.

The importance of medical technology and smoking behavior can also be seen in the simultaneous decline and rapid convergence of mortality rates for all developed countries, particularly for men, and particularly for cardiovascular disease. Smoking causes cardiovascular disease, with a relatively short lag, and lung
cancer, with a much longer lag. In consequence, the increase in smoking among men in the second quarter of the twentieth century contributed to the slowdown in mortality decline in the third quarter of the century, and the reduction in smoking, which is now widespread throughout the rich countries, is currently acting in concert with technical progress in medicine. Women began to smoke later than men, and have been slower to quit, and women’s smoking rates are still rising in some European countries. As a result, the current gap between men’s and women’s life expectancy is low by historical standards; the decline and convergence of mortality from cardiovascular disease has been slower for women than for men; and women’s mortality rate from lung cancer is still rising in many countries, though it has recently reversed in some, including the United States.

Some analysts would continue to give nutrition the primary role in reducing life expectancy, even after 1870 and well into the twentieth century, rather than public health and medical care. In most countries of the world, although not the United States, people are continuing to get taller, and relationships that link mortality to physical characteristics will predict much of the recent mortality decline based on body size and function alone. Fogel (2004, p. 103) gives the personal health care system much credit for reducing morbidity (hip replacement, cataract surgery, and so on) but none for mortality decline: “The main thing that physicians do is to make life more bearable: reduce morbidity and tell people how to take care of themselves.”

The Long-Term Reach of Early-Life Factors

The “fetal origins” (or “womb with a view” hypothesis) of Barker (1990) posits that in conditions of nutritional deficiency, the developing fetus will differentially compromise functions that are operative only late in the life-cycle, beyond the normal age of reproduction, thus maximizing the chances of survival through reproductive ages and the number of offspring. The consequence for modern populations is that better nutrition decades ago could be having its effect only today.

The fetal origins theory is supported by the robust correlation between health in adulthood and birth weight, a marker for in utero nutrition (though a relatively poor one according to the theory, which stresses body shape more than weight), even when controlling for current socioeconomic status. A more compelling analytical approach is to examine late-life health of children who were in utero during famines. Children who survived in utero the brief Dutch famine at the end of World War II had higher levels of risk factors associated with coronary heart disease at age 50, though this is not true of mortality from heart disease itself. By contrast, no exposure effects were found for families in Finland in 1866–68 and in Leningrad during the siege of World War II (for a review of these and other results, see Rasmussen, 2001). Even if the theory is correct, it is likely that the effects of nutritional improvements in utero on cardiovascular mortality in adulthood are small compared to the effects of reductions in risk factors in adulthood.
The seasonality of life expectancy with respect to month of birth is another type of evidence. Doblhammer and Vaupel (2001) have shown a relationship between month of birth and longevity at age 50; those born in the northern hemisphere in October to December live about as much as 0.6 year longer than those born in April to June. As expected, the southern hemisphere is out of phase with the northern hemisphere by six months. After examining alternative explanations, such as selective infant mortality, these studies conclude the month-of-birth effect is most likely due to the seasonal availability of fresh fruit, vegetables and eggs to the pregnant mother in the first and second trimesters.

After birth, the environment during childhood, including disease prevalence and conditions at home, predicts the onset of disease in adulthood. For example, Case, Fertig and Paxson (2005) use the data from the 1958 British birth cohort to calculate that each chronic condition at age seven raises by 4 percent the probability of reporting a chronic condition at age 42; if the condition is still present at 16, the effect is twice as large. Lindeboom, Portrait and van der Berg (forthcoming) look at cohorts born in Holland between 1812 and 1912 and find that per capita GDP up to age seven is associated with large effects on mortality at ages above 50; in fact, the health effects from economic conditions at age seven are larger than the effects of contemporaneous macroeconomic conditions at ages 50 and above.

Overall, childhood factors such as nutrition and the disease environment have the potential to significantly affect mortality at older ages, although the magnitude is open to some debate. For this reason, some of the decline in mortality at the end of the twentieth century might be attributable to improvements in diet and public health many years ago, with the possibility of still more improvement to come.

Summary

Looking at this evidence as a whole, we see the history of mortality reduction as encompassing three phases. The first phase, from the middle of the eighteenth century to the middle of the nineteenth century, is the one where improved nutrition and economic growth may well have played a large role in health—although this is hotly debated, and incipient public health measures were certainly important as well. In the closing decades of the nineteenth century and into the twentieth, the second phase occurred, in which public health mattered more—first negatively, because of high mortality in cities, then positively in the delivery of clean water, removal of wastes, and advice about personal health practices. The third phase, dating from the 1930s on, has been the era of big medicine, starting with vaccination and antibiotics, and moving on to the expensive and intensive personal interventions that characterize the medical system today.

Determinants of Mortality in Poor Countries

Life expectancy is much lower and mortality rates are much higher in poor countries than in rich countries, as shown in Figure 1 and Table 1. There are also
marked differences in who dies and from what. In poor countries, 30 percent of deaths are among children, compared with less than 1 percent among rich countries. In rich countries, most deaths are from cancers and from cardiovascular diseases; in poor countries, most deaths are from infectious diseases. Most of these diseases are but a historical memory in rich countries; today they kill people in poor countries almost exclusively.

Yet there have been enormous improvements in life expectancy over the last half century in today’s poor countries. In India and China, life expectancies have risen by nearly 30 years since 1950 and, even in Africa, where there has been much less economic progress, life expectancy rose by more than 13 years from the early 1950s to the late 1980s, before declining in the face of HIV/AIDS. The worldwide decline in mortality after the World War II happened because 200 years worth of progress against mortality in the now-rich countries was rapidly brought to bear on mortality in the rest of the world. Measures such as improvements in water supply,
cleansing the environment of disease vectors (like *anopheles* mosquitoes that carry malaria or rats that carry lice), the use of antibiotics and the widespread immunization of children—the combined development of which had taken many years in the West—were introduced to the rest of the world over a relatively very short span of time. Because those who had previously died were mostly children, and because subsequent reductions in fertility followed only slowly (and in some countries not at all), this rapid deployment of life-saving public health led to the population explosion of the last half century.

Table 1 shows that there is a great deal more to be done before health in poor countries resembles that in rich countries today. As is also clear from the table, the problem is not primarily lack of suitable treatments. Diarrheal disease and respiratory infections—the first and fourth leading causes of death worldwide—are easily and cheaply treatable, with oral rehydration therapy (a mixture of salts and sugar that stops the dehydration that kills children with diarrhea) and with antibiotics. Malaria has been fully controlled in the rich world by environmental measures and can arguably be controlled by similar measures—although it will certainly be more difficult in low-income countries given the more difficult environmental conditions—or by the use of insecticide-impregnated bed nets in those countries. (Induced evolutionary changes in the *anopheles* mosquitoes will eventually make ineffective any given insecticide, and the outcome of the resulting arms race is unclear.) Cheap and effective antibiotics exist for most kinds of tuberculosis, though therapy must be maintained for a considerable period of time. The infectious “children’s diseases” of whooping cough, tetanus, polio, diphtheria and measles kill more than a million children each year, and all have been eliminated in rich countries by nearly universal immunization. Deaths within the first seven days of life are rare in the West, where pre- and post-natal health care are routinely available, but common in the rest of the world. The anti-retroviral drugs that have controlled mortality from HIV/AIDS in the rich world are expensive and not generally available in sub-Saharan Africa, not even in South Africa which is by far the wealthiest country in the region.

Of course, the fact that treatments already exist for many conditions does not deny that new technology could be valuable. It is speculated that vaccines could be developed for many of the key killers in developing countries (especially AIDS, tuberculosis, and malaria; see Kremer, 2002), which would make disease prevention much easier. Easier-to-use therapies could also be important, to the extent that the difficulty of using therapies such as anti-retroviral medications explains their low use. However, cheap and easy-to-administer treatments that are already available for many diseases are not being used.

These diseases themselves are the result of other risk factors and disease exposures. The World Health Organization (2003) has identified a set of risk factors for mortality in poor countries. Included in the risks are unsafe sex (certainly important for HIV/AIDS), unsafe drinking water (one cause of diarrheal disease) and a variety of other factors such as malnutrition and indoor smoke from burning solid fuels (important in respiratory conditions). This list of underlying
factors is—more or less—the right one, but the quantitative magnitude of the particular factors is unknown. In thinking about ways to improve health in poor countries, we focus less on particular risk factors and more on the overall medical and economic environment that can affect those factors.

Health delivery is often of low quality in both public and private sectors. Absenteeism among medical staff is often a problem (with a third or more not showing up for work), particularly in rural areas (Chaudhury, Hammer, Kremer, Muralidharan, and Rogers, 2005). Recent surveys in India have shown that, while public doctors are more likely to be qualified, they are also more likely to be absent and to have insufficient time or medicines to provide effective treatment. Private providers are often ill-qualified, and face competitive pressure to overtreat: for example by giving everyone an injection of antibiotics without any prior testing (Das and Hammer, 2004). Countries which are unable to provide effective public health care are often also those that do not have the institutional ability to regulate and to monitor the private sector. At the same time, many countries spend so little on health care that, no matter how organized, it is unlikely to be effective.

Paradoxically, many consumers in these low-income countries report that they are well-satisfied even with objectively unsatisfactory provision, so that there is little political demand for improvement. Whatever the reasons, many countries cannot deliver the cheap, effective, and widely available drugs that currently exist—a fact often used as a counterargument against those blaming antiretroviral-drug patents for the lack of success in treating HIV/AIDS in Africa.

Many of the most successful health programs in poor countries—such as immunization campaigns, the (successful) eradication of smallpox and the (close to successful) eradication of polio—have been “vertical” campaigns run from outside the country by international organizations such as the World Health Organization (WHO) or UNICEF. Some critics argue that, although these programs have been successful, they have also done little to improve (and may bid resources away from) the domestic health care systems on which further progress in reducing mortality may arguably depend. However, it seems unlikely that health inputs would be so inelastic in the long run. More consequentially, there is some evidence that the international immunization campaigns have run out of steam in recent years. They are less well-funded than in the past, and past campaigns may have successfully targeted the easiest-to-reach segments of the population (Bloom, Canning and Weston, 2005). Perhaps in consequence, there has been a worldwide slowdown in the rate of reduction of infant mortality in the 1990s compared with the 1980s (Ahmad, Lopez and Inoue, 2000).

Broader social factors are also important for reductions in mortality. In his pioneering work, Preston (1980) attributed about half of the gain in life expectancy in developing countries (excluding China) from the 1930s to the late 1960s to the combined effects of changes in income, literacy and the supply of calories, although the last was not significant in his regressions. He attributed the rest of the gain to the public health measures newly implemented in the third world, although he recognized the difficulties of attribution, if only because of likely interactions;
income or education may facilitate the adoption and effectiveness of some public health measures.

The importance of education, particularly women’s education, has been confirmed in many subsequent studies. The importance of women’s education is likely a result of the fact that, as primary caretakers, women are most likely to implement the behaviors that can improve their children’s health. To the extent that education improves an individual’s ability to undertake these changes, more educated mothers will have healthier babies.

The role of economic growth in health improvements in poor countries has been as controversial as it is in the history of mortality decline. If Figure 1 were a causal relationship, it would show that the effects of income on health are strong at low levels of income, where absolute deprivation (including lack of food and clean water) is common. Such income-based explanations emphasize the nutritional factors brought up in the historical account, as well as the fact that higher income makes it easier to provide the infrastructure of public health, such as water and sanitation. In recent years, a number of authors have followed Pritchett and Summers (1996) and argued from cross-country regressions that income is more important than any other factor, and have endorsed policies that downplay the role of any deliberate public action in health improvement. According to this view, if countries are growing, the health of their inhabitants will look after itself. As was certainly intended, Pritchett and Summers’s title, “Wealthier is Healthier,” has become a banner under which some economists defend economic liberalization against claims by the public health community and others that it has harmed health.

Yet the cross-country data show almost no relationship between changes in life expectancy and economic growth over 10-, 20-, or 40-year periods between 1960 and 2000. Many countries have shown remarkable improvements in health with little or no economic growth, and vice versa. For the two largest countries, India and China, there is a negative correlation between decadal rates of economic growth and progress in reducing infant and child mortality. Almost all of China’s remarkable post–World War II reduction in infant mortality happened prior to the acceleration in economic growth after 1980, after which there was relatively little progress in child health. Similarly, in India, the acceleration of the rate of growth after the economic reforms in the early 1990s was accompanied by a slowdown in the rate of decline in infant mortality (Drèze and Sen, 2002, chapter 4). Drèze and Sen also argue that the slowdown in progress in China was a direct result of the change in policy and the switch in resources that generated the growth.

As with the historical record, then, the cross-country evidence does not suggest that economic growth will improve health without deliberate public action. This may seem paradoxical if only because income brings so many things that favor better health for the poor: better nutrition, better housing, the ability to pay for health care, as well as the means for the public provision of clean water and sanitation. There are a number of possible hypotheses as to why income is not more important. As we have seen in the historical account, income growth and health are
not always associated. As was the case in Europe, economic growth has been accompanied by urbanization in much of the poor world, and with some of the same consequences that attended urbanization in Europe. This seems unlikely to be the explanation for India or China, however.

More importantly, nutrition and housing may have limited effects without macro public health measures, which require political action. Income can only buy so much if the disease burden is overwhelming. Some rapidly growing economies have not provided a good public health environment—indeed, have substituted away from it in their attempt to promote economic growth. Other countries that rely on more of a command and control economic system have used their command over labor to undertake public health measures that might not be feasible in a more democratic state. Examples from China range from the coerced mobilization of whole villages to deal with health threats or pests, such as mosquitoes, to the one-child policy itself. Similarly, Cuba has a program of local doctors that is the envy of many countries, even as its overall economy is in shambles. Ironically, the weakness of their economy as a whole may make it easier for some countries to afford the distribution of resources to health care.

Finally, and on the opposite side of the argument, there is an old view, recently endorsed by Acemoglu and Johnson (2005), that improvements in health technology and the associated reduction in child mortality should reduce GDP per head, at least temporarily, if health innovations result in large increases in population. Since economic growth and health improvements are almost uncorrelated in the data, the negative effects of health improvements through increased population must have been almost completely offset by some positive effect of economic growth on health. If this story is right, growth does indeed improve health, but the effect has been hidden for much of the post–World War II period by the negative effects of population growth on income per head.

Determinants of Mortality within Countries

A vast literature shows that individuals with low income, low wealth, low education, or low social status often die younger than those who are better off or better educated; and this is true for many countries and for many (if not all) periods. The British census of 1851 showed differences in mortality across (occupationally defined) social classes, with those in lower (manual) classes having higher mortality than skilled workers or professionals (Macintyre, 1997). More recently, the famous study of Whitehall civil servants in Britain shows a difference in mortality rates across groups defined by their civil-service ranks (Marmot et al., 1991); all-cause mortality diminishes with rank, as does mortality from most causes, although the effects are much stronger for cardiovascular disease than for cancer.

In the United States, an array of studies has found similar patterns by income, education, and race (for a compendium, see Rogers, Hummer and Nam, 2000). The National Longitudinal Mortality Study, which matches death certificates with
earlier data from the Current Population Survey, shows inverse (partial and total) correlations between both education and income and mortality (Elo and Preston, 1996), as well as correlations between mortality and race, urban/rural residence, and other factors.

These socioeconomic differences in health extend even to babies. White infants of mothers with less than twelve years of education have a mortality rate that is twice as high as that of white infants of mothers with a college degree (10 per 1,000 versus less than 5 per 1,000). Infants of black mothers have higher mortality rates than whites for every education level—furthermore, children of black mothers with a college degree have higher mortality rates than children of white high school dropouts (Pamuk, Makuc, Heck, Ruben and Lochner, 1998, Figure 9). Income gradients in nonfatal health begin in early childhood, and grow larger as the child moves into adulthood (Case, Lubotsky and Paxson, 2002).

Similar “gradients”—the term is used to emphasize that there are “graded” differences in health running across ranked groups, not just between poor and rich—are found in Canada and in European countries. Although data on adult mortality are lacking in many poor countries, the World Bank has documented a strong negative link between infant and child mortality and an index of living standards based on the ownership of durable goods, and sometimes more directly on income or consumption. In most places, mortality differences by social class are particularly well-defined for cardiovascular disease and for lung cancer. Mortality differences are a good deal less sharp for other cancers, and are reversed for breast cancer in women, where highly educated, high-income women are more likely to die.

The elimination or at least reduction of differences in health by income, race or geography has become a major focus of health policy in many countries, including the United States and Britain. Our concern here is why these inequalities exist and whether their existence is consistent with our accounts of historical and contemporaneous mortality decline, and with differences in mortality between rich and poor countries.

Medical Care

One possible answer is that those with high incomes receive more health care. Health insurance is related to income in the United States, and while health care coverage is universal in most other countries, better and less well off have access to different physicians and sometimes hospitals. For example, in the United States, standards of care appear to be lower in hospitals that mostly treat blacks (Bach et al., 2004; Skinner, Chandra, Staiger, Lee and McClellan, 2005).

But access to health care cannot explain everything. As several studies show, including the Whitehall study in the United Kingdom and the Health and Retirement Study in the United States, the incidence of adverse health conditions is higher among those of lower rank or lower education, even before the health care system has become involved. Moreover, some large changes in access to health care have had only minor effects on health gradients. The introduction of Medicare in 1965
had no clear effects on the mortality of the elderly (Finkelstein and McKnight, 2005), and no effect on U.S. relative to British mortality rates for the relevant age groups (Deaton and Paxson, 2004). Strikingly, Britain’s class-based differences in health survived the introduction of the National Health Service after World War II.

**Resources**

An alternative theory of resources is that money matters because of the non–health care things it can buy. This theory may have made more sense in the past, when adequate food, clothing and shelter were constant struggles, but it makes less sense today, at least in rich countries. Indeed, access to cheap food is a risk factor for poor health in the United States and many countries (Cutler, Glaeser and Shapiro, 2003). There is no evidence that, as living standards rise, the health gradient disappears. In fact, according to some measures, health gradients appear to be increasing in both the United States and Europe.

**Differences in Health-Related Behaviors**

More educated people are less likely to smoke, and this difference has increased over time; between the mid-1970s and the mid-1990s, the difference in prevalence of smoking between high-school graduates and college graduates grew from about 9 percentage points to 15.5 percentage points (Pamuk, Makuc, Heck, Reuben and Lochner, 1998). Smoking is a substantial factor in differences in lung cancer and cardiovascular disease mortality across education groups. Drinking, exercise, eating habits, use of preventive care (such as annual mammography), adherence to therapy, and other health behaviors are also correlated with measures of socioeconomic status (Adler et al., 1994; Goldman and Smith, 2002).

But again, observed behavior is not everything. Health gradients by socioeconomic status persist even when differences in smoking, drinking, and other factors are taken into account (Marmot, 1994). In the study of Whitehall civil servants, looking at nonsmokers only eliminates the mortality differences between the top and bottom groups for lung cancer but not for coronary heart disease. Over a ten year follow-up, the mortality ratio between the top and bottom group was reduced from 2.7 times to just over twice by adjusting for a range of risk factors, including smoking. In the United States, non-Hispanic Caucasian Americans are more likely than blacks to have ever smoked or to have smoked heavily (Rogers, Hummer and Nam, 2000, p. 245), so smoking does nothing to help explain the black–white differences in U.S. mortality patterns.

Moreover, a behavioral explanation for gradients in mortality does nothing to explain why people from different socioeconomic groups behave differently. Economic theories of differences in health behaviors across groups (pioneered by Grossman, 1972) are generally based on differences in information, prices, the value of long life, or discount rates. Information differences between those of different socioeconomic status are an easy explanation, but less promising when examined closely; to take but one example, knowledge about the harms of smoking is nearly universal in the United States. Prices, too, are similar for rich and poor.
Social Structures, Stress and Health

Outside of economics, the currently dominant theory of health differentials is that the poor health of low status individuals is caused by “psychosocial stress”—the wear and tear that comes from subordinate status and from having little control over one’s own life. This account is heavily influenced by both the Whitehall evidence, and by accounts of rank and health within other primates. For example, Sapolsky’s work on baboons in Kenya (1993) shows that subordinate baboons have worse levels of various markers of chronic stress, such as glucocorticoids, and are in poorer health. Furthermore, the stress-related symptoms emerge after hierarchies become established and change when the hierarchy changes, suggesting they are related to an individual’s rank rather than to fixed individual characteristics such as genetic traits.

Some biological evidence supports this theory. The mechanism that helps animals deal with stress in the wild, “the flight–fight response,” is a series of short-run responses that help save the animal’s life from an immediate threat, at the expense of other functions relevant for long-term survival. Individuals who are in low-status and subordinate situations, who are subject to arbitrary demands by others, or who are discriminated against because of their race, are continually having these biochemical responses triggered in a way that eventually causes permanent malfunction, a buildup of what is known as “allostatic load” (Seeman, Singer, Rowe, Horwitz and McEwen, 1997). This cumulative distress leads to an increased probability of disease, particularly cardiovascular disease.

One concern with this work is that the concept of socioeconomic status is often a convenient catchall for a range of variables—including income, education, occupation and race—but it is not helpful for thinking about how these variables might have separate effects on health. Nor is grouping these variables together helpful for policy analysis, which requires knowing which variables to alter.

A second concern is that the relationship between socioeconomic status and health must work in both directions. Income is a case in point—there is substantial evidence that poor health leads to low income, rather than the other way around. In the United States and elsewhere, ill-health is a leading reason for retirement (Smith, 1999, 2005) or for dropping out of the labor force (Case and Deaton, 2003), each of which are typically accompanied by a substantial drop in income. Conditional on education, which acts as a form of protection against new episodes of illness, changes in income do not predict changes in health, and lagged income does not predict future incidence of ill health (Smith, 2003; Adams, Hurd, McFadden, Merrill and Ribeiro, 2003). Similarly, if income were the main factor, it would be difficult to explain why mortality fell most rapidly in the United States in the period after 1970, during which median real income growth had slowed to a crawl, or why it is that different European countries, with different economic performance in the post–World War II period, should have such convergent experiences of mortality decline (Deaton, 2004; Deaton and Paxson, 2004). The behavior of health and income over the business cycle is also inconsistent with a strong effect of income on health; Ruhm (2000) documents that recessions actually improve
health, because individuals are more likely to exercise, and less likely to drink, smoke or engage in other health-damaging activities during downturns.

The effects of education are more consistent than the effects of income with theories that health is determined by socioeconomic status. Looking at the United States, Currie and Moretti (2003) find that women in counties where colleges opened were more likely to attend college and had healthier babies. Lleras-Muney (2005) finds that the populations of states that first enacted compulsory schooling laws subsequently lived longer; Oreopolous (2003) also finds that increases in minimum schooling laws in England and Ireland improved the health of the population. Also, as noted earlier, maternal education is strongly inversely correlated with infant and child mortality in developing countries.

Education is likely to provide general human capital that can be used to maintain and improve health in a wide range of circumstances. As emphasized by the “fundamental causes” literature (Link and Phelan, 1995), educational differences (like other forms of power differences) will maintain a gradient in health whenever there exists a mechanism or technology that more knowledgeable and educated people can use to improve their health. Such explanations, unlike psychosocial stress, help to explain shifting gradients over time in specific diseases—for example, that lung cancer and cardiovascular disease were once relatively more common among the relatively more educated population. The explanation also predicts that, if breast cancer screening becomes more effective, it will diminish or reverse the current gradient where highly educated women are more likely to die of breast cancer (Link, Northbridge, Phelan and Ganz, 1998).

But as was the case for income, there is also evidence of a reverse relationship running from health to education, certainly among children, and poor health in childhood may predict poor health later. Case, Fertig and Paxson (2005) find that children who experienced poor health in childhood entered adulthood with significantly lower educational achievement. Miguel and Kremer (2004) and Bleakley (2002) find that provision of deworming drugs significantly improved schooling in contemporary Kenya and the pre–World War II American South, respectively. But it is not clear how much of the observed relationship between education and health in adulthood can be explained by the fact that children in poor health obtain fewer years of schooling.

Summary

The link between social status and health is complex, perhaps too complex for a single explanation. It seems clear that much of the link between income and health is a result of the latter causing the former, rather than the reverse. There is most likely a direct positive effect of education on health. While the exact mechanism underlying this link is unclear, stress and the differential use of health knowledge and technology are almost certainly important parts of the explanation.

These cross-sectional findings have implications for our time-series analysis as well. If better education leads to better health, some of the post-1970 decrease in mortality in the United States and elsewhere might be attributable to the large
increases in the average education of the population, with correspondingly less attributed to medical care.

**Conclusion**

What sense can we make of all of these disparate accounts in different contexts, and what can we expect for the path of mortality in the future? There is no consensus on these issues. Here, we hazard our own best guess, recognizing that the evidence is weak or missing for many of the links in our argument.

Knowledge, science and technology are the keys to any coherent explanation. Mortality in England began to decline in the wake of the Enlightenment, directly through the application to health of new ideas about personal health and public administration, and indirectly through increased productivity that permitted (albeit with some terrible reversals) better levels of living, better nutrition, better housing and better sanitation. Ideas about the germ theory of disease were critical to changing both public health infrastructure and personal behavior. Similarly, knowledge about the health effects of smoking in the middle of the twentieth century has had profound effects on behavior and on health. Most recently, the major life-saving scientific innovations in medical procedures and new pharmaceuticals have had a major effect, particularly on reduced mortality from cardiovascular disease. There have also been important health innovations whose effect has been mainly in poor countries: for example, the development of freeze-dried serums that can be transported without refrigeration, and of oral rehydration therapy for preventing the death of children from diarrhea.

Perhaps controversially, we tend to downplay the role of income. Over the broad sweep of history, improvements in health and income are both the consequence of new ideas and new technology, and one might or might not cause the other. Between rich and poor countries, health comes from institutional ability and political willingness to implement known technologies, neither of which is an automatic consequence of rising incomes. Within countries, the lower earnings of people who are sick explain much of the correlation between income and health, rather than a causal relationship from higher income to better health.

There seems no reason to suppose that the flow of health-enhancing knowledge and technology will slow. Indeed, there are enormous incentives for the discovery of new basic health-enhancing knowledge, as well as for the development of new drugs and new medical treatments. Richer people are prepared to pay more for longer lives, and people who live longer are prepared to pay more to cure diseases, such as Alzheimer’s, that few people used to live long enough to contract. Of course, the pace of progress is hard to predict. Optimistic assessments can sound truly fantastic. Oeppen and Vaupel (2002) show that in the 160 years from 1840, life expectancy in the leading country or region of the world has increased by three months per calendar year. If this trend continues, the leading country will have a
life expectancy at birth of 100 by the middle of this century, and even a laggard like the United States will get there before the century is out.

However, changes in knowledge, science and technology will often increase the gradient in health, at least for a time. There was no health gradient between English aristocrats and ordinary people prior to the Enlightenment, but one developed soon thereafter, so that average life expectancy and the gap between rich and poor rose together. There was no gradient in infant mortality between the children of physicians and non-physicians prior to an understanding of the germ theory of disease. More educated people quit smoking faster after the health consequences were understood. Our hypothesis is that greater speed of introduction of new health-relevant knowledge and technology will tend to raise the health gradient, a hypothesis that is consistent with rising gradients in rich countries in the recent past.

If our analysis of the gradient is correct, our prediction of an acceleration in the production of new knowledge and new treatments is likely to make the health gradient steeper, with increasing gaps across educational and social class (occupational) groups, and possibly race as well. Gaps between countries may also widen. The incentives for research and discovery are much weaker or absent for the diseases, such as malaria or tuberculosis, that are largely confined to the poor of the world. Even when treatment is available in rich countries, there is no guarantee that it can be made available elsewhere, as we have learned during the AIDS pandemic and indeed from the several million people who die each year from vaccine-preventable diseases. Steepleing gradients within and between nations are likely to provoke much soul-searching, and it is clearly an appropriate aim of public policy to improve equality of access for everyone to new, life-saving technologies. Yet, if we are right, increases in the gradient also have a silver lining. They indicate that help is on the way, not only for those who receive it first, but eventually for everyone.

A longer version of this paper, which includes an extended bibliography, is available as a National Bureau of Economic Research working paper at (http://www.nber.org). We are grateful to Jeffrey Hammer and to participants in the NIA-sponsored workshop on the Determinants of Mortality, Princeton, July 16, 2004. The views here are those of the authors and not of the participants. David Cutler and Angus Deaton gratefully acknowledge financial support from the NIA through NBER, Grant No. P01 AG05842-14. David Cutler acknowledges support from the NIA through grant No. R01 AG015911 and Angus Deaton acknowledges support from the NIA through Grant No. R01 AG20275-01.
References


