

Detecting the effect of medical care on mortality

Ralph Catalano,^{a,*} John Frank,^{a,b}

^a320 Warren Hall, School of Public Health, University of California, Berkeley, Berkeley, CA, 94720, USA.

^bUniversity of Toronto, Canadian Institute for Advanced Research, (Population Health Program)

Received 17 August 2000; received in revised form 8 November 2000; accepted 17 January 2001

Abstract

Objectives: To demonstrate an important limitation of empirical tests of the proposition that medical care has contributed to long-term declines in mortality. Quasi-experiments cannot detect the effect of care that sustain, rather than change, a downward trajectory. We demonstrate this limitation by testing two hypotheses. One is that isolation of patients and antibiotic treatment coincided with declines in tuberculosis (TB) mortality in Massachusetts between 1850 and 1950. Another is that the introduction of Medicare in the 1960s increased life expectancy at age 65. **Results:** The first hypothesis is supported, for both patient isolation and streptomycin. The second is not. The circumstances that could yield such results are explored. **Conclusions:** Epidemiologists and historians should cooperate to devise methods that can resolve the issue of whether medical care has sustained the steady downward trend in mortality witnessed over the last century-and-a-half. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Medical care; Mortality; Life expectancy; Tuberculosis; Time-series

1. Introduction

The growing debate over health insurance and access to medical care has brought new attention to an old question. That is, how much has medical care contributed to declining mortality over the last century-and-a-half [1–8]? Research has failed to converge on an answer. Among the reasons for this failure is that there is no agreed-upon methodology for estimating the effect of medical care on mortality. Indeed, the methods used in the literature have ranged from classical historical argument [5] through descriptive epidemiology [6,7] to multivariate modeling [4,8].

Further attempts to quantify the historical benefit of health care will probably introduce yet more methods into the debate. The time-series quasi-experiment is likely to be prominent among these because the economists and evaluators likely to be attracted to the task often apply the method in their work. Medical researchers, on the other hand, have traditionally exhibited ambivalence toward the method [9]. We suspect that this ambivalence arises primarily from a lack of familiarity with a method that appears far removed from the clinical trial. We hope to make readers more familiar with the method by clarifying its connection to experimental logic and applying it to questions concerning the effect of medical care on mortality.

Our intent is not to convert ambivalence to acceptance, but rather to better inform opinion about the method. In fact, there may be good reason for medical researchers to be ambivalent about using the method to assess medicine's contribution to declining mortality. We illuminate this circumstance in two steps. The first is to establish the sensitivity of the method by testing the hypothesis that two specific medical interventions, patient isolation in sanatoria and streptomycin treatment, each contributed to the decline in cause-specific mortality from tuberculosis (TB) in Massachusetts between 1850 and 1950. The second step is to test an hypothesis more germane to the debate over whether expanded access to medical care has saved lives. That is that the Medicare program increased life expectancy for persons aged 65 when introduced in the 1960s.

2. Time-series quasi-experiment

We believe that clinically oriented researchers concerned with the historical effects of medical care on mortality will be attracted to the time-series quasi-experiment if its connection to the clinical trial is made clear. Both methods share the logic that an intervention cannot be assumed effective unless the observed values of the outcome differ, as intended, from those expected under the null hypothesis (i.e., under the assumption that the intervention had no effect). The clinical trial derives values expected under the null hypothesis from an unbiased control group. The princi-

* Corresponding author. Tel.: 510-642-3103; fax: 510-642-6981.

E-mail address: rayc@uclink4.Berkeley.edu

pal challenge in any quasi-experiment is to arrive at the expected value under the null hypothesis without a control group.

Most quasi-experiments use statistical assumptions and manipulations to derive values expected under the null hypothesis. The simplest assumption is that the mean of pre-intervention observations will be observed post-intervention if the intervention had no effect. Time-series, however, typically exhibit the tendency, referred to as autocorrelation, for succeeding values of a variable to be correlated with each other. This tendency complicates quasi-experiments because the expected value of an autocorrelated series is not its mean.

Mortality rates and life expectancy have been observed to exhibit several types of autocorrelation. “Epidemiological momentum,” for example, is induced by long and highly variable latent periods between exposure to risk factors and the onset of symptoms [10]. Biological and seasonal phenomena have also been suspected of inducing cycles in mortality [11,12]. Mortality and life expectancy also exhibit the tendency to oscillate after unusually high or low values because these extremes imply that deaths among frail and unwell persons were accelerated or delayed [13,14]. Oscillations set in because accelerations deplete the pool of susceptibles while delays augment it.

Methods for deriving expected values under the null hypothesis have been developed for quasi-experiments that use autocorrelated time-series data [15]. These methods use statistical manipulation to “decompose” a time series, such as life expectancy, into its expected and unexpected components, based on the patterns observed in the series. The test then hinges on whether the observed values are outside, in the intended direction, the 95% confidence interval of the expected values in the time periods specified *a priori* as those in which the intervention should have taken effect. If they are, the data are judged consistent with an effect of the intervention.

3. Example: TB mortality in Massachusetts

We demonstrate the sensitivity of this approach by testing the hypothesis that two medical innovations, widely separated in time, affected TB mortality in Massachusetts between 1850 and 1950 [16]. Newsholme, working in the UK in 1903–08, offered striking data consistent with the claim that isolating patients in work-houses, although not intended as a public health intervention, hastened the decline of TB mortality [17,18]. If there were such an effect in the US, it would most likely have appeared in the decade prior to World War I (1905–1914), when TB sanatoria were built and filled at a rapid rate. Bryder estimates that the US had only 34 such facilities in 1900 and some 536 by 1925 [19]. For the purposes of demonstrating our method, we specified that isolation should have caused a downward shift in TB mortality sometime in the 5-year period 1909 through 1913. Shifting the interval later in time would cre-

ate an overlap with the war years and flu pandemic, both of which might be expected to alter TB mortality [20].

We further demonstrate the sensitivity of time-series methods by examining the trend in TB mortality after the introduction of streptomycin. Historical commentators universally assert that this treatment greatly reduced TB death rates [6,17–21]. Empirical support for this assertion, however, is limited to the observation that logarithmically transformed annual TB mortality rates exhibit a slope change after World War II [17,18]. The exact date at which this initially scarce antibiotic became routinely available for TB patients is poorly documented and probably varied by place. Based on a review of international data, Lancaster infers that the effect of antibiotics on TB mortality could not have begun before 1946 [22]. We therefore would expect to see the onset of a sustained mortality decrease in the years 1946 through 1950.

The above considerations imply the following hypotheses.

H1: TB mortality rates in the years 1909 through 1913 will be lower than expected from the rates observed from 1850 to 1908.

H2: TB mortality rates in the years 1946 through 1950 will be lower than expected from the rates observed from 1850 through 1945.

We use Massachusetts TB mortality to demonstrate our method because the series has been widely cited as a source of evidence that medical innovations had an effect on TB mortality [16]. Fig. 1 shows a time plot of these data.

We transformed the annual TB mortality rates into natural logarithms to make their distribution more normal. We arrived at the expected values of TB mortality for the years 1909 through 1913 and 1946 through 1950 in two steps. We first identified the time-series models that best fit the data for the periods 1850 through 1908, and 1850 through 1945. We then used these models to forecast values for 5 years (i.e., 1909 through 1913 and 1946 through 1950).

We reject the null hypothesis of no association between the interventions and TB mortality if two conditions are met. The first is that more than one of the observed values during the 5-year period is below the lower 95% confidence bound of the expected values derived from the forecasts. The second is that there be no immediate “rebound” or observed value above the upper boundary of the confidence interval. This condition guards against the circumstance in which a temporary “postponement” of TB mortality is confused with a lasting effect of the intervention in question. Rejecting the null hypotheses establishes a statistical precondition for the assertion that an innovation had an effect. The possibility would remain that a third variable induced a coincident reduction in mortality.

3.1. Modeling the TB mortality series

We used the model-building and forecasting methods attributed to Box *et al.* to implement our test [23]. The strat-

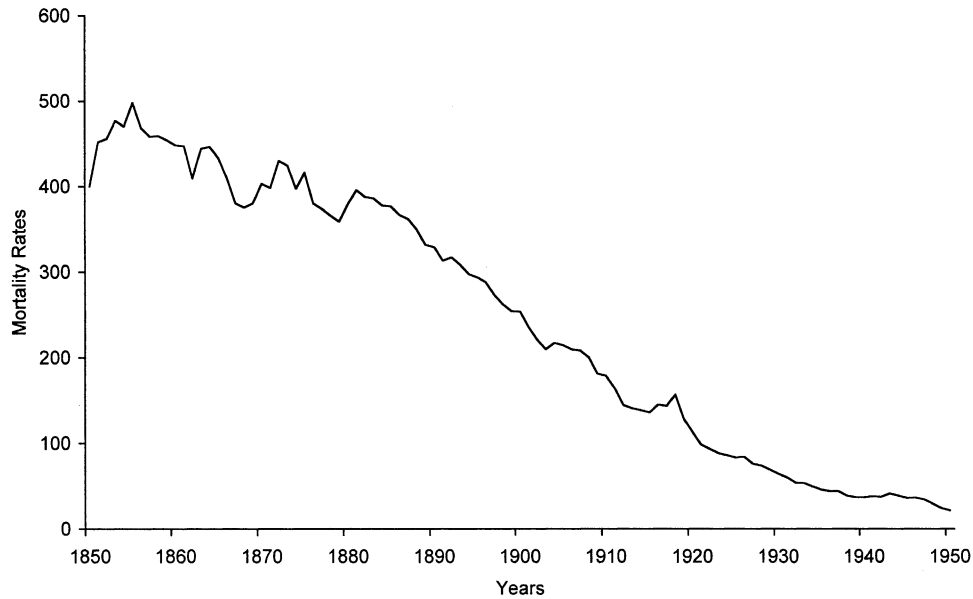


Fig. 1. TB mortality rates (per 100,000) in Massachusetts for the years 1850 through 1950.

egy, Auto Regressive, Integrated, Moving Average (ARIMA) modeling, allows any of a large family of possible models to be empirically fit to a time-series.

We tested Hypothesis 1, that Massachusetts TB mortality rates in the years 1909 through 1913 would be lower than expected from the rates observed from 1850 through 1908, in three steps. The first was to estimate the following ARIMA equation for the years 1850 through 1908.

$$\nabla_d Y_t^e = c + \frac{(1 - \theta B^q)}{(1 - \phi B^p)} a_t \quad (1)$$

∇_d is the difference operator that indicates a series was differenced at lag n to remove secular trends or cycles (i.e., to render the series stationary in its mean).

Y_t^e is the natural logarithm of the TB mortality rate for Massachusetts.

c is the Y intercept or mean of the series.

ϕ is an autoregressive parameter. Autoregressive parameters measure a series' tendency to remain above or below its means after a perturbation.

θ is the "moving average" parameter. Moving average parameters measure the tendency of perturbations to be present for more than one year.

B is the "backshift operator" or value of the conditioned variable at year $t - q$ or $t - p$

a_t is the error term at year t .

The second step was to use the above equation to forecast the values for the years 1909 through 1913 and to derive the 95% confidence intervals of the forecasts. The third step was to determine if the observed values were below the lower bound of the 95% confidence interval.

The second hypothesis, that TB mortality rates in the years 1946 through 1950 would be lower than expected from the rates observed from 1850 through 1945, was simi-

larly tested. The principal difference was that the model for Hypothesis 2 had to include the effect of the flu pandemic of 1919 and any effect of isolation discovered in testing Hypothesis 1.

Two flu-related variables were added to Eq. 1. The first was a binary variable scored 1 for 1919 and 0 otherwise. The parameter for this variable measures the abrupt increase in TB mortality associated with the flu pandemic. The second variable was also a binary variable but scored 1 in 1920 and 0 otherwise. The specified parameter structure for this variable assumes that the abrupt increase in TB mortality in 1919 depleted the vulnerable population that otherwise would have died in subsequent years [20]. This population would have included the chronically ill, including many TB sufferers, as well as those not previously exposed to a related influenza strain. The parameter structure assumes that TB mortality in 1920 will be depressed and gradually increase to a level unaffected by the pandemic.

The estimated equation for the second hypothesis is as follows:

$$\nabla_d Y_t^e = c + \omega_1 I_{1t} \frac{\omega_2}{1 - \delta_1 B} I_{2t} + \omega_3 I_{3t} + \frac{(1 - \theta B^q)}{(1 - \phi B^p)} a_t \quad (2)$$

I_1 is a "flu variable" scored 1 for 1919 and 0 otherwise.

ω_1 is an effect parameter that estimates the excess TB mortality associated with the flu pandemic of 1919.

I_2 is a variable scored 1 for 1920 and zero otherwise.

ω_2 is an effect parameter that estimates the reduction of TB mortality associated with the depletion of vulnerable persons during the flu pandemic.

δ_1 is the "drag" parameter that estimates the proportion of ω_2 that is carried from 1920 to 1921 and beyond.

I_3 is a variable scored 1 for any years between 1909 and 1913 in which an effect of patient isolation and 0 other-

wise. This variable is included in the equation only if effects of isolation are discovered in the test of Hypothesis 1.

ω_3 is an effect parameter that estimates the reduction of TB mortality associated with patient isolation.

3.2. Results of the TB mortality demonstration

The best fitting ARIMA model for TB mortality during the years 1850 through 1908 was as follows.

$$\nabla Y_t^c = (1 + 0.2968B^{12})a_t \tag{3}$$

The model suggests that the obvious downward trend in TB mortality included a less-apparent tendency for unusually high or low values to be followed 12 years later by similar, but less pronounced, peaks or valleys. The statistical efficiency of the model is excellent in that the trend and 12-year cycle parameters account for 97.2% of the variance in the original data.

As shown in Table 1, four of the observed values were below the 95% confidence interval expected from history. There were no values above the confidence interval. We reject the null hypothesis that the observed values of TB mortality were not different from the expected values. There is, in other words, a statistical association between patient isolation and the decline in TB mortality in the 5 years preceding World War I.

The best fitting ARIMA model for the years 1850 through 1945 is shown below.

$$\nabla_d Y_t^c = -0.0174 + 0.1025I_{1t} + \frac{-0.1397}{1 - 0.7736B}I_{2t} - 0.0652I_{3t} + (1 + 0.2479B^2 + 0.3464B^{12})a_t \tag{4}$$

All the parameters were at least twice their standard errors. The residuals exhibited no autocorrelation. The model accounted for 99.2% of the variance in TB mortality over the tested period.

The best fitting model for the years 1850 through 1945 was more complex than that for the initial period. In addition to the downward trend and 12-year cycle, the data exhibited weak oscillations such that a high or low value was followed by an opposite, though much reduced, movement 2 years later. The “1918 Spanish flu effect” was also strong and persisted for at least 2 years.

Table 1 shows that the years 1949 and 1950 had rates below the confidence interval of the expected values. We reject the null hypothesis of no statistical association between the introduction of antibiotics and reduced TB mortality in Massachusetts. The effect appears to begin in 1948 when the observed value (i.e., 29.4) falls below the point estimate (i.e., 32.2) of the expected value.

A curious pattern detected in the TB mortality time-series is the presence of a 12-year moving average, implying that perturbations in mortality have tended to “echo” after a dozen years. While we can only speculate on the cause, we note that historical plots of TB incidence and mortality ver-

Table 1

Point estimate (i.e., if intervention had no effect) and observed TB mortality rates per 100,000 persons (Confidence interval (95%) of point estimate is also shown)

Innovation and year	Point estimate of			Observed TB mortality rate
	Lower bound of point estimate	TB mortality rate	Upper bound of point estimate	
Isolation				
1909	181.6	197.2	214.2	181.5
1910	173.6	195.1	219.2	179.1
1911	167.9	193.7	223.4	164.1
1912	164.2	193.6	228.3	144.4
1913	158.4	190.5	229.0	141.2
Antibiotics				
1946	32.1	34.9	38.0	36.7
1947	29.2	32.9	37.1	34.5
1948	27.4	32.2	37.8	29.4
1949	26.4	31.9	38.7	24.1
1950	24.4	30.4	37.8	21.5

sus age tend to show two peaks, one in the preschool period and another beginning in the late teens, often extending to young adults in their twenties. We have therefore a historical distribution of TB-latent periods in early life, with a mean of perhaps a dozen years, albeit with a large standard deviation, as surviving children infected before age five reach an age of high re-activation risk. Perhaps perturbations in TB incidence in young children, which would have been closely accompanied by mortality changes prior to the antibiotic era, presaged analogous “echo” changes in teenage incidence and death.

We used the longest time-series of TB mortality in the US available to us. It could, however, be argued that using a long series favored a type 1 error because the confidence interval of the forecasts was narrowed by having many time points before 1909. We repeated the tests with the fewest pre-1909 years that allowed use of the *t*-statistic for modeling autocorrelation (i.e., 30 years). The results of the tests were the same as reported above. The observed values of TB mortality were lower than expected when isolation was practiced and when antibiotics were introduced.

4. Testing access to care: Medicare and life expectancy

The argument that improved access to medical care would reduce mortality in the US inevitably leads to the hypothesis that the advent of Medicare should have increased life expectancy for persons at age 65. Indeed, the justification for the program was to protect the health of a vulnerable population that had relatively little access to care. Medicare, created by a 1965 amendment of the Social Security Act, provides federally financed health insurance for persons 65 years of age or older. The program began on July 1, 1966, at which time 19.1 million persons were enrolled. Approximately 3.7 million persons received health services paid for by Medicare in the last 6 months of 1966 alone. The

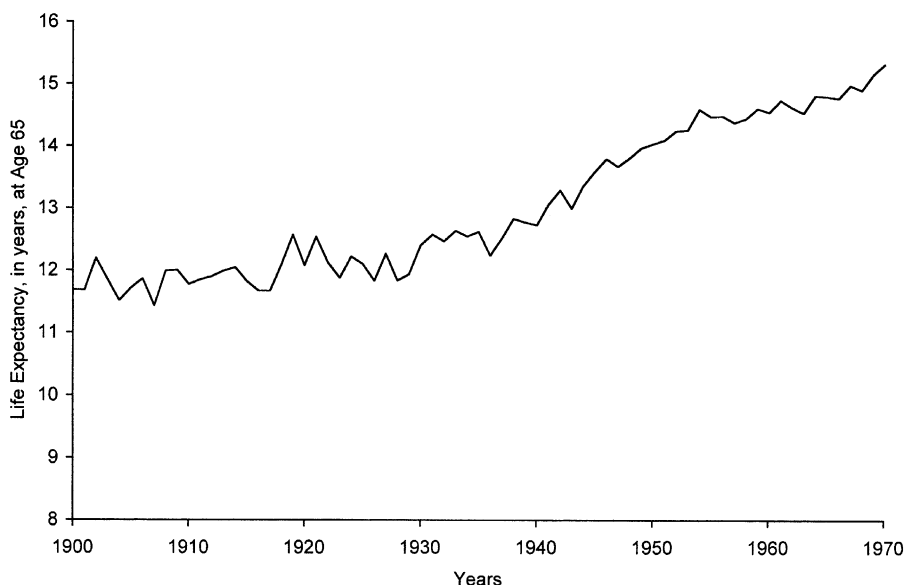


Fig. 2. Life expectancy for persons aged 65 for the years 1900 through 1970.

number of persons treated continued to climb in succeeding years. We, therefore, would expect to see effects of Medicare on life expectancy at age 65 sometime in the 5-year period 1966–1970, inclusive.

The Medicare hypothesis was tested for the years 1900 through 1970. Fig. 2 is a time plot of life expectancy, in years, at age 65 in the USA, from 1900 to 1970 [24].

4.1. Results of the medicare test

The best-fitting ARIMA model for the natural logarithm of life expectancy at age 65 during the years 1900 through 1965 was as follows.

$$\nabla Y_t^c = .0033 + \frac{1}{(1 + .4494B + .3768B^2)} a_t \quad (5)$$

The three parameters were at least twice their standard errors. The residuals exhibited no autocorrelation. The model suggests that the obvious downward trend included a less-apparent tendency for high or low values to decline over 2 years. The model accounted for 95.3% of the variation in the life expectancy series.

As shown in Table 2, none of the observed data was below the 95% confidence interval of values expected from history. We must accept the null hypothesis that observed life expectancy at age 65 from 1966 through 1970 was not different from the expected values. There is, in other words, no statistical association demonstrated between the implementation of Medicare and changes in life expectancy at age 65.

As with the TB test, it could be argued that using fewer pre-Medicare years would have produced different results. We, therefore, repeated the test using 30 pre-Medicare years. The results were the same as those reported above.

The null hypothesis that life expectancy was not different from that forecast from the 30 prior years was accepted.

5. Discussion

The time-series quasi-experiment can detect the hypothesized effect of two very different health care innovations on TB mortality in Massachusetts between 1850 and 1950. The same method, however, finds no change in life expectancy among persons aged 65 when Medicare was implemented. How can these results be reconciled?

The most parsimonious explanation is that Medicare truly had no effect on longevity. This interpretation would be consistent with the view that modern medicine, at least as practiced in the late 1960s, was largely palliative among the elderly and had little or no measurable effect on the upward trend in their life expectancy [6]. This says nothing, of course, about the contribution of the program to the quality

Table 2
Point estimate (i.e., if Medicare had no effect) of observed life expectancy for persons aged 65 (Confidence interval (95%) of point estimate is also shown)

Year	Lower bound of point estimate of life expectancy at age 65	Point estimate of life expectancy at age 65	Upper bound of point estimate of life expectancy at age 65	Observed life expectancy at age 65
1966	14.291	14.798	15.323	14.780
1967	14.317	14.898	15.503	14.990
1968	14.335	14.948	15.588	14.910
1969	14.288	14.979	15.703	15.170
1970	14.285	15.037	15.829	15.340

of life among the elderly or to the social cohesion of American society [1].

A second and, to us, more intriguing explanation is that the effect of medical care on the population studied was to sustain, rather than accelerate, the previously observed upward trajectory in life expectancy. The implication is that the trajectory could not have been sustained without Medicare. Quasi-experiment methods cannot detect such an effect because, without an unbiased control group, the expected value under the null hypothesis is a continuation of the trajectory observed before the intervention.

Medicare could have sustained the upward trend of life expectancy under at least one set of circumstances. These circumstances are the following.

1. Uninsured persons over age 65 in fact received considerable health care—especially that affecting mortality in the short run, such as care for acute serious illness and injuries—before the advent of Medicare. Much of the care could have been charitable or otherwise uncompensated. Other care could have been compensated out-of-pocket by the patient or his or her family.
2. Without Medicare, the fraction of ill, uninsured elderly persons receiving *charitable* care would have decreased in the late 1960s. This would have occurred because: (a) their number was increasing faster than before, and (b) the willingness or capacity of physicians to provide uncompensated care may have declined, as more expensive medical interventions (e.g., ICU/CCU care) for extending longevity in the elderly became commonplace.
3. The fraction of ill, uninsured elderly whose care had been previously paid for out-of-pocket, by themselves or friends/relatives, would have begun to decline simply because the cost of care began to rise.

Under the above circumstances, Medicare could have allowed the ill elderly, although increasing in number, to go on receiving health care at a level previously maintained by charity and out-of-pocket expenditure. In this way, Medicare would have maintained, but not accelerated, the trend in life expectancy already evident in Figure 2 before the 1960s.

This possibility raises at least two issues for further research and discussion. The first and most obvious is, were the circumstances described above actually present in the 1960s? This empirical question is perhaps best answered by demographers and historians, whom we welcome into this debate. Efforts should focus on dating the emergence and possible effects of phenomena such as the increasing numbers of elderly, the ethical mandate to provide the best possible care to all patients, and the increasing cost of such care.

The second and more general issue is, which, if any, *a priori* considerations would allow researchers to separate interventions under study into those that should accelerate

versus sustain pre-existing trends in life expectancy? This issue should be addressed early in any effort to measure the historical contribution of health care to declining mortality, in order to avoid misleading conclusions being drawn from quasi-experiments.

More work needs to be done to devise quasi-experimental tests for interventions believed to sustain rather than accelerate trends. Prototypes of such methods are available [15], but they require comparison groups and have not been applied to the effects of medical care on longevity.

The prospects of a troubling circumstance should motivate us to support the development of alternative methods. That circumstance is one in which our faith in the clinical trial leads us to accept its near cousin, the time-series quasi-experiment, as the methodological convention for assessing the efficacy of health insurance reforms. For the reasons described above, these assessments are likely to support the position that increased access to medical care has no measurable effect on longevity. This conclusion may be statistically accurate, but it says nothing about the contribution of medical care to maintaining the documented long-term increase in longevity. We are concerned that the uncritical use of these methods, without full appreciation of their strengths and weaknesses, may do as much harm as good.

References

- [1] Bunker J, Frazier H, Mosteller F. Improving health: measuring effects of medical care. *Milbank Q* 1994;72:225–54.
- [2] Fogel R. Nutrition and the decline in mortality since 1700: some additional preliminary findings. Cambridge, MA: National Bureau of Economic Research, 1986.
- [3] Fogel R. Economic growth, population theory, and physiology: the bearing of long-term processes on the making of economic policy. Cambridge, MA: National Bureau of Economic Research, 1994.
- [4] Mackenbach JP. The contribution of medical care to mortality decline: McKeown revisited. *J Clin Epidemiol* 1996;49:1207–13.
- [5] Szreter S. Fertility, class, and gender in Britain, 1860–1940. Cambridge: Cambridge University Press, 1996.
- [6] McKeown T, Record R, Turner R. An interpretation of the decline in mortality in England and Wales during the twentieth century. *Population Studies* 1975;29:391–422.
- [7] McKeown T. Determinants of health. *Human Nature* 1978;1:54–9.
- [8] Newhouse J. Medical care expenditure: a cross-national survey. *J Human Resources* 1977;12:115–25.
- [9] Gruchow W, Rimm A, Hoffman R. Alcohol consumption and ischemic heart disease mortality: are time-series correlations meaningful? *Am J Epidemiol* 1983;118:641–50.
- [10] Hertzman C, Frank J, Evans R. Heterogeneities in health status and the determinants of population health. In: Evans R, Barer M, Marmot M, editors. *Why are some people healthy and others not?* Hawthorne, NY: Aldine de Gruyter, 1994. ch. 3.
- [11] Catalano R. The effect of deviations from trends in national income on mortality: the Danish and USA data revisited. *Eur J Epidemiol* 1997;13:737–43.
- [12] Brooker A, Frank J, Tarasuk V. Back pain claim rates and the business cycle. *Soc Sci Med* 1997;45:429–39.
- [13] Phillips D, King E. Death takes a holiday: mortality surrounding major social occasions. *Lancet* 1988;2(8613):728–32.
- [14] Phillips D, Christenfeld N, Ryan N. An increase in the number of deaths in the United States in the first week of the month; an associa-

- tion with substance abuse and other causes of death. *N Engl J Med* 1999;341:93–8.
- [15] Catalano R, Serxner S. Time series designs of potential interest to epidemiologists. *Am J Epidemiol* 1987;126:724–31.
- [16] Patterson R. A century of tuberculosis control in Massachusetts. Boston, MA: Tuberculosis and Health League, 1951.
- [17] Fairchild AL, Oppenheimer GM. Public health nihilism vs pragmatism: history, politics, and the control of tuberculosis. *Am J Public Health* 1998;88:1105–17.
- [18] Wilson L. The historical decline of tuberculosis in Europe and North America: its causes and significance. *J Hist Med Allied Sci* 1990;45:366–96.
- [19] Bryder L. *Below the magic mountain: a social history of tuberculosis in twentieth century Britain*. Oxford: Clarendon Press, 1988.
- [20] Noymer A, Garenne M. The 1918 influenza epidemic's effects on sex differentials in mortality in the United States. *Pop Devel Review* 2000;26:565–81.
- [21] Dubos R, Dubos J. *The white plague*. Boston: Little, Brown and Company, 1954.
- [22] Lancaster H. *Expectations of life*. New York: Springer-Verlag, 1990.
- [23] Box G, Jenkins G, Reinsel G. *Time series analysis: forecasting and control*. London: Prentice, 1994.
- [24] United States Bureau of the Census. *The statistical history of the United States, from colonial times to the present*. New York: Basic Books, 1976.