

The health transition: the cultural inflation of morbidity during the decline of mortality*



S. Ryan Johansson

Ethics in Society, Stanford University, California 94303-2155, USA

Abstract

It has become commonplace to observe that as mortality falls, morbidity levels rise. The question is why? The explanation offered here stresses the multidimensional nature of morbidity, and the important role that diverse cultural forces have on the patterns of behaviour which underlie reporting behaviour during modernization. These forces involve rising health expectations on the part of ordinary people, including their ability to perceive illness and their willingness to seek professional help, and institutional pressures on medical professionals which reward them for discovering and treating an ever-growing set of non-fatal diseases. Since non-Western developing countries are training physicians to practice scientific medicine, are educating their citizens to think about disease along modern lines, and measure morbidity as developed countries do, there is every reason to suppose that as mortality falls in these countries, morbidity will rise, just as it has done in the developed world.

Introduction. The Health Transition

Most of the world's developing countries are somewhere in the middle of what has been termed the health transition (see Caldwell 1990). This transition is related to, but separate from the mortality and epidemiological transitions which are already familiar to many social scientists. According to Caldwell the health transition relates to the role that the cultural, social and behavioural determinants of health play in rising life expectancy at birth (the mortality transition) and the decreasing proportion of all deaths caused by infectious diseases (the epidemiological transition). The institutional changes associated with the cultural aspects of the health transition include the systematic promotion of the techniques of disease control and health care associated with modern medicine at the expense of folk medicine.

In the narrowly quantitative sense, the health transition relates to observed levels of morbidity during the decline of mortality. Although it would be logical to assume that morbidity levels will decline as the health transition proceeds on a worldwide basis, in this case logic has certain limits. In

* Many people have commented on versions of this article. I am particularly grateful for John Caldwell's interest and encouragement. The first version was given at a talk at the University of Michigan's Population Center in October 1987; Lois Verbrugge made many helpful comments. A second version stressing the multidimensional nature of morbidity was presented at the Social Science History Association in Chicago in November 1987; Paul David and Partha Dasgupta commented on the paper at that stage. The final version was written for a conference called 'A Research Agenda on the Compression of Morbidity', held at Asilomar Conference Center, Pacific Grove, California in March 1990, sponsored by the National Institute on Aging (Grant No. 1 R13 AGO8966) and the Kaiser Permanente Medical Care Program of Northern California; comments were made by Jay Olshansky, Mitchell LaPlante, Elizabeth Barrett-Conner, David Ballard, George Myers and Robert Wallace. None of the above readers is responsible for any remaining errors or the risks taken on behalf of morbidity history. My observations on Japanese morbidity are part of a project on the history of death and disease in modern Japan undertaken with Carl Mosk.

general, it is not high-mortality developing countries which have high levels of morbidity, and low-mortality (generally developed) countries which have low levels of morbidity, but quite the reverse. Currently, reported morbidity levels are highest in the wealthiest countries, as are the per capita amounts of money spent to secure health and control disease (Fuchs 1990). In the United States the health sector's share of the gross national product in the late 1980s was estimated to be about 11 per cent; in the 1940s it was under 5 per cent and in the early twentieth century it was under 2 per cent, indicating through indirect means a rise in the amount of money spent on health and disease. Since little or nothing in the modern health history of Western and developed countries suggests that net morbidity levels decline with mortality it is unlikely that the health transition in the developing countries will lead to better health, if better health means lower reported morbidity levels (see Riley 1990a,b, but also Schach 1985).

But that is the problem. 'Health' has more than one meaning, and so do the concepts 'morbidity' and 'disease'. All of these concepts are complex, and intrinsically difficult to define and measure, because they have a strongly cultural character which permits their meaning to change over time and space. The cultural nature of health, disease and morbidity must be confronted if we are to understand why morbidity tends to rise during the health transition.

At this point, depending on the disciplinary background of any particular reader, emphasizing the importance of cultural influences on morbidity levels could seem either an exercise in the obvious or too 'flaky' to deserve the consideration of serious social scientists.

The first response would be most likely to come from sociologists and medical anthropologists, who over the last fifteen years have been working to sensitize the medical establishment to the cultural dimensions of health. By now, the literature has reached a sufficiently advanced state to require surveys which summarize what has been discovered about cultural influences on health and illness (e.g. Helman 1990). But even the recent literature on culture and health remains relatively ahistorical and largely unquantitative. Very little cultural research has dealt with the role of cultural influences on long-run trends in morbidity.

The second response is most likely to come from economists, who often regard the descriptive nature of the cultural research as a sure sign that it lacks scientific utility and explanatory power (see Johansson 1990). Most health economists would probably identify wages, productivity, technology and insurance as the most likely culprits in the rise of morbidity, not only because they are economists, but because each of these factors is clearly defined and potentially measurable (Fuchs 1990). High wages make it possible for consumers to buy more health care; lower costs make it possible for doctors to sell more health care. In non-market situations, increasingly generous benefits for the sick and disabled can encourage people to try to claim such benefits (Rice & Laplante 1988). Thus narrowly economic approaches to the study of morbidity make it seem as if the increase in morbidity and disability is a form of mass hypochondria, mass malingering or medical salesmanship. But in any case, economists who study the impact of development on human welfare find it difficult to explain why rising life expectancy, which is so frequently used to index human welfare, is associated with rising morbidity, which is difficult to interpret as a sign of increasing welfare (Dasgupta n.d.¹).

As the set of specialists most influential in generating and interpreting modern morbidity data, epidemiologists fall somewhere between these two extremes. The basic premise of epidemiology is that disease occurs in patterns which reflect the operation of their underlying biological causes (Fox, Hall &

¹ I am grateful to Partha Dasgupta for allowing me to read the manuscript of his forthcoming book on welfare economics.

Elveback 1970). But at the level of any specific disease, a rise in incidence may result from better detection and reporting, not a real underlying biological change (see Wynn et al. 1989). Multiple sclerosis is only one of many diseases which have been tracked for decades, but now that it is detected with increasing accuracy, incidence rates are rising although the number of real cases may be declining. Although this awareness is compatible with the idea that there are diverse cultural influences on reporting behaviour, it does not necessarily lend itself to the belief that health and morbidity are intrinsically 'cultural' phenomena. To some epidemiologists the claim that morbidity trends have important cultural dimensions seems to threaten the scientific status of epidemiology itself. If epidemiologists study disease, which is defined as biological, they are scientists; if they study illness, which is defined as sociological, they are, at best, only social scientists.²

Neither health historians nor demographers have been able to settle these debates by bringing their own particular perspectives to bear on the relationship between morbidity and mortality over time. They too cannot agree among themselves to what extent the rise of morbidity during the mortality transition is biologically real.

Health historians relate the progress of civilization in its recent forms to the rise and triumph of public health, and the control of infectious disease which is its hallmark (Sigerist 1943). That this triumph involves cultural influences on health and medicine is recognized but not stressed (Dubos, 1968). To other health historians the triumph over infectious disease is ambivalent, since it involves the replacement of one set of diseases identified as leading causes of death with another equally numerous set of causes (Omran 1971). Thus the substitution of chronic causes of death for those which were infectious may simply leave net morbidity unchanged during the health transition (McKeown 1988). Still other historians have been convinced that the rise of morbidity levels during the health transition is a sign of genuinely deteriorating health among ordinary people. If those people most likely to die of infectious disease in traditional mortality regimes were among the sickest and most frail members of their population, then by dying at an early age the frail left behind a population of survivors which was temporarily at least, healthier than it would have been had they stayed alive (see Alter & Riley 1986; Riley 1987). Therefore when mortality increases and the frail die, morbidity decreases; when mortality falls and the frail are saved, morbidity increases. Some physicians adopt a similar line of argument, by pointing out that in the very recent past doctors have developed a much more efficient technology for saving and extending the lives of people who would have otherwise died. When modern medicine succeeds in saving marginal lives, it automatically increases the real (biological) burden of morbidity and disability in low-mortality populations (Gruenburg 1977).

But since rising morbidity and declining mortality, or rising life expectancy at birth, seem so fundamentally inconsistent with one another most demographers continue to believe that measures of mortality are and remain the best indicators of a population's 'real' health status, and thus the best measurable indicator of morbidity levels (see Ruzicka & Kane 1990). This solution is very unsatisfactory because it sabotages the systematic study of morbidity data. In policy-related terms, it leaves open the possibility that the rise of morbidity is indeed illusory, and it lends support to the idea that, in order to prevent mass hypochondria and induced malingering, generous health benefits and disability programs should be cut back in the developed countries, and not extended to people who live in developing ones.

² Generally epidemiologists believe that by use of data which come from blood tests, or other laboratory data, the study of disease can be separated from the social study of illness (Helman 1990:271). However, most of the information used by epidemiologists to estimate long-run trends in morbidity does not come from blood tests, or laboratory analysis; it comes from individuals and doctors who report data on perceived illness.

The persistent belief that biological morbidity must decline as life expectancy rises has led a small set of health optimists in the developed countries to predict an impending 'compression of morbidity', a future state in which life expectancy will finally reach its 'natural' upper limit, and most reportable disease will be confined to the very last year or years of life. In this scenario, after a long period of rising morbidity, which remains inadequately explained, individuals in the developed countries are finally liberated by science to live longer *and* healthier lives at costs which are no greater than those already borne, and possibly less.³ Thus the nightmare of rising morbidity ends, and the health transition ends happily (but see Kaplan 1990).

The cultural influences on such optimistic scenarios are obvious. Westerners prefer stories with happy endings. But it is unlikely that the anticipated compression of morbidity will occur, since nothing is being done to modify or change the cultural forces which drive morbidity upward, and which are intrinsic to the health transition as an historical process involving simultaneous changes in the biology of health and disease, and how disease is perceived, named and reported.

Thus the general study of cultural influences on morbidity history is not an attempt to deny the relevance of biology, medical science or economics to the study of health. Its aim is to integrate all these diverse influences into one coherent story. This is a very different undertaking from the specialist study of specific cultural practices in various local contexts. Although in the past, culturally diverse forms of folk medicine may have made valuable contributions to the demographic viability of pre-modern societies (Micozzi 1983), none were able to reduce the high parasite loads common among past and present rural populations in developing countries (*Science*, March 23, 1990:1415) or to efficiently eradicate the lethal communicable diseases which once caused extensive premature death (Preston 1976). Scientific medicine had and has the potential to control a range of lethal diseases that were beyond the reach of folk medicine. The potential of modern scientific disease control can only be realized once certain cultural and institutional 'reforms' become universalized; but these very reforms tend to set in motion the inflation of morbidity.

Thus, the purpose of this article is to demonstrate how the study of cultural influences on health and sickness interacts with the changing biology of disease, the institutionalized conventions which govern reporting, and the economics of being sick during the health transition in such a way as to produce a long-run, but not steady, inflation in morbidity. The article discusses the policy implications which flow from taking an historically and culturally integrated approach to the study of morbidity history during the health transition, as opposed to approaches with a narrowly biological or economic perspective on health and sickness in the relatively recent past.

The cultural foundations of health and disease

Until 1980 or so, when modern medical technology finally made death more biologically complex than it used to be, not much cultural training was required to tell the difference between the living and the dead (Veatch 1989). But the difference between sickness and health has never been completely clear; this intrinsic fuzziness has always generated the need for interpretation, which can either be left to the imagination of each individual or can be culturally standardized. When the individual's interpretation of his or her own health status has social and economic implications for the other members of a household, family, lineage or community, society is unlikely to leave the individual decision makers to their own imaginative devices. Instead, cultural pressures designed to standardize perception and behaviour will

³ The 'compression of morbidity' concept was first developed by Fries (1980). The core of his argument is that there is a fixed upper limit to the natural human lifespan of 85 years. As individuals and their societies learn how to minimize preventable disease, that which is not preventable will take place only in the last few years of life.

be brought to bear on 'private' decision making with respect to health related preferences. These social pressures will focus on the identification of a breakpoint in the health-sickness continuum.

What is the health-sickness continuum? Between feeling wonderful and being so sick that it is difficult to remain conscious, there are a whole range of health gradations recognized in ordinary language, and which the ordinary person must be taught to distinguish between according to some criteria. Thus, in modern American English a person can feel very healthy, healthy, so-so, not so good, not so bad, bad, awful, near death.

In this health-morbidity continuum the precise breakpoint where 'healthy' ends and 'sick' begins is not obvious; but identifying it is not a trivial matter, since under most circumstances a judgement of 'sick' is usually a signal that the routine of the newly sick person (going to school, going to work or even the ordinary activities of daily living involved in dressing, eating and personal hygiene) has been, or will be disrupted, and that other people in the household, neighbourhood or the community will be tapped for assistance. Insofar as the disruption of an individual's routine has social and economic consequences for others it cannot be a matter of personal tastes or preferences. Rules must be established which link biological states to behavioural outcomes. These rules are contained in the culturally constructed values and norms associated with health and disease, the purpose of which is to induce a limited degree of self-organization with respect to the disruption of normal routines and the systematic delivery of assistance and material resources to the sick and disabled. The standardized breakpoint on the health continuum is socially constructed through various forms of institutionalized negotiation designed to culturally integrate the biology of disease with the resources available to cope with sickness in various places and times.

The subset of people who decide that they are sick must also decide whether or not they are sick enough to seek professional assistance from a folk practitioner or doctor, instead of home treatment. The decision to seek professional assistance also involves a culturally standardized breakpoint in the health continuum, because it also involves a covert claim on household or social resources. Those individuals who seek professional assistance for a health problem, either their own or one of their dependants', will find themselves dealing with culturally trained health professionals who have also been trained to interpret symptoms along culturally standardized lines. That is, a doctor's interpretation of a patient's health status is not a matter of the doctor's personal medical tastes and preferences, because there are cultural rules designed to constrain medical judgement. These rules govern how people with certain symptoms should be classified into malingerers, hypochondriacs or patients genuinely suffering from specific diseases which require certain specific types of treatment. A range of institutions from medical schools to law courts have been set up to monitor the cultural orthodoxy of medical professionals, and thus to enforce a reasonable degree of standardization on medical interpretation and decision making.

In modern societies various institutions have also been socially constructed to track the decisions made by individuals about the state of their health, and the decisions made by doctors about the extent and nature of disease. In the developed countries the institutionalized reporting of morbidity involves annual or more frequent health surveys designed to gauge the extent to which people say they are sick or well in one limited time period. The reporting of morbidity also encompasses doctors who count their patients and report on the number who have specific diseases. Hospitals are ordered to report on the number of people they admit, and what procedures are used to deal with diverse physical and mental health problems. In addition private agencies which insure health and pay for home and institutional care also count the number and type of cases they process, and the expenditure associated with each case.

This list does not exhaust the kinds of morbidity data available in modern societies but it reminds us that the diversity of morbidity data makes them much more complex to analyse than mortality data. In general social scientists cannot or should not attempt to relate morbidity and mortality during the health transition because morbidity is made up of phenomena of several different kinds, each of which relates differently to sickness and death.

The multidimensional nature of morbidity

We can identify and measure four quasi-independent dimensions of morbidity which make it more complex to report and measure than mortality. The first dimension of morbidity tracks its proportional extent in the total population in some specified time period. The extensiveness of morbidity has two subdimensions: incidence and prevalence (Peterson & Thomas 1978). The incidence of morbidity measures the proportion of the total population (usually the total number of people in a specific age group) who regard themselves as having become sick, or who were diagnosed as sick, in general or with a specific disease, by a health professional during some limited time. The more diseases there are, that is, the more diseases for which patients and professionals have names, the more likely are individuals to think of themselves as sick or to be diagnosed by a professional as sick. In this way the incidence rate is a function of the culturally recognized stock of diseases, along with the propensity of ordinary people to classify biologically suboptimal states as sickness according to the culturally standardized breakpoint on the health continuum.

But the incidence of morbidity does not measure the proportion of people who were already sick at the beginning of the time period under observation. Whatever the period of observation, some people were already sick when it started. Since new and old cases of sickness exist side by side, the prevalence of morbidity is measured as the proportion of people who were already sick at the beginning of the time period selected for measurement. Obviously the prevalence rate must be influenced by the length of time sick people are 'normally' sick with the 'average' disease. But this time period is sensitive to the types of disease which are reported or treated most frequently, and subject to the amount of support society is willing to provide for sick children, adults who must take time off work, or older people who need help with the ordinary activities of daily living.

The second dimension of morbidity is duration. All episodes of morbidity end in recovery or death. Obviously, the longer the 'average' reported disease lasts before it leads to recovery or death, the higher the probability that the average sick person will be reported as sick or feel motivated to seek treatment in some time period. But duration is also influenced by the stage of development in which a disease can be medically detected, and by the extent to which medical technology can keep a seriously ill person alive; and keeping sick people alive and in need of treatment is costly: therefore the average duration of reported diseases will be influenced by how much money or time is available to support the sick and disabled. Obviously the farther income levels are from mere subsistence, the more resources can be devoted to the care of the sick and disabled, if that is made a social priority. But income in itself does not set a fixed level of investment in health. The role of health care in an overarching value system matters just as much, since it determines how health related investments will fare, vis-à-vis other alternatives, in the competition for resources at the household and community level.

How much and what kind of support the sick and disabled require is related to the severity of the conditions which afflict them (dimension 3). Some forms of morbidity require hospitalization, but most require nothing more than staying home and taking a few inexpensive medicines. Some forms of disability require total community support, others require partial support. The relative severity of various forms of disease and disability can be measured in several ways, but generally the more severe a disease, the more likely it is to end in death rather than recovery in a limited time period and therefore the more likely it is to require some form of professional intervention. In epidemiology, case-fatality

rates are generally designed to distinguish between diseases which greatly increase the probability of dying in a specified (and fairly short) period of time, and those which do not. Diseases which are classified as non-fatal, for example, have no clear relationship to the probability of dying in some fairly limited period of time.

Disease-specific case fatality rates are not constant over time, however, since the severity of any particular disease, which relates its onset to the probability of dying in a set time period, can be altered by the evolution of medical technology which extends the time between onset and death. This development would tend to increase its prevalence, if incidence rates remained constant or even declined to some extent. Similarly the extent to which disabling conditions require full or partial support can be influenced by the progress of medical technology.

Lastly, individuals in the morbid state can be sick for more than one reason. This gives morbidity a fourth dimension which can be called its depth, but is conventionally called co-morbidity. Since there is no clear limit to the number of physiological conditions that can be identified as diseases, individuals can have and report (or have reported for them) several diseases at once. At the individual level co-morbidity greatly complicates the decisions made by clinicians with respect to various treatment strategies. But at the aggregate level, where institutions operate, the depth of morbidity creates what epidemiologists call the 'iceberg' of disease (Peterson & Thomas 1978:21).

This metaphorical iceberg is a reminder that much more disease always exists than is normally perceived and reported by health professionals. Because so many suboptimal forms of health exist, the medical establishment must assign priority to detecting the existence of those diseases which do most harm. But as more and more scientific research is done, more and more new diseases are 'discovered', and as scientific knowledge advances, the meaning of 'harm' can change. The more new diseases are admitted to an official list of diseases, with which doctors are expected to be familiar, and about which ordinary people are expected to learn, the greater the extent of perceived co-morbidity. This tendency is further increased by the professional drive to detect named diseases or forms of disability at increasingly earlier subclinical phases. This in turn leads to the conversion of vague symptoms, for which there is no treatment, into specific diseases, for which some form of treatment is available, or for which new forms of treatment ought to be invented. All of these trends increase the perceived prevalence of disease, the perceived extent of co-morbidity, and the fiscal costs of coping with a higher volume of sickness (Fried 1990).

Mortality lacks the multidimensional complexity of morbidity. Even now comparatively few individuals exist in a biological state which requires doctors, relatives and institutions to decide whether or not a person is alive or dead, by interpreting ambivalent biological signs; although this is becoming more frequent. Similarly, once individual bodies are declared dead very few come to life again. This biological crispness makes it comparatively easy to measure the incidence of mortality in a given time period, and deprives mortality of a durational dimension which can usefully be tracked, other than for personal reasons such as the anniversary of a relative's death. Neither does mortality have relative severity, since all the dead are equally dead. Finally, although a person can die from more than one cause, once dead, multiple causes of death do not make it necessary to treat dead bodies differently, or lead us to expect that we will discover (through more research) previously unsuspected cases of dead people among the living.⁴

As long as the biologically definite character of death made it relatively easy for anyone to 'diagnose', culturally constructing a shared interpretation of which bodies were alive or dead was

⁴ Manton and Stallard (1984) discussed the history and evolution of reporting multiple instead of single causes of death. There is still disagreement about the biological and demographic utility of multiple cause-of-death data.

unnecessary. But the recognition of sickness and the legitimate disruption of ordinary routines on its account, has always been subject to cultural negotiation, except perhaps when sickness takes its most extreme forms, and the afflicted person can barely remain conscious. With the exception of these most extreme forms of morbidity, the negotiation over sickness extends to defining and measuring morbidity, including which of its several dimensions should be reported, and how reporting will be institutionalized.

Failure to respect the intrinsically complex and multidimensional nature of morbidity makes it possible to assume or prove contradictory things about the relationship between sickness and death during the health transition, including its most recent decades (Palmore 1987). Because each of the dimensions of morbidity has a different and quasi-independent relationship to mortality during the health transition, it is impossible to find some timeless and natural law of morbidity and mortality. Since morbidity has multiple, context-influenced meanings, and the way that information about it is gathered differs from one institution to another, it is not easy to discern the relationship between any particular set of morbidity data and mortality trends, particularly in the short run, outside fairly culturally specific contexts (Hansluwka 1985:5).

We can systematically consider the general history of reported net morbidity, in some general, or age-specific form, and the reported incidence of mortality, as the death rate in its summary or age-specific forms. This history of incidence rates will be cultural to the extent that the reported incidence of disease is influenced by how people and health professionals are trained to report sickness, as well as how many legitimate diseases there are to report. There is also the general history of mortality and the prevalence of disease, which is related to the type of disease most frequently reported, and the medical management of disease. Next, there is the history of mortality and the duration of reported diseases, including the material resources available to support extended periods of sickness and disability. Subsequently, there is the history of mortality and the severity of most reported diseases, which includes the history of medical technology and its ability to eradicate or manage various types of diseases; but it also includes how ordinary adults are taught to manage health and sickness at the household level. This management includes knowledge about nutrition and sanitation, and a limited amount of training about how to treat various forms of illness without professional assistance, so as to amplify the body's natural ability to recover from sickness. Lastly, there is also the history of mortality and co-morbidity which relates to science, medicine, health education, economics and reporting conventions.

Each of these histories deserves separate treatment, but the more specialization is encouraged, the less easy it is to get a coherent picture of the overall evolution of the health transition, and the more likely is the belief that the history of morbidity during its course is beyond the reach of systematic generalizations. By a culturally integrated approach to the history of the health transition, and what its study requires in terms of quantitative data at both the micro and macro level, the set of culturally influenced relationships between multidimensional morbidity, and unidimensional mortality, can be considerably clarified.

By considering the cultural aspects of morbidity, for example, we can see how things cultural conspire to inflate the amount of reported morbidity during the health transition, even though the severity of the most frequently reported diseases is decreasing and in a fundamental biological sense there is less life-threatening exposure to particularly dangerous pathogens than formerly (see Johansson & Mosk 1987). The inflation of morbidity is thus biologically real, but it involves a shift from the reporting of diseases which are fatal or relatively severe (that is, diseases which materially increase the probability of dying in a specified and fairly short period of time) to those that are less fatal and even non-fatal. At the same time this very shift involves an increase in the duration of reported disease,

which begins with the substitution of chronic for infectious disease during the cause of death transition, but also includes the advance of life-extending medical technology, and the increased amount of income and resources created by economic development, and made available (or not) through policy for the community support of disease and disability. As the management of morbidity is effectively severed from the management of fatal diseases, levels of morbidity and disability, including co-morbidity and multiple disabilities, can keep rising, and the costs associated with them can keep rising as well, without having an apparent relationship with life expectancy at birth, currently the most popular summary measure of mortality. In fact the amount of money spent on health can rise indefinitely until the cultural willingness or perceived fiscal capacity to pay for culturally recognized manifestations of sickness and disability has been exhausted.

The health transition can be modelled as a culturally managed process. Economic and biological realities are interpreted by both lay people and professionals according to their cultural theories of health. Standardized perception leads to standardized reporting behaviour which ends with quantitative morbidity data. No part of this process is constant over time, making the health transition intrinsically historical and cultural rather than timeless and natural. The absence of true constants invalidates methods which model the health transition as if one of its components could be allowed to vary while the others remained constant (Isaac & Griffin 1989). Since all of the components of this process can change, and do in the short run at least, it is better to track the evolution of this complex system with the descriptive logic imposed by narrative forms, than with the analytic logic characteristic of formal models required by regression techniques. Most regression techniques assume, in one way or another, that time can be tamed, or removed from a multifaceted evolutionary processes, and they deny the reality of multiple structural shifts, which exist because each facet of a complex process can change over time at a different rate from the other facets to which it is loosely linked.

The rest of this article attempts to summarize and simplify the long and complex history of the health transition. Making the health transition simple involves telling it in before-after terms, as if it had a clear beginning and a definite end, although as any historian would suspect, this requires oversimplification. The simple contrasts sketched below necessarily sacrifice detail for generality, in the hope that a big picture (if not a gross violation of that reality) will be more useful for policy purposes than a detailed study of any one of its aspects (see McNicoll 1990:911).

The multidimensional history of the health transition

The cultural story of the health transition begins at the micro level with how individuals are trained to make culturally standardized decisions about their health or the health of their dependants. It proceeds at a higher more institutionalized level where health professionals are trained to make decisions about the health status of their patients, including the detection of specific diseases and the differential assignment of sick individuals to other health-related institutions. Above this level there is the bureaucratically visible layer of macro-level institutions, explicitly constructed to track and measure the decisions made by individuals and professionals at lower levels, and convert those decisions to data on trends in morbidity, whether in the form of incidence rates, prevalence rates, case-fatality rates, or measures of co-morbidity. Both micro-level and macro-level behaviour must be integrated in order to understand the health transition as an observable and measurable phenomenon with pervasive cultural foundations.

At the micro level the cultural history of morbidity begins by noting that children are necessarily taught to understand health and sickness in culturally specific terms which can be translated into rules about continuing or disrupting their own normal routines (Lau & Klepper 1988). Children are generally not in charge of their own health, but the ideas they absorb and the behaviour patterns they learn become the basis of how they will probably manage illness in later life, barring some sort of

culturally mandated continuing education or a disruptive and difficult process of cultural reprogramming in adulthood which requires that, to a certain extent, they develop a totally different orientation to processing biological and economic data (Johansson 1990). In any case decision-making strategies with respect to health and sickness are not abstractions separate from other aspects of human life; they quickly become part of the cultural construction of a social identity in which the perception of the body becomes an object of the mind (Scheper-Hughes & Lock 1987) and pain is given a socially sanctioned meaning beyond the individual recognition that it hurts (Das 1990).

The culturally influenced perception of health and sickness at the micro level and the willingness to decide that some suboptimal state of health is a sign of sickness, is the foundation for the downstream institutionalized tracking of the extent of disease. At the micro level what will eventually become some kind of incidence rate, for example, starts with the two breakpoints in the health continuum which were discussed earlier. The first breakpoint sets the difference between being healthy and being sick, and the second identifies when being sick requires professional assistance.

Both before and after the health transition being alive and being dead are clearly distinct biological states; whereas being well or sick involves a transition along an unbroken health continuum which has no natural biological markers as obvious as those which distinguish the living from the dead. Thus, short of virtual collapse and total immobility, being ill or disabled is subjective to some extent, and different cultures can vary the point at which biologically suboptimal states become sickness. But under normal circumstances the breakpoint 'sick' must be one which will leave sick people in a minority most of the time. Otherwise the resources allocated to their care will have to flow from a minority of the population to the majority, which would be virtually impossible for any length of time, particularly in low-productivity societies.

Biologically, therefore, sick people must have some symptoms which healthy people do not normally have. Thus cultural influences must constrain the subjectively variable perception of pain and discomfort, so that the hundreds of diverse sensations which can cause a departure from an extreme sense of well being and vitality do not chaotically disrupt productive activities any more than absolutely necessary, or overwhelm the capacity of a community to aid the extremely sick and disabled.

But in any case the probability of being sick or seeking help depends on where the breakpoints on the health continuum are set. The closer the breakpoints are to death the lower the reported incidence of disease will be, and the less frequently people will seek the services of doctors or healers for conditions which are merely troublesome or associated with continuing pain and discomfort. As long as the biological conditions which justify a judgement of 'sick' are relatively severe, – that is, they are associated with a high probability of soon dying – the average sick person will not be sick for long, before either dying or recovering. In any case, recovery will not necessarily mean the restoration of perfect health, but the abatement of certain acute symptoms. Thus a person who is no longer 'sick', in a functional sense, may continue to have one or more diseases which simply remain undetected. By extension, the real extent of co-morbidity will go unrecognized, as will the true prevalence rate; with 'real' and 'true' defined according to some set of scientific, biological criteria which are as culture-free as possible. At the beginning of the health transition, as long as productivity is low and poverty extensive, there is no point in designing incentives to encourage expert healers or bureaucrats to establish the real extent of co-morbidity, or to measure the true prevalence of suboptimal biological states.

Thus, in pursuit of simplified before-after contrasts it would have to be said that before the modern health transition the economics of being sick interacts with the biology of disease in such a way that, under cultural management, the incidence rate is kept low. Ordinary people are trained to disrupt their routines only in response to relatively severe diseases of comparatively short duration. No incentive

system exists to support the scientific (acultural and biological) study of disease. Nor would it be logical to construct such a system, given the fact that not much could be done to address so much disease. To the extent that the social and institutional monitoring of sickness exists, it is more or less organized to keep people working in suboptimal states of health for as long as possible, until the point of total collapse has been reached, and the probability of death has markedly increased. Towards this end various philosophies of pain are developed which attribute a high value to stoic endurance, and which teach people to think of pain in a positive light, such as that suffering pain in silence will be rewarded in heaven. People are also encouraged to think about disease in fatalistic terms which makes public intervention seem foolish or impious.

Obviously the evidence for the truth of a 'before' story of this nature cannot be found in the form of high-quality data on incidence and prevalence rates, or case-fatality or co-morbidity rates, since in the early phases of the health transition no institutions existed to gather data of this nature, except in a few limited and atypical cases. In fact, gathering morbidity data is part of the health transition, but like the transition itself, data-gathering evolves differently in various cultural contexts. This differential evolution makes it difficult to compare morbidity levels in time, between various cultures, or over time in one culture. It is the range and diversity of cultural influences on the bureaucratic gathering of morbidity data which makes some historians and demographers throw up their hands and look for other types of 'hard' data which will proxy morbidity during the mortality transition in a less culturally influenced way (Hanslowska 1988:5).

But as a substitute for non-existent or very limited types of quantitative data on morbidity it is possible to use historical descriptive evidence and contemporary descriptive evidence from developing countries. Anthropologists who have studied the culture of health in very poor countries write descriptions like the following. In parts of Africa where bilharziasis, a parasitic disease, affects the majority of a local population, most people pass blood in their urine. This condition, therefore, cannot be accepted as a legitimate reason for believing that the transition between well and sick has occurred, and that the daily work routine should be interrupted, or assistance sought (Barry et al. 1963:21). Similarly where most children have runny noses, or parasitic infections, or frequent diarrhoea, these conditions are not thought of as symptoms of illness which signal that a child is not basically healthy, and therefore needs time-intensive or special care. Mull and Mull (1988) report of rural Pakistan that mothers find it hard to take childhood diarrhoea as seriously as health professionals because it is so common among children and most episodes do not lead to death, even without ORT treatment. Thus, even when the government distributes the ORT packets free, mothers will not necessarily use them because they think it unnecessary.

Some low-productivity traditional societies once imposed something like formal limits on the duration of a 'natural' illness. Several decades ago anthropologists working in various parts of Africa observed that adults who felt sick for longer than three days were trained to consider their condition serious enough to suspect sorcery, and to consult a local healer. If the traditional healer could not effect a sudden cure, the next step was a fatalistic resignation to death. Since neither the household nor the community was organized to offer extended assistance to very sick people, death followed with surprising swiftness (Barry et al. 1963:21). Since the average community was not organized to support the household in managing diseases of long duration, elaborate forms of learned behaviour designed to delay death had not evolved as an ordinary part of the support offered to the stricken. But low levels of support for the sick were likely to make any particular disease more severe than it needed to be, that is, more likely to cause death, because home care was not designed to enhance the recuperative powers of the human body and its multifaceted ability to resist and recover from disease.

Obviously, the poorer a community and its households, the more likely it is that breakpoints on the health continuum will involve the onset of acute symptoms, and the more likely that it will take a close-to-death state to justify seeking professional help, either from a folk healer or a scientifically trained health professional. The literature summarizing recent research by health anthropologists in a wide range of developing countries stresses the reluctance of working adults in relatively poor households to take time off for sickness particularly if the household itself must entirely pay the cost of sickness (Wirsing 1985; Dennett & Connell 1988). Generally, when an adult stops working, the family's already marginal income declines still further, and healthy people must work still harder (Goodman et al. 1988:189–192).

Similarly, it is well known that minor forms of disability cannot count for much in impoverished areas where many people have poor vision, hearing, or other sensory impairments, along with stiffness in their joints, and other skeletal and muscular disorders which limit their mobility. In the past, people who were severely disabled had to find a way of turning their disability into an occupation, rather than a reason to stop working. Wherever begging was culturally sanctioned, being blind or being crippled could become a source of regular income. Disabilities had to be actively turned into an occupation before they could attract community support in the form of voluntary contributions. In this cultural context families of professional beggars have been known to mutilate their own children in order to put them to work soliciting alms.

Modern research suggest that high levels of underlying but unrecognized co-morbidity were intrinsic to the conditions which existed before the health transition began. When modern doctors, trained in the conventions of scientific medicine, do biologically oriented research on co-morbidity in traditional high-mortality societies today they almost always find surprising amounts of moderate to relatively severe unperceived disease. A recent study of the health of rural women living in two typical low-income villages in India unexpectedly revealed, after careful medical examination and some laboratory tests, that 92 per cent of those women examined had one or more gynaecological and sexual diseases, with an average of 3.6 per woman. This was not the sum total of their disease load; 98 per cent were anaemic, 58 per cent suffered from vitamin A deficiency, 12 per cent had filariasis, and 10 per cent had leprosy (Bang et al. 1989).

In other words, almost all the women had more than one serious disease, even though none were hospitalized at the time, and none had disrupted their daily routines to any great extent. In a cultural sense the women were healthy, although many had symptoms which they regarded as painful or unpleasant. As long as they shared those symptoms with many or most women, they considered them a natural part of a woman's lot in life, rather than the manifestation of a disease which required professional assistance. As long as none of the multiple diseases from which the village women suffered were severe enough to disable or kill quickly, the villagers saw no good reason to interrupt their daily routines or seek costly treatments.

Extensive co-morbidity can actually accelerate the biological progress of any particular disease towards death. The epidemic of AIDS, a disease which assaults the immune system itself and thus gives rise to a multiplicity of clinically expressed diseases, has also fostered more consciousness of co-morbidity and its role in the relative severity of other diseases. For example it is now recognized that people simultaneously infected with the human immunodeficiency virus and mycobacterium tuberculosis experience an accelerated progression to overt tuberculosis in both primary and reactivated cases (Harries 1990). The reported rise in tuberculosis which has taken place recently in several African countries, and to a lesser extent in the developed countries, is not necessarily related to the failure of public health, or increasingly inadequate nutrition, but a side effect of AIDS on the incidence and prevalence of another more routinely detected disease. Very probably in the past the declining

apparent severity of any specific disease might well have reflected declining but unobserved levels of co-morbidity, which for many individuals had previously shortened the period between onset and death, in addition to reducing the overall probability of a natural recovery among the afflicted.

After the health transition is completed (insofar as it is completed) conditions are dramatically different. People who live in modernized developed societies divide the health continuum in a very different way from that of their ancestors. They have moved the breakpoint 'sick' from near-death states to suboptimal biological states which merely indicate the lack of perfect health. This shifting breakpoint is clearly exemplified in the official definition of health adopted by the World Health Organization. Thus the United Nations has committed itself to the position that health is a state of complete physical, mental and social well-being, not merely the absence of disease or infirmity. Several decades ago it was recognized that this new definition was utopian, and that if taken literally it would mean (on a worldwide basis) that most people were sick most of the time (WHO 1983; see also Dubos 1968:87-88). The definition was not constructed by ordinary people living in ordinary environments, but by exceptional people living in environments which had been radically changed by the application of science to minimizing the harm done by the most severe diseases which caused the most premature deaths. This transformation in the disease environment was funded by the increasing productivity of developing countries. In such countries at present, if modern adults or their dependants are unwell, they need medical care or social assistance, even though the sick individuals are in no apparent danger of dying. Quite obviously a shift of this nature would automatically raise the reported incidence of morbidity (in the form of net morbidity, as opposed to the incidence rate for any one particular disease) as detected through health surveys, or any measure of morbidity designed to track the frequency with which individuals think of themselves as sick, disrupt their routines, use professional services, or buy medicines.

Even before the application of science to disease control, people who lived in privileged subpopulations were economically free to gravitate towards more generous definitions of what it meant to be sick. As early as the eighteenth century wealthy people in England had an almost modern sensitivity to pain, and an unambiguous desire to minimize it, even if that meant disrupting their normal routines until recovery occurred (see Porter & Porter 1988). Although the rich were already culturally positioned to afford much more illness than the poor, their high per capita incomes did not produce continually rising morbidity levels of the type we observe in the modern world. For one thing there were no institutions designed to measure such a rise. More fundamentally, the medical professionals with whom the pre-modern rich dealt had a very limited set of diseases which they were prepared to diagnose, and a very limited repertoire of techniques for curing or managing disease. Hospitalization was unfashionable and surgical techniques were comparatively primitive. In that context it was difficult for the high income levels of the already rich to fuel a continually rising level of morbidity, even if such a rise could have been funded without additional community assistance.

Thus, although the modern inflation of morbidity requires increasing income levels, various cultural changes must become institutionalized within the medical community and the bureaucracy before rising incomes can be 'successfully' converted into rising morbidity levels. The most fundamental cultural forces which drive morbidity upwards are new scientific theories of health and disease, which permit the continual discovery of new diseases; new techniques for diagnosing and treating diseases, each of which has real costs; new or revised institutional solutions for dealing with relatively severe forms of illness and disability; and the cultural legitimization of setting a lower breakpoint on the health continuum. Among other things this lower breakpoint converts formerly perceived hypochondriacs into perceived patients, and it reduces the pervasive suspicion that many sick people who are seeking community assistance are malingering.

Once again the most potentially subjective changes, those which relate to the amateur and professional perception of disease, are the most fundamental, because they alone legitimate increasing the amount of perceived morbidity. From the perspective of scientifically trained health professionals the before-after nature of the health transition can be simplified by considering what happens to the iceberg of disease.

Before the health transition the disease iceberg was very large. Only a small tip of the iceberg was medically visible in the form of named diseases which were diagnosed and treated. Most of these named diseases were of the acute infectious type with a comparatively short natural duration and high severity. The epidemiological transition began as diseases in this class were brought under control. If the list of recognized diseases had been held constant, the net incidence rate, assuming it was reported, would have declined. But historically all that happened is that medical attention moved to the chronic diseases, which by their very nature are less severe and of longer natural duration (Omran 1971). Moreover, as modern medical research explores an ever wider range of disease progresses, more and more diseases are discovered. Thus the iceberg of disease can diminish from the point of view of biology, while more and more of its volume becomes medically visible during the health transition.

Even if the iceberg of disease shrinks in some real biological sense the new, smaller disease iceberg is associated with a lower clinical threshold for detecting and reporting a wider spectrum of diseases, more and more of which are chronic and therefore by definition last longer. This type of change could continue into the future, so that if both heart disease and cancer were eliminated as leading chronic causes of death, an even broader spectrum of less severe diseases could replace them. Twenty-first-century people could be healthier than ever in some deep biological sense, while continually reporting higher and higher levels of less and less severe forms of morbidity, and thus more and more morbidity and co-morbidity. As long as each reported disease required some form of privately or publicly subsidized expenditure, the costs of health care would continue to rise until they reached a culturally perceived and negotiated limit. This observable and measurable rise of morbidity would be biologically real, but it would reflect an evolving biological reality perceived in a changing cultural context, including an increasingly sophisticated medical establishment routinely dealing with patients who have a much higher standard of living, and who therefore can afford to buy their way out of even minor pain. But this is quite different from seeing the rise of morbidity as a sign of mass hypochondria or large-scale malingering, both of which can be controlled by reducing benefits or raising costs.

The people who are harmed by harsh strategies are not just opportunists or the victims of imagined diseases. They almost all have real and painful diseases, but not severe and not clearly linked to the probability of dying. They will suffer as limits are imposed on the inflation of morbidity. Nevertheless, if most people, in particular the most powerful people, believe that limits must be placed on the level of community funding for health care, they must face the fact that they have a cultural problem. Should limits be placed on the medical establishment with respect to how many old diseases will be subsidized, or how many new diseases can be discovered? Or should limits be placed on the perceptions of ordinary health consumers, so that they will have to learn to accept a certain level of pain as a normal part of life and work, not only in old age, but in youth and mid-life as well? The increasing reluctance to accept pain of any kind (mental or physical) in any stage of life is illustrated by recent research on elderly patients in a sample of nursing homes. Not surprisingly the researchers found that as many as 80 percent were depressed, and that those who were depressed had a 59 per cent greater chance of dying in the course of a year than those who were not. Instead of accepting depression as a 'natural' response to aging, sickness and the impending end of life, the researchers concluded that elderly patients should be treated for depression, and their lives thereby extended (see Goleman 1991:4). Whatever is decided,

setting limits always involves cultural negotiation, based on various ethical considerations about the nature and value of human life and the implications of that value for the formulation of health policies (Callahan 1987).

It is probable that even a very rich country like the United States, which can afford to support extensive illness and disability, must face the fact that it does not have the resources to keep everyone in perfect health all the time. Nor can it afford to extend the lives of all those who could benefit from full use of the sophisticated forms of death-postponing technology already available. Nevertheless, because modern medical technology can do so much to keep people alive for long periods in a marginal condition, it also becomes increasingly necessary to pay explicit cultural attention to whether afflicted individuals *should be* kept alive, if the life they are leading is a 'low-quality' life, which is nevertheless very socially costly. Thus medical ethicists in America and elsewhere are taking an increasingly active role in the practice of medicine, because modern medicine involves so many value-laden decisions about how to use the material resources available for extended and costly health care on an individual and collective basis (see *Time* 1990). Medical ethicists, for example, are already deeply involved in the increasing controversy over the right to die, a controversy which spending money cannot settle (Gervais 1986).

But no applied ethicists consider the implications of the incentive system within the medical community itself. Doctors are symbolically and materially rewarded for discovering new diseases and for diagnosing their onset at the earliest possible stages; but they do not have to pay for assistance for those who are invited to think of themselves as sick. Instead they benefit from increasingly 'lax' or 'humanitarian' criteria depending on the philosophical point of view. Recently medicine has admitted four new disabling diseases to the official list of diseases: stress, depression, alcoholism and drug addiction. These conditions, which used to be personal problems or moral failings, now qualify sufferers for community assistance, sometimes of an extended and costly nature.

It would be a mistake to think that these new diseases are not really diseases. Although epidemiological research is still not definitive, both stress and depression seem to increase the probability of dying from chronic diseases like heart failure or cancer (Joint Economic Committee 1984), although their fatal character can only be detected over a comparatively long time. But people who live in less affluent countries, and who suffer from the same symptoms, will not be recognized as suffering from any of these diseases until their societies can afford to be more generous. In the meanwhile, the afflicted will have to keep working until they collapse, unless their relatives will support them in their distress. In contrast, Californians who are diagnosed as suffering from stress or depression can now apply for state-funded disability leave for three or four months, while receiving 70 to 90 per cent of their normal pay. Under these circumstances, the extent and duration of the disabling conditions will increase until the funds run out according to some culturally determined criteria.

In sum, nothing at either the micro or the macro level supports the stabilization or compression of morbidity, including disability, in modern, developed societies; everything fosters the continued increase of the non-severe types of morbidity, including non-severe disability (see La Plante 1989). The only controls which can limit this process are cultural, and include the evolution of new norms for perceiving and responding to 'disease'.

The health transition in historical context

There are advantages in taking a greatly simplified before-after approach to the role of cultural influences on the multidimensional history of morbidity during the health transition, but it is no substitute for detailed historical analysis, which focuses on the specific contributions of various cultural differences to the history of health in particular countries. Such research is extremely important, because the historical truth of the 'big picture' sketched above can only be verified by detailed historical

research which is context-specific. But the focus on context also reveals many national or local variations on common themes. Inevitably it will also turn up counter examples which must be dealt with. But detailed research alone rarely provides the grand themes most relevant to applied, policy-oriented history, which can be defined as the attempt to think in time, or extend the experience of one area to another (Neustadt & May 1986). All too often, specialist research provides the mind with so many separate windows into the past, that the big picture gets lost from sight. Thus, for many decades historians have worked on the health history of developing Europe without being forced to think systematically about what it means for the health history of the currently developing countries.

Although quantification is often presented as the best approach to deriving generalizations useful for policy related purposes, the history of Europe's health transition cannot be told in 'hard' non-cultural quantitative terms. For example, by the time that case-fatality rates were devised in the nineteenth century, largely in connection with the public health campaigns to control acute infectious disease, the diseases which were tracked had a normal duration of six to eight weeks before they ended in recovery or death, rather than a course which was limited to a week or less between onset and recovery (Riley 1990:165). This fact alone indicates that far more resources had been created and channelled towards the management of disease than was typical in much poorer societies.

Even with the limited amount of quantitative data for the early and middle phases of the health transition, we can only make informed guesses about the incidence and prevalence of disease in traditional, pre-industrial Europe, and we can only speculate about the real extent of co-morbidity among traditional peasants and early industrial urban workers. Thus, we have no idea how long the average eighteenth-century adult European peasant was sick before disrupting his or her work routines, even though at the time many physicians already suspected that the average European peasant suffered from a number of simultaneous diseases, which were simply not recognized as a legitimate reason to be sick (Peter 1975). As clinicians, however, they had no recourse but to identify, treat and count the single disease most likely to cause death in a short time: the most severe disease.

Morbidity began to be carefully measured in Europe in connection with adult males who were members of privately organized societies designed to provide financial benefits for those unable to work. It was not until the early twentieth century that morbidity, as a measurable phenomenon, began to be tracked among children, housewives or older people. The absence of long-run quantitative data on morbidity for all age groups and both sexes has justified the scholarly resort to other forms of 'hard' data which can be used to proxy the course of morbidity during the beginning and middle phases of the health transition, and thus relate it to mortality.

But are the data available for studying morbidity by proxy any less cultural than actual morbidity data? The most quantitative proxies for morbidity trends have focused on the amount of money spent on health and disease (directly and indirectly), or on the mortality data as a proxy for real health. But using either kind of data to proxy morbidity continues to beg all the important cultural questions about what expenditures on health have to do with morbidity, and how death itself relates to disease, particularly to chronic disease.

As development proceeded, more and more community sponsored support was eventually extended to the sick and disabled, including the elderly who were widely regarded as normally in a sick or disabled state. The fact that the amount of money spent on sickness and disability rose during the health transition suggests that morbidity and mortality are inversely related. In general, a narrowly economic approach to health can also make it look as if rising incomes cause the inflation of morbidity, because they are correlated over time.

In cultural reality the extension of community support to the sick and disabled is a complex story in which the leading community actors are not necessarily the richest people in their national

neighbourhood. Germany and France were usually well ahead of England and the United States, both much wealthier countries, in terms of providing publicly funded welfare benefits for the sick, disabled and old. England, with a much higher level of per capita income, was usually last to provide the least generous benefits. In England, pregnant women employed in market work received no paid leave for sickness until 1911, several decades after such benefits had been granted in France and Germany; and even then, sickness benefits related to pregnancy were closely monitored and frequently denied. Part of the problem was the biological confusion over what constituted genuine morbidity during pregnancy; but this confusion was compounded by the deeply ingrained prejudice in England against letting the community assume responsibility for the care of the sick, a responsibility which so many people believed should be left to families (Bonfield 1914).

Generally, those who believed in family responsibility for the care of the sick fought for very stringent definitions of what sickness constituted; those who were philosophically inclined to let the state assume increasing responsibility for the ordinary individual's health, were in favour of less and less stringent criteria for sickness. Since the criteria used to determine morbidity differed from place to place, the per capita amounts of money paid out to assist the sick and disabled also varied. But these variations tell us much more about the cultural supports for morbidity than they do about comparative health. Therefore per capita or other macro-level expenditures on health are not an objective proxy for morbidity trends in developing Europe, and ought not be used as a substitute for them in comparative research without a great deal of careful thought. Similarly, the idea that sickness benefits represent some objective biological measure of sickness at the micro level is extremely naive, because it overlooks the struggle that routinely went on between the potentially sick (those who formally applied for benefits) and those who were in charge of distributing benefits according to socially constructed criteria.

Problems like these drive historical demographers back to mortality data as the best observable proxy for biologically real but unobservable morbidity trends. Unless it is conceivable that poor health positively promotes rising life expectancy, the decline of mortality must mean that there is less poor health and less real sickness as the health transition proceeds. If morbidity, like mortality, had only one dimension, it might be wise to leave it at that; but this is not possible. If the mortality transition involves the replacement of fast-killing diseases by slower-acting less severe diseases, the net burden of morbidity could indeed increase and thus legitimately inflate community health care costs as mortality declines.

The most interesting problems with respect to the use of mortality data as a proxy for morbidity involve the interpretation of cause-of-death data. Gathered since the middle of the nineteenth century, cause-of-death data have been used by some demographers to show that a certain kind of contradictory progress was intrinsic to the health transition. At the core of the idea of an epidemiological transition is the gradual substitution of chronic but slow killers for acute infectious diseases among the leading causes of death. But other demographers are not so sure that the real but unrecorded burden of chronic disease is faithfully reflected in cause of death data. They point out that heart disease and other chronic forms of illness were simply lost among the substantial proportion of deaths initially reported as 'indefinite and unknown' during the early stages of single-cause reporting (Preston 1976:28–31). If it was not easy for nineteenth-century physicians to tell the difference between infectious diseases like tuberculosis and various chronic forms of bronchitis, it must have been even more difficult to accurately diagnose the real extent of most chronic diseases and the role they played in causing death (Szreter 1988). It is possible that deaths caused by, or complicated by, chronic diseases like heart disease, cancer and diabetes may have been more common in the late nineteenth century than they are today. Diabetes was not discovered as a disease until the late nineteenth century, and before the 1920s,

when insulin was developed and made available, doctors had little incentive to diagnose or report it. Thus a rising proportion of deaths caused by chronic diseases like diabetes is not a meaningful indicator of real trends in its incidence, until medical reporting is thought to be complete. The same kind of logic extends to heart disease and even to certain forms of cancer.

In other words, the real extent of the chronic diseases before and after the health transition is not easy to determine using cause-of-death data. Nor can the failure or reluctance to report multiple causes of death on most death certificates be interpreted as an indication that co-morbidity was once less extensive than it is today. The human body has always been open to multiple forms of exposure, including bacterial diseases, viral diseases, any one of a number of other macro parasites, and inorganic pathogens like toxic substances and excessive radiation. Even now, if we were fully aware of the extent to which our bodies are constantly exposed to disease-causing agents, and constantly resisting their effects, we would all think of ourselves as suffering from many diseases at once (Sell 1987). Very probably, in the early phases of the health transition, more individual deaths were caused by two or more relatively severe diseases, experienced simultaneously, than at present. But as long as doctors were forced to identify and treat only the most obviously fatal diseases, they tended to report as principal causes of death only those diseases which had the most clearly defined symptoms, particularly if an individual was being treated for that disease at the time of death. That conclusion cannot be used to justify the belief that less reporting of chronic disease meant less chronic disease, or less extensive co-morbidity. It probably means that those people who suffered from the chronic diseases, which in most instances are intrinsic to the aging process itself, were forced to carry on, working and living in pain and slowly growing weaker, until they developed some acute manifestation of disease, which legitimized the disruption of their routine and clearly signalled the need for treatment.

Some forms of chronic disease have actually increased in the last century. But the extent and nature of such increases requires careful investigation, including a review of historical data, in both qualitative and quantitative forms. The relationship in the developed countries between dietary patterns and high reported levels of heart disease is still medically and epidemiologically problematic. For a review of the literature see Smil 1989. It is possible that as the fat content of the average diet increased, the real level of heart disease also increased. Similarly there seems little reason to doubt that as tobacco consumption increased, the real incidence of lung cancer increased as well. In general using only cause-of-death data to study the history of morbidity trends is not a good basis for deriving conclusions about the biological magnitude and direction of changes in disease levels during the health transition. Ultimately there are no good 'objective' proxies for morbidity trends. Fiscal and mortality derived proxies for morbidity are as cultural as they are biological.

These lessons about the health transition and the historical measurement of morbidity in developing Europe apply with equal force to developing countries today. In the still developing countries per capita health care expenditures vary, so do reported cause-of-death patterns and life expectancy levels, but what the patterns of observable variation are telling us about unobservable morbidity levels remains unclear. For that reason, quantitative data from currently developing countries cannot be interpreted intelligently without a great deal of cultural clarification, almost on a case-by-case basis. Culturally insensitive applications of regression analysis are not likely to have much empirical value for modern or historical analysis. The study of morbidity levels during the world's continuing health transition should start at the micro level, with how ordinary people are taught to recognize sickness, and what they should do about their perceptions. Subsequently research must proceed at the institutionalized level where health-care professionals are trained to detect and report diseases to officials who are charged with collecting and publishing that data. Morbidity data in turn are only gathered because higher-level officials have assigned priority to it as part of a country's official

policies about health, and these policies themselves have deep philosophical-cultural roots in the value assigned to the preservation of human life, and political-cultural roots in how the value of human life, and therefore of human health, is differentially conceptualized in both democratic and hierarchical political systems.

Therefore, the cultural foundations of health and sickness cannot be ignored, because they are implicit in the very idea of sickness and health, and even in the hardest quantitative data on the economics of health and the causes of death. That is what makes the cultural history of health and morbidity so relevant to present policy issues in both the developed and the developing countries. If developing countries follow the path of the developed countries they will encounter the same inflation of morbidity those countries are already experiencing. That is, if public health reforms are made along standard lines, including the bureaucratized gathering of morbidity data, and health professionals in the developing countries are trained along modern Western lines; if ordinary people are educated to accept modern theories of disease and to act accordingly, by seeking treatment from formally trained doctors instead of traditional healers; and if all this takes place in the context of rising per capita incomes; the inflation of morbidity, particularly in the form of prevalence rates and co-morbidity rates, is inevitable.

We already have limited evidence that this type of inflation is under way. Kerala is one of the poorest states in India, a country which itself is one of the poorest countries in the world. Nevertheless Kerala state, by efficiently institutionalizing public health reforms and controlling infectious disease, has achieved a life expectancy at birth of 70 years. Although this is by far the highest level of life expectancy in India, Kerala's people also report the highest levels of morbidity, according to government health surveys (Bhattacharya, Chatterjee & Pal 1988; Sen 1990). This is only a problem if the inflation of morbidity is interpreted in a negative light, instead of as a positive sign that the health transition, organized on Western lines, is proceeding along its expected course.

The problem of interpreting morbidity data is even more acute at the local level where small projects designed to improve local health conditions need to be periodically evaluated. Pebley (1990) touched on the general difficulty of evaluating morbidity data from rural areas of developing countries when the data sets are small, and short-term trends show no clear form of change. In general, the organizations doing this kind of research in developing countries do not seem to be taking enough care to distinguish between improved health as an underlying form of biological change, and improved health as a culturally perceived and reported form of change. If they took an historically informed perspective on the course of morbidity change during the health transition, they would look for evidence in the small-scale, short-run, community-level data sets, which are notoriously difficult to interpret under the best circumstances, that the morbidity levels of the targeted population were rising, not falling, as human lives were ameliorated by educational and economic assistance (but see Wirsing 1985).

If the developing countries do not want to adopt policies which are virtually certain to raise reported morbidity levels, then they must change the way people are being trained to perceive sickness and report disease, or change the incentive system to which modern doctors are responding, or change the way the data provided by health professionals, or people themselves, are gathered and published during the health transition. On the other hand, those countries which make accelerating the health transition a top policy priority, can learn to interpret rising morbidity as a sign that they are achieving their goals, provided that mortality continues to decline, and less severe disease is being reported.

Japan might be the most appropriate model for such health-conscious developing countries, since its particularly rapid rise in life expectancy after the end of the Second World War was indeed accompanied by an even more dramatic rise in reported morbidity at all ages (*Statistical Yearbook*). This rise, particularly as it is reflected in health surveys, is clearly a function of the mass

reprogramming of the ordinary Japanese mind. People in Japan now think of themselves and their children as afflicted with all manner of previously unrecognized diseases, like 'one-day flu' and high blood pressure, which their ancestors had, but without being aware of it. High blood pressure, thought to be linked to the high amounts of salt long consumed by Japanese, is certainly linked to the disproportionately high number of adults in Japan who die from stroke and its complications. But high blood pressure itself is a non-severe disease which can last for many years before the sufferer dies of either stroke or some unrelated cause. Once it became recognized and reported as a disease after 1950, however, it permanently inflated prevalence rates in Japan, particularly among adults past middle age. Meanwhile, modern Japanese have become convinced that they or their children are sick when they get a cold or influenza which lasts for one or two days. The level of reported morbidity in Japanese health surveys now fluctuates annually in response to the number of people who happen to have very mild diseases of this nature. But colds and 'flu' can inflate incidence rates quite dramatically in the short run while prevalence rates continue their slow rise in response to the increasing public awareness of chronic diseases. Meanwhile both forms of rising morbidity have less to do with short-run changes in mortality. Mortality just keeps declining as reported morbidity soars.

Given the exceptional position Japan occupies in the history of world development its historical experience with mortality and morbidity is an unusually good source of data which fills in the descriptive picture of the health transition (Johansson & Mosk 1987a). Japan started both its mortality and morbidity transitions when Europe's had been under way for a century or more. Borrowing from the West, post-Meiji Japanese officials began the modernization of their country by fully appreciating the need for certain kinds of explicit data on the incidence and prevalence of infectious disease. Thus there are case-fatality rates for various infectious diseases which permit us to track their incidence and severity as well as their relationship to mortality in surprising detail (Johansson & Mosk 1987b). But as soon as the government succeeded in reducing the incidence of most of the major infectious diseases by 1900, bureaucrats whose function was to gather morbidity data began to track the less acute infectious diseases, most notably tuberculosis.

At the time most Japanese factory workers (the majority of whom were very young girls working under labour contracts) exhibited one or more symptoms which indicated that they had already contracted tuberculosis. If employers were prepared to grant leave, even unpaid leave, for one or more of these symptoms, most of their workforce would have been in factory-run hospitals or at home, rather than at work in the factory. Since it was costly to recruit and train new workers, employers were reluctant to let individuals leave the factory before their labour contracts expired. Thus cultural means had to be devised and negotiated with reforming officials for determining what kinds of symptoms meant that tuberculosis was severe enough to justify time off work. Only the most acutely ill workers were sent home or given treatment in the company infirmary. Some died there, but some began to recover. When their symptoms abated, recovering workers were forced to return to the factory floor, often wearing a mask on their mouths to prevent them infecting other workers. Thus as tuberculosis began to spread rapidly among the factory workers of Meiji Japan, they were generally forced to keep on working, even if they were coughing and spitting blood, but able to carry on their simple repetitive tasks. Only complete physical collapse excused them from work, or got them some medical attention.

Thus the earliest morbidity data produced in Japan make it look as if textile workers were frequently ill, but only for short periods. But these 'facts' about the duration of illness among textile workers are not facts about the biology of tuberculosis or other infectious diseases: they indicate how long factory owners were prepared to support sick workers by keeping them in the factory infirmary, before insisting on the resumption of work, or the loss of employment (Johansson & Mosk 1987a). The government tolerated these harsh standards because they were also unwilling to provide increasing

community support to those who suffered from tuberculosis or the chronic diseases. Thus the government repeatedly said it could not afford to fund tuberculosis sanatoria, meaning that it accorded a much higher priority to funding an expensive military buildup. In fact improving the health of the nation was held to be fiscally justifiable only if it improved the country's ability to fight future wars.

In the meantime, although reported morbidity levels in early twentieth century Japan were lower than they are today, the 'real' health status of the average peasant was obviously much worse than that of the average urban worker today. This conclusion is strongly suggested by an elaborate medical survey carried out by Western-trained Japanese doctors in 134 representative Japanese villages from 1918 to 1928, when life expectancy at birth was 40–45 years. Upon examining the adult villagers, both men and women, the doctors concluded that the 138,462 adults examined had 214,137 cases of serious disease, or 1.5 reported diseases per person. Three quarters of those examined had a parasitic disease, and many had more than one.⁵ Leprosy was still a serious health problem. On the whole, only 10 per cent of the entire set of peasants were considered disease free by the doctors' standards; and the doctors were not yet looking for cancer, heart disease, high blood pressure or venereal diseases (cf. Embree 1939). But by local standards very few villagers were sick, in the sense of being willing to disrupt their daily work routines (Ryoichi 1937). For one thing, most adults could not afford to withdraw from field work or domestic production. Most adults routinely accepted a certain amount of pain as the price of ordinary survival, just as they disregarded the fact that most of their children had lice, and were rarely free of runny noses and other signs which indicated suboptimal health. For older people, even moderate amounts of disability were regarded as a normal part of life, rather than a sign of illness. Work was simply reassigned to fit the abilities of aging parents with limited abilities.

The history of health in pre-war Japan reassures us that the rise of reported morbidity in post-war Japan is not really the result of the massively deteriorating health of the ordinary adult and child, but of the culturally influenced changing nature of perceived and reported illness. Other developing countries with a non-European cultural background cannot expect to escape the cultural inflation of morbidity during the health transition if they educate their citizens, doctors and bureaucrats along Western lines.

Those Western countries which were uniquely culturally configured to initiate the health transition in the first place must continue to consider how to manage morbidity in cultural terms. Given the nature of modern life, it would probably be impossible to cap the expansion of knowledge about health and disease, or to undermine the incentive system which encourages researchers and practitioners to detect and treat as many new diseases as possible at their earliest possible stages.

It is also doubtful that current health expectations could be lowered by promoting more stoic fatalism about the inevitability of pain and disability, even in old age. By nineteenth century standards the ordinary senior citizen may be in exceptional health today, but historical standards have no relevance to people who did not live them. To people who feel energetic and free of pain most of their lives, any deviation from that state constitutes a real loss of health. In a culture that promotes the idea that everyone has the same right to good health, the old are not likely to relinquish their rights on account of age.

⁵ Extensive parasitism was also documented in the American south in the early twentieth century. In the southern states the rural population was heavily infested with parasitic diseases related to widespread vitamin deficiency disease. Nevertheless even children who by modern standards would be considered sick, were judged well enough to work on the farm or in factories by their own parents. See Kunitz, 1988. Parasitologists often argue that most parasitic diseases are not severe enough to cause their host's death (E. Barrett-Conner, personal communication). But the parasitic diseases are currently called the great neglected diseases precisely because they have not been intensively studied by modern scientists. For example, we do not have good data about the role of parasitic diseases (as a form of co-morbidity) in hastening death from other causes.

Barring any changes in the cultural aspects of perception, detection and reporting, it seems that health care must be rationed through some culturally devised means. The developed countries which provide virtually free medical care to their citizens have already imposed covert forms of rationing by extending the waiting times for 'free' medical care, especially for procedures requiring hospitalization and surgery. In the United States medical care is openly allocated on the basis of income level, which is correlated with ability to buy private health insurance. Only the extremely poor receive truly free medical care (through Medicaid) but the quality of that care is very uneven. If the newly devised Oregon plan diffuses to other states, health care for the poorest will be rationed.⁶ Meanwhile, the less generous and more stringent disability programs become, the fewer disadvantaged people will apply for assistance. In fact only the most desperate will seek community assistance; the more stoical individuals will indeed continue to work in pain until their health breaks down, and they receive emergency care for near-death conditions. High-income groups will continue to get good health care, and low income groups will not. But this situation is not something a very wealthy country can officially enshrine in the form of an explicit policy unless some formal justification could be provided for differential health care, an unlikely event in a democratic society.

But do modern conditions justify the rationing of health care, for either the poor or the not-rich? From the narrowest economic standpoint the answer appears to be yes. Developed countries such as Canada, West Germany Japan, Sweden, the United Kingdom and the United States, which once spent less than 2 per cent of their gross domestic product on health care, currently spend from 6 to 11 per cent. There is no sign that these rising levels of expenditure will level off (Verbrugge 1989; Rice 1990). At present the United States, for example, spends more money on protecting health and managing morbidity than any other nation. In 1986 its expenditures totalled 458 billion dollars; on a per capita basis this equalled \$1,837 (National Center for Health Statistics 1989; see also Lee & Etheredge 1989).

In constant dollars the amount of money the United States spends on health is about equal to the entire per capita income of developing countries such as Mexico and Brazil (World Bank 1978). Yet life expectancy at birth levels in these countries, 66 and 65 years respectively, are only ten to twelve years below current United States levels, and equal to life expectancy levels for American males of African descent. These disproportions between expenditures on health and life expectancy at birth are so enormous that it becomes natural to wonder what the United States is buying for its money that Mexico is not. Without a theory which explains the cultural inflation of morbidity it becomes necessary to suspect that the United States is wasting vast amounts of money on unnecessary or inefficiently delivered health care (Fuchs 1990). While this is true to a certain extent, the United States also pays for diseases which no longer have a discernible statistical relationship to longevity. Instead these diseases and forms of disability may be merely painful, including mentally painful; or they may represent the attempt to detect or prevent the development of one or more diseases in their earliest stages; or they may represent the attempt to address medically various forms of addiction, now reconceptualized as diseases instead of moral failure for which the individual is personally responsible.

At any rate, in 1988, 46 per cent of those Americans who answered a question about how satisfied or dissatisfied they were with health conditions in their country said they were 'very satisfied': in Japan which spends much less on health, only 14 per cent answered the same question in the same way. Similarly only 13 per cent of Americans included in a five-country survey said they were 'fairly' or 'very dissatisfied' with health conditions, while in Japan 26 per cent reported fair or high levels of

⁶ Kitzhaber 1990. The Oregon plan assigns priorities to the kinds of health care the state will provide free to the poorest individuals and families, based on the perceived reality of fiscal limits. As expected, the forms of health care which are accorded the highest priority are those which are the least expensive and most likely to reduce mortality.

dissatisfaction (Suzuki 1989). No other developed country reported comparable levels of satisfaction as high as those in the United States except for Great Britain. Since Japan has a higher level of life expectancy at birth than the United States, but a citizenry which reports much lower levels of satisfaction with health conditions, many Japanese people may indeed be working and living in pain to a greater extent than is true in the United States.

Is it possible that the United States has gone further than most developed countries in matching the high health expectations of modern people? Only further culturally oriented research could answer that question. But if it were true, how could the government refuse to stop meeting high health expectations, even if they are rising? Currently, the American government spends six to eight billion dollars paying some farmers not to grow food, and other farmers to grow such crops as sugar which could be raised elsewhere at a lower cost, or tobacco which is known to cause deadly forms of chronic disease. Since it is willing to invest its resources in this way, it could hardly refuse, through explicitly formulated policies, to subsidize the kind of health care which its citizens, especially its older ones, have worked hard to be able to afford, and which they show every sign of wanting in greater measure (Schneider 1990).

While there is no clear moral basis for rationing health care in very high per capita income countries, there is some basis for wondering how much health care would be 'enough' for people who increasingly believe that any deviation from a very high standard of health justifies the costly attention of a trained professional, and social benefits such as paid sick leave, and early retirement. Thus there is a great need for formally developed philosophies of health care, and more trained medical ethicists taught how to apply those philosophies to concrete decision making.

In the meanwhile there is still potential but unexpressed demand for health care among selected disadvantaged subpopulations in a rich country like the United States. Ethnically distinct subpopulations often cling to health standards and practices that prevailed in earlier decades. Accordingly, they use preventive or regular health care services less frequently than hospital emergency services designed to treat life-threatening conditions, even if non-emergency forms of health care are low-cost or free. Other things being equal, if disadvantaged Americans adopted the culturally influenced standards and practices of advantaged Americans, reported morbidity would rise still further, and so would the costs of coping with increasing utilization rates. In addition to the kinds of cultural differences associated with class, there are culturally derived influences on morbidity linked to race (Freund 1982). In general, African-American males have all the same health problems that White males have, but, in addition, they suffer from violence-related morbidity at much higher rates. Not only are Black men and women wounded in violent incidents more often than Whites, but African-American men who live in New York's Harlem are also murdered at ten times the rate of other men.⁷

This is not just a simple economic side-effect of poverty: other equally income-disadvantaged subpopulations have much lower levels of injury and violent death. The high level of violence which makes ghettos more dangerous to young and adult males than a Vietnam battlefield, has evolved as part of a cultural response to pervasive racism. But once evolved, cultural arrangements have a life of their own, which can only be changed by a recognition that certain fundamental cultural reforms, directed at reducing and reversing the effects of past racism, will have to be undertaken if violent morbidity and mortality are to be brought under social control.

⁷ McCord & Freeman 1990:174. Black men in Harlem have a lower probability of reaching age 65 than men in Bangladesh, where life expectancy at birth is 49 years. More of Harlem's men survive to age 20 than in Bangladesh, but subsequently they do much worse. The leading causes of premature death in Harlem are cardiovascular disease and cirrhosis, accounting for about 40 per cent of deaths, and homicide, another 15 per cent.

To a lesser extent cultural differences continue to exert an influence on excess mortality among all men, irrespective of their race. Statistically, excess male mortality has been linked to male lifestyles which encourage smoking, drinking and other forms of health risk taking. But men are also still less likely to seek medical help than women, and to receive treatment in the later stages of the diseases from which they die. These poor health habits contribute to the fact that, although men earn higher incomes and control more resources than women in developed economies, they nevertheless have a substantially lower life expectancy at birth (Verbrugge & Wingard 1987). If men in the developed countries adopted the same modernized health standards and practices as women, this would also increase morbidity levels and health care costs; but it might also be a small step towards closing the life expectancy gap between the two sexes. Larger steps would involve the virtual eradication of culturally rooted lifestyle differences between men and women.

Given the intrinsically cultural nature of differential morbidity and mortality, policies designed to minimize differences between subpopulations must be explicitly cultural. So must policies designed to minimize differences between nations. It is difficult to deny that the world's death rates are unprecedentedly low because scientific knowledge and technology were used to understand the biological causes of various infectious diseases, and subsequently applied to preventing them, curing them, or minimizing the damage they do. Preston (1976) suggests that the cost of eliminating infectious and parasitic disease has been falling in the twentieth century, and that low-income countries can afford the elimination of these diseases if they have the political will. Economic development created the material resources needed to support the reduction of exposure to disease, and to enhance the body's natural capacity to resist disease. But rising levels of knowledge, technology and income only force us to face more squarely the cultural foundations of health, which generally tend to inflate morbidity as mortality declines.

Thus, from the cultural standpoint the history of the health transition is not a simple demographic story about declining mortality, nor is it a straightforward scientific and epidemiological story about the biology of disease, and the means of controlling and reporting it. Nor is it a sad story about the physiological penalties of saving the innately frail, nor is it a rational economic story about the induced effects of an increasing ability to pay for sickness and disability. The interdisciplinary history of the health transition is about how all of the above are linked over time under cultural management. This story, if told at a sufficiently general cultural level, can transform rising morbidity into a positive sign of increasing human welfare, which governments should continue to anticipate and willingly fund.

References

- Alter, G. & Riley, J. 1986, The historical relationship of morbidity and mortality, Paper presented to the Social Science History Association, St Louis, October.
- Bang, R.A. et al. 1989, 'High prevalence of gynaecological diseases in rural Indian women', *Lancet*, 8629:85-88.
- Barry, G., Bronowski, J., Risher, J. & Huxley, J. 1963, *Health and Wealth: Man's Fight Against Sickness and Want*, Macdonald, London.
- Bhattacharya, H., Chatterjee, G.A. & Pal, P. 1988, 'Variations in level of living across regions and social groups in India', in *Rural Poverty in South Asia*, eds T.N. Srinivasan and R. Bardhan, Columbia University Press, New York.
- Bonfield, M. 1914, *The National Care of Maternity: A Scheme Put Forward as a Basis for Discussion*, London.
- Caldwell, J.C. 1990, 'Introductory thoughts on health transition', in Caldwell et al. 1990.
- Caldwell, J.C., Findley, S., Caldwell, P., Santow, G., Cosford, W., Braid, J. and Broers-Freeman, D. 1990, *What We Know about Health Transition: The Cultural, Social and Behavioural Determinants of Health*. The

- Proceedings of an International Workshop, Canberra, May 1989*, 2 vols, Australian National University, Canberra.
- Callahan, D. 1987, *Setting Limits: Medical Goals in an Aging Society*, Simon & Schuster, New York.
- Das, V. 1990, 'What do we mean by health?', in Caldwell et al. 1990.
- Dasgupta, P. n.d., *An Inquiry into Well-Being and Destitution*, forthcoming.
- Dennett, G. & Connell, J. 1988, 'Culture and health in the Highlands of Papua New Guinea', *Current Anthropology*, 29:273–299.
- Doan, D.D.H. (ed.) 1985, *The Future of Health and Health Care Systems in the Industrialized Societies*, Praeger, New York.
- Dubos, R. 1968, *Man, Medicine and Environment*, Mentor Books, New York.
- Embree, J. 1939, *Suye Mura, a Japanese Village*, University of Chicago Press, Chicago.
- Fox, J., Hall, C. & Elveback, L. 1970, *Epidemiology: Man and Disease*, Macmillan, Toronto.
- Freund, P. 1982, *The Civilized Body: Social Domination, Control and Health*, Temple University Press, Philadelphia.
- Fried, L. 1990, The incidence of disability: the bottom of the iceberg and the key to prevention, Paper presented to National Institute on Aging Conference on Compression of Morbidity, Asilomar Conference Center, California, March 18–20.
- Fries, J. 1980, 'Aging, natural death and the compression of morbidity', *New England Journal of Medicine*, 303,3:130–135.
- Fuchs, V. 1990, 'The health sector's share of the Gross National Product', *Science*, 347:534–538.
- Gervais, K. 1986, *Redefining Death*, Yale University Press, New Haven.
- Goleman, D. 1991, 'Risk found in depressed nursing home patients', *New York Times*, Feb.27:4.
- Goodman, A. et al. 1988, 'Biocultural perspectives on stress in prehistoric, historical and contemporary population research', *Yearbook of Physical Anthropology*, 31:169–202.
- Gruenberg, E. 1977, 'The failures of success', *Milbank Memorial Fund Quarterly/Health & Society*, 55:3–24.
- Hansluwka, H. 1985, 'The changing health spectrum: reflections on the future', in Doan 1985.
- Harries, A.D. 1990, 'Tuberculosis and human immunodeficiency virus infection in developing countries', *Lancet*, 335:387–390.
- Helman, C. 1990, *Culture, Health and Illness*, 2nd edn, Wright, Butterworth & Co., London.
- Isaac, L. & Griffin, L. 1989, 'Ahistoricism in time-series-analyses of historical process: critique, redirection and illustrations from U.S. labor history', *American Sociological Review*, 54:873–890.
- Johansson, S.R. 1990, 'Cultural software, institutional hardware, and health information processing in social systems', in Caldwell et al. 1990.
- Johansson, S.R. 1991, 'Continuity and change in the explanation of male/female mortality differentials', *Continuity & Change*, forthcoming.
- Johansson, S.R. & Mosk, C. 1987a, 'Exposure, resistance and life expectancy: disease and death during the economic development of Japan', *Population Studies*, 41,2:207–235.
- Johansson, S.R. & Mosk, C. 1987b, Morbidity, mortality and their complex relationship in the Japanese experience with sickness and death in the twentieth century, Paper presented at Social Science History Association Conference, New Orleans, Oct. 31.
- Joint Economic Committee, U.S. Congress 1984, 'Economic change, physical illness, mental illness, and social deviance', Subcommittee on Economic Goals and Intergovernmental Policy, Government Printing Office, Washington, DC.

- Kaplan, G. 1990, Epidemiologic observations on the compression of morbidity, Paper presented to Conference on Compression of Morbidity, Asilomar, March 18–20.
- Kitzhaber, J. 1990, 'A healthier approach to health care', *Issues in Science and Technology*, 7:59–65.
- Kunitz, S.J. 1988, 'Hookworm and pellagra: exemplary diseases in the new South', *Journal of Health & Social Behavior*, 29:139–148.
- LaPlante, M. 1989, 'Disability risks of chronic illness and impairments', *Disability Statistics Report*, 2:1–38.
- Lau, R. & Klepper, S. 1988, 'The development of illness orientations in children age 6 through 12', *Journal of Health & Social Behavior*, 29:149–168.
- Lee, P. & Etheredge, L. 1989, 'Clinical freedom: two lessons for the UK from the US experience with privatization of health care', *Lancet*, Feb.4:263–265.
- Manton, K. & Stallard, E. 1984, *Recent Trends in Mortality Analysis*, Academic Press, Orlando.
- McCord, C. & Freeman, H. 1990, 'Excess mortality in Harlem', *New England Journal of Medicine*, 322:173–177.
- McKeown, T. 1988, *The Origins of Human Disease*, Basil Blackwell, Oxford.
- McNicol, G. 1990, 'Comments on policy aspects of health-transition research', in Caldwell et al. 1990.
- Micozzi, M. 1983, 'Health care', *Human Organization*, 42:351–353.
- Mull, J. & Mull, D. 1988, 'Mothers' concept of childhood diarrhoea in rural Pakistan: what ORT planners should know', *Social Science & Medicine*, 27:53–67.
- National Center for Health Statistics 1989, 'Health, United States, highlights 1988', *Publication No. (PHS) 89-1232*, Government Printing Office, Washington DC.
- Neustadt, R. & May, E. 1986. *Thinking in Time: The Uses of History for Decision Makers*, Free Press, New York.
- Omran, A.R. 1971, 'The epidemiological transition', *Milbank Memorial Fund Quarterly*, 49:509–538.
- Pebley, A. 1990, Women's health in rural Bangladesh, Paper presented at Stanford University, March 7.
- Peter, J. 1975, 'Disease and the sick at the end of the eighteenth century', in *Biology of Man in History*, eds R. Foster & O. Ranum, Johns Hopkins University Press, Baltimore.
- Peterson, D. & Thomas, D. 1978, *Fundamentals of Epidemiology*, D.C. Heath & Co. Lexington.
- Porter, R. & Porter, D. 1988, *In Sickness and in Health*, Oxford University Press, Oxford.
- Preston, S.H. 1976, *Mortality Patterns in National Populations*, Academic Press, New York.
- Palmore, E.B. 1987, 'Trends in the health of the aged', *Gerontologist*, 26:298–302.
- Rice, D. & La Plante, M. 1988, 'Chronic illness, disability and increasing longevity', in *The Economics and Ethics of Long-Term Care and Disability*, eds S. Sullivan & M. Lewin, American Enterprise Institute for Public Policy Research, Washington DC.
- Riley, J.C. 1987, 'Ill-health during the English mortality decline: the friendly societies experience', *Bulletin of the History of Medicine*, 61:563–588.
- Riley, J.C. 1990a, 'Long-term morbidity and mortality trends: inverse health transitions', in Caldwell et al. 1990.
- Riley, J.C. 1990b, 'The risk of being sick: morbidity trends in four countries', *Population & Development Review*, 16,3:403–432.
- Ruzicka, L.T. & Kane, P. 1990, 'Health transition: the course of morbidity and mortality', in Caldwell et al. 1990.
- Ryoichi, I. 1937, *Population Pressure and Economic Life in Japan*, University of Chicago Press, Chicago.
- Schach, E. 1985, 'Health status of the population: changes in the past and predictions for the future of life expectancies', in Doan 1985.
- Scheper-Hughes, N. & Lock, M. 1987, 'The mindful body: a prolegomenon to future work in medical anthropology', *Medical Anthropology Quarterly* (New Series), 1:6–41.

- Schneider, E. 1990, The options for responding to increased prevalence and costs of chronic disease disability, Paper presented at Conference on Compression of Morbidity, Asilomar, March 18–20.
- Sell, S. 1987, *Immunology, Immunopathology and Immunity*, 4th edn. Elsevier, New York.
- Sen, A. 1990, Values, capability and well-being in development assistance, Paper presented at Hoover Institute Conference on Wealth of Nations in the 20th Century, Stanford University, March 6–8.
- Sigerist, H. 1943, *Civilization and Disease*, University of Chicago Press, Chicago.
- Smil, V. 1989, 'Coronary heart disease, diet and Western mortality', *Population & Development Review*, 15,3:399–424.
- Statistical Yearbook of Japan*, various years, Tokyo.
- Suzuki, T. 1989, Cultural link analysis: its application to social attitudes – a study among five nations, Paper presented at ISI 47th Session, Paris, Aug.29–Sep.6.
- Szreter, S. 1988, 'The importance of social intervention in Britain's mortality decline c.1850–1914', *Social History of Medicine*, 1:1–38.
- Time* 1990, 'The right to die', March 19:62–68.
- Veatch, R. 1989, *Death, Dying and the Biological Revolution*, Yale University Press, New Haven.
- Verbrugge, L. 1989, 'Recent present and future health of American adults', *Annual Review of Public Health*, 10:333–361.
- Verbrugge, L. & Wingard, D. 1987, 'Sex differentials in health and mortality', *Women & Health*, 12:103–145.
- Wirsing, R. 1985, 'The health of traditional societies and the effects of acculturation', *Current Anthropology*, 26:303–322.
- World Bank 1987, *World Development Report 1987*, Washington DC.
- World Health Organization 1983, *Basic Documents*, 32nd edn. Geneva.
- Wynn, D. Rodriguez, M. O'Fallon, W. & Kurland, L. 1989, 'Update on the epidemiology of multiple sclerosis', *Mayo Clinic Proceedings*, 64:808–817.