

## Comment



### **From a high mortality regime to a high morbidity regime: is culture everything in sickness?**

**James C. Riley**

*History Department, Indiana University, Bloomington, IN 47405, USA*

In the inaugural number of *Health Transition Review*, S. Ryan Johansson notes that the health transition has been marked by declining mortality but rising morbidity and suggests that this can be explained as the result chiefly or exclusively of a cultural inflation of sickness (Johansson 1991). So much common sense lies behind the idea that cultural forces have played a role that, until Johansson's essay, no one had taken the trouble to specify which forces should be included under the rubric 'cultural'. Johansson provides a long and inclusive list of items that she believes capture changes in behaviour and belief, rather than in experience. She makes a conventional argument about them. Cultural forces, she claims, account for the inflation of morbidity, and the effect is sufficiently self-evident that no demonstration of it is needed.

To the contrary. While it is almost certainly true that changes in behaviour and belief account in some part for rising sickness, it is also true that non-cultural forces account for some part. Most recent research has been directed toward identifying non-cultural forces and assessing their effect. That research incorporates arguments that are more rigorous in their demands for evidence and logic than the arguments brought to bear by Johansson, who takes little note of the effects of non-cultural forces and who, in assessing the role of cultural forces, sets the task of merely showing what may be plausible. The argument for cultural forces needs not only to be specified in the quite useful way that Johansson has done by identifying certain forces. It also needs to be tested in the light of what is known about changes in beliefs and behaviours across periods of declining mortality and rising morbidity and across periods of rising mortality and declining morbidity. How much and in what precise manner are changes in behaviour and belief meant to have contributed to rising sickness? Furthermore, much of Johansson's argument relies on defining cultural forces so broadly that they include not only behaviours and beliefs but also a wide range of biological, medical, economic, social and even demographic forces. Cultural factors will count for everything if they are defined broadly enough. But broad definitions are misleading and vague and they impede the effort to address high morbidity as a human problem that may be alleviated, in a way analogous to the manner in which once-high mortality has been reduced.

Johansson's case for a chiefly cultural inflation of sickness is built on a series of assumptions and analogies, of which some are valid and helpful and others are invalid and misleading. In the nature of scholarly debate it is the invalid and misleading assumptions and analogies that deserve discussion in this note. But it is worth remarking that I have selected for discussion weak parts of an essay that has many strengths.

### **Inverse health transitions**

Material conditions improved in the developed countries during their mortality decline, which promoted higher expectations about health. People suffering sickness learned to perceive it better and became more willing to seek professional help. Medical practitioners learned to recognize a growing number of diseases that rarely or never cause death. Summarizing her essay in this way, Johansson sets forth a seemingly familiar group of changes in behaviour and belief that account for some part of changes in the incidence and duration of diseases and injuries. The problems, however, are these: are these forces correctly specified as cultural? Are they accurately stated? What part can they be shown to have played in the inverse health transition? On one hand, people now die later in life. On the other hand, holding age constant, people also more often regard themselves as sick, they take more time off from work, they spend not only more of their income on health services but also a rising proportion of their income (Riley 1990a). In these and other ways, too, people act as though their health deteriorated rather than improved during the mortality decline.

One of the signal contributions of Johansson's essay lies in its acknowledgement that sickness rates have increased. In countries that achieved low death rates early, such as most of western Europe; in those that began later but made more rapid progress toward low death rates, such as Japan; and in regions that have only recently achieved low death rates, such as the Indian state of Kerala, sickness prevalence has increased. If people now live longer, it is tempting to suppose that in the process they escape health hazards that their predecessors, who lived shorter lives, did not. Hence their additional sickness may be illusory. In terms of the threat to survival that each sickness poses, the additional sickness is illusory. But the tendency people have shown more often to act as though they are sick, even to the point of sacrificing income and enlarging their health spending, argues that the illusion lies not in behaviour but in the way sickness is graded. A gauge often used to distinguish serious sicknesses from other sicknesses is the threat a given malady poses to survival. In the traditional scheme of things diseases with high case-fatality rates were rated as important. Beginning in western Europe in the eighteenth century public authorities devoted a rising share of resources to sanitary and public-health measures. To a substantial degree those investments paid off, reducing mortality and lowering the average lethality of episodes of infectious disease. But people did not respond to the mortality decline by being well most of the time. Instead they responded to it by believing and acting most of the time as though they were sick. The central problem is to discover why this happened.

### **Behaviours and beliefs**

Using lethality as a gauge served public authorities and individuals well in an era in which the leading health losses consisted of deaths so premature that they made life expectancy at birth little more than 30 years. But already at the beginning of the health transition, which in western Europe was the eighteenth century, individuals were not acting as though lethality was the leading motive behind the decisions they made about health status and care. Already the ratio of medical practitioners to people – combining the licensed practitioners consulted by wealthier people together with the unlicensed practitioners and apothecaries that ordinary people consulted – was close to or above the modern ratio. Already people offering remedies of various kinds advertised their wares aggressively, using claims and appeals quite similar to those employed by their modern counterparts (Ramsey 1988). There was plenty of literature giving people advice about health maintenance and a range of home encyclopaedias describing commonplace diseases and remedies for them. What is more, people expected clergymen and socially prestigious fellow citizens to be able to provide them with useful health advice and care, supplementing that received from apothecaries, folk healers, physicians, surgeons and others.

Before the health transition began, individuals took a lively interest in their health, and strove to maintain it when well and to regain it when sick. They fretted about infections that, in their experience, might result in death. They fretted also, and attempted to treat, many maladies that could not be expected to result in death (Porter and Porter 1988). Ralph Josselin, a seventeenth-century clergyman, who kept a diary in which he commented on the health experiences of himself, his wife and children and a wide range of other matters, recorded colds as readily as he did smallpox and even took the trouble to notice occasions when his eyes gave off a watery discharge (Macfarlane 1976). The same story can be retold many times for many different people who provide us with autobiographical accounts of their health (Porter and Porter 1988). It was in the eighteenth century that hypochondria was transformed from a disease of the abdomen into an imagined ailment with no anatomical location. In short, in that high-mortality regime ordinary people were not guided in their thinking about sickness and wellness by the idea that potentially fatal diseases bulked much larger in importance than other ailments. Their diaries, correspondence and letters show that they wished to avoid discomfort and pain. The same sources reveal how difficult it was to avoid discomfort and pain because, when compared to modern remedies, they show how little the available advice and therapy did to restore people to health.

Nevertheless the important point, from the perspective of the cultural inflation of sickness, is this: before the health transition started, west Europeans fretted about their health, tested themselves for signs of sickness, bought medications and medical services and in other ways acted the part of people seriously concerned about their own health. Their expectations about health were already high and carefully formulated. It is almost certainly true that those expectations could increase and that they have increased since the eighteenth century. It is an error to suppose that health expectations began to increase from circumstances in which people expected to be sick and, when sick, disregarded their ailments. In short, the boundaries for the merely cultural inflation of sickness are in many ways rather narrow. If the point is to compare the mortality and morbidity experience of people in countries now developed before and after the health transition, it is essential to explore the culture of sickness and wellness in both periods.

It is seriously misleading to make inferences about health experience in eighteenth-century Europe from what is surmised about health experience in developing regions of the globe in the mid-twentieth century, as Johansson does, referring, for example, to a practice attributed to 'various parts of Africa' in the 1960s according to which people sick for more than three days were denied aid and comfort. Whether or not this is true about some parts of Africa in the 1960s, it is untrue about Europe in the eighteenth century or at any point since. Moreover, the profile of diseases that people suffer, the ways disease is treated and attitudes toward recovery are profoundly different between high-mortality regions of the present-day world and Europe before its mortality decline began. Long before the eighteenth century European folk medicine had incorporated Hippocratic and Galenic advice about what people should do when they are sick, and the remedies that people used were often efficacious in the terms of those regimes. That is, they had the desired and expected result: they made the bowels move and they caused people to sweat, bleed and vomit. Long before the mortality decline began people attempted to treat their maladies and reacted to ailments in ways that can readily be accommodated within the modern range of reactions: self medication, consulting friends and associates, taking time off from work, seeking professional help and seeking institutional care.

### **Cultural forces and policy choices**

In the developed countries of the world the health transition has progressed to the point at which it is now feasible to discuss redirecting attention from a primary focus on efforts to reduce mortality, salvaging for survival more of the potential life expectancy of humans and toward a primary focus on reducing morbidity, salvaging a healthier rather than a longer life expectancy. For this reason

especially the health experience of the countries that are now developed, and which had made some substantial progress toward economic development before their health transition began, has been an important source of insight about how mortality might be reduced in countries that entered the health transition later. If Western European experience furnishes information about means and measures of mortality control, it is also an important source of information about why morbidity has increased and therefore about what may be done to influence that trend. The point, of course, is not to find the most efficacious path to high morbidity, but to find in the experience of developed countries paths that short-circuit the inverse trends of mortality and morbidity.

By describing the problem as one that is chiefly or exclusively cultural, Johansson implies that the strategies for addressing the problem of too much sickness should be chiefly or exclusively cultural. The problems she addresses certainly exist. They help account for additions to the sum of sickness that are both real and perceived, that are serious and frivolous. As Johansson recognizes, most of the cultural factors she cites as having contributed to rising sickness rates can be regarded as achievements. It is, for example, a good thing that more people are able to recognize diseases that, in earlier days, they might not have been able to identify. But identification should not be confused with existence. Judging from the assurance with which their clergymen diagnosed smallpox as a cause of death in eighteenth-century Sweden, where most localities lacked trained medical practitioners, it appears that lay people readily recognized that disease. They did not recognize tuberculosis so readily and they had a difficult time identifying many diseases whose external signs are weak or which create no signs on the outside of the body, such as heart disease and many neoplasms. Nevertheless, they suffered tuberculosis and it is plausible to suppose that they also suffered heart disease and cancer. It is true that cultural changes concurrent with the health transition have added to the list of diseases that people report and to the security of the identifications now made. But it does not follow that people did not suffer from a disease before they or their physicians could name it.

### **Inflation and deflation**

The list of diseases that physicians and people identify has grown longer. Some diseases that seem to have been long familiar have been identified; many more have been subdivided because enough was learned about their characteristics to recognize separate strains, such as in influenza; and a few new diseases have appeared to afflict humankind. Some of these newly identified diseases are associated chiefly with biological characteristics, such as micro-organisms, and others with emotions.

It is also true that there has been a deflation of sickness. Diseases familiar before the mortality decline have become unfamiliar and a few, such as smallpox, have disappeared. That is a larger matter than may appear at first glance. Smallpox was the leading cause of death in eighteenth-century Europe, accounting for ten per cent or more of all deaths. Virtually the entire population was exposed to smallpox and most people – and in cities nearly everyone – contracted the disease. Across much of the nineteenth century tuberculosis was equally widespread as an infection and a cause of death. To withdraw merely these two diseases from the profile of maladies that people suffer in the late twentieth century – taking smallpox completely out and counting tuberculosis only by the small proportions of people who have died in recent years or have tested positive – is to withdraw two diseases that in their heyday were nearly universal in occurrence. If maladies as commonplace as these have disappeared, or at least become uncommon, then how can sickness rates have increased?

### **How sickness rates have increased**

Recent research has explored an array of forces that have contributed to rising sickness. The mortality decline itself changed the composition of the population (Alter and Riley 1989). Mortality has often been a weaker force than fertility in shaping the population's age structure, but in recent decades

declining mortality among the old has far outstripped fertility as a force behind population ageing. In the aggregate, today's population demands more health services and spends more time in sickness because it is older and thus at a stage of life where the risk of sickness is much higher and where also the propensity and the ability to pay for health services are greater.

The point of particular importance is that, and at specific ages, the sickness rate has increased as mortality has declined. The average age-specific duration of sickness episodes among insured males in Britain increased between the 1870s and the 1890s, a period in which the mortality of adults fell sharply (Riley 1987). In the era of health surveys in Japan, the United States and Britain, sickness prevalence has increased as mortality has decreased (Riley 1990b). An inverse association has obtained also in Hungary, the one country where disease prevalence has been surveyed across a period of increasing mortality. Health surveys undertaken there in 1981 and 1986 show that the prevalence of chronic sickness among adults decreased while mortality increased (Riley 1991). If the addition of new survivors augmented sickness rates in populations experiencing a mortality decline (Verbrugge 1984), Hungarian experience suggests that rising mortality promotes a withdrawal of non-survivors whose earlier death contributes to lower sickness rates.

A striking feature of comparisons of mortality and morbidity trends is that each one so far reported shows an inverse association: when mortality falls, morbidity rises and *vice versa*. Another striking feature is that this inverse association occurs in comparisons across short and long periods. It is plausible to suppose that cultural forces might have changed in Britain between the early 1870s and the mid-1890s or in Japan between the early 1950s and the mid-1980s, between which mortality decreased and morbidity increased. Hence it is plausible, even without specifying which cultural forces operated, and how they operated, to suppose that changes in behaviour and belief pushed morbidity upward. But the surprising finding is that inverse shifts in mortality and morbidity as large in proportion have occurred across short periods in these two societies and elsewhere. What are the grounds for supposing that cultural forces may have promoted a substantial decrease in sickness in Hungary between 1981 and 1986, or a substantial increase in the United States and Britain during the 1970s? What specific beliefs and behaviours changed so substantially in those brief periods? Johansson calls attention to imposing changes in behaviours and beliefs that occurred over decades, even centuries. But the increase of sickness has not required the passage of decades or centuries.

Another non-cultural factor that has been shown to have contributed to rising sickness rates is the successful treatment but not the successful resolution of diseases, a phenomenon Gruenberg designated the 'failures of success' (Gruenberg 1977). Many diseases cannot be cured. Some of these will eventually result in death. Others will not, but the people suffering them will die from other causes before they recover from the original malady. In treating the chronic diseases of adulthood, Gruenberg points out, modern medicine has been especially adept at finding ways to push back the point at which death occurs. This has added substantially to the sum of sickness time because sickness episodes have been prolonged so much. The same force has added yet more substantially to spending on medical services.

Many of the sicknesses associated with rising morbidity do not threaten survival. One recent estimate suggests that, even in disabling diseases, only 36 to 41 per cent of sickness occurs because of diseases that may cause death (Chapman, LaPlante and Wilenski 1986). Most disabling diseases do not cause death. Since it is disabling, this sickness is serious. In terms of individual and social costs, this kind of sickness is the modern equivalent of the often fatal diseases of the high-mortality regime.

To these explanations may be added other economic, social and medical forces that have contributed to rising sickness rates. As the discussion above indicates, it cannot be shown that the ratio of medical practitioners to people increased between the eighteenth and the twentieth century. But it

can be shown that the proportion of their incomes that people spent on health services and health insurance increased, with particular force beginning in the nineteenth century. It can also be shown that the earlier detection of disease has added to the length of time during which people acknowledge that they have a disease, although not necessarily to the length of time for which they suspend ordinary activities. These and other contributors have a cultural component, but that component is not by itself sufficient to explain why sickness has increased.

### **Trends**

It is too early to know what will eventually be learned about sickness trends that will shed light on the power behind each of the biological, medical, economic, social, demographic and cultural forces contributing to sickness increase. But enough is known already to formulate an hypothesis (Riley 1989). In the high-mortality regime that existed before the health transition began, sickness was commonplace. But it was also mostly brief in duration. Measured by incidence, the pre-transition society lived under a heavy burden of sickness, one that over time would shrink. Measured by sickness duration, however, the burden would grow rather than shrink. It would grow especially in this form: both the incidence and the average duration of sicknesses suffered by infants and children would diminish. People in those age groups have come to be better protected from the risk of falling sick and, especially since the 1930s, they have benefited also from new medications that shorten the course of many childhood maladies. In the developed countries today's children are sick much less often than were their counterparts in the eighteenth century. But the reverse is true of adults. In the pre-transition scheme adults, like children, suffered often from infectious diseases having an acute course, as well as from chronic diseases. Since then their susceptibility to acute diseases has diminished, but their susceptibility to chronic diseases has increased. Most importantly, the sickness time added in the form of more and longer chronic diseases far exceeds the sickness time subtracted in the form of fewer acute diseases. For the individual, and for society, the burden of sickness has increased chiefly because the average duration of sickness episodes has expanded.

That, in itself, is an important insight. Discovering that the average duration of sickness episodes has increased weakens the force of any argument that cultural factors alone are responsible for rising sickness. That is so for the simple reason that what has increased is not the number of experiences that people count as sickness, but the duration of those experiences.

### **Conclusion**

The point of investigating the reasons why sickness rates have risen is to discover how to create a low-morbidity regime. That evidently will require different strategies from those followed to reduce mortality. To identify strategies appropriate for reducing sickness requires that the causes of rising sickness be understood. To allocate resources among strategies requires that the relative importance of all of the forces contributing to rising sickness be understood. Johansson wishes to assign to cultural forces a leading or even an exclusive role without first showing how far changes in behaviour and belief go toward explaining rising sickness rates. The evidence now available suggests that sickness rates have risen and are continuing to rise chiefly because of an extension in the average duration of the chronic maladies of adults. Cultural forces may have played a role in prolonging the average duration of sickness episodes, but that role appears to be much smaller than the part played by medical advances, which have deferred death in fatal diseases and in old age, and by the wide-ranging forces that have added new survivors to the population.

## References

- Alter, George and James C. Riley. 1989. Frailty, sickness, and death: models of morbidity and mortality in historical populations. *Population Studies* 43:25–45.
- Chapman, Steven H., Mitchell P. LaPlante, and Gail R. Wilensky. 1986. Life expectancy and health status of the aged. *Social Security Bulletin* 49 (10):24–48.
- Gruenberg, Ernest M. 1977. The failures of success. *Milbank Memorial Fund Quarterly/Health and Society* 55:3–24.
- Johansson, Sheila Ryan. 1991. The health transition: the cultural inflation of morbidity during the decline of mortality. *Health Transition Review* 1:39–68.
- Macfarlane, Alan. ed. 1976. *The Diary of Ralph Josselin, 1616-1683*. London: Oxford University Press.
- Porter, Roy and Dorothy Porter. 1988. *In Sickness and in Health: The British Experience 1650-1850*. London: Fourth Estate.
- Ramsey, Matthew. 1988. *Professional and Popular Medicine in France, 1770–1830*. Cambridge: Cambridge University Press.
- Riley, James C. 1987. Ill health during the English mortality decline: the Friendly Societies' experience. *Bulletin of the History of Medicine* 61:563–588.
- Riley, James C. 1989. *Sickness, Recovery and Death: A History and Forecast of Ill Health*. Iowa City: University of Iowa Press.
- Riley, James C. 1990a. Long-term morbidity and mortality trends: inverse health transitions. Pp.165–188 in *What We Know about Health Transition: The Cultural, Social and Behavioural Determinants of Health* (2 vols), ed. John Caldwell, Sally Findley, Pat Caldwell, Gigi Santow, Wendy Cosford, Jennifer Braid and Daphne Broers-Freeman. Canberra: Health Transition Centre, The Australian National University.
- Riley, James C. 1990b. The risk of being sick: morbidity trends in four countries. *Population and Development Review* 16:403–432.
- Riley, James C. 1991. The prevalence of chronic diseases during mortality increase: Hungary in the 1980s. *Population Studies* 45:489–496.
- Verbrugge, Lois M. 1984. Longer life but worsening health? Trends in health and mortality of middle-aged and older persons. *Milbank Memorial Fund Quarterly/Health and Society* 62:475–519.

## Measuring the cultural inflation of morbidity during the decline in mortality

**S. Ryan Johansson**

*Department of History, Stanford University, Stanford, California 94305-2024, USA*

### Barriers to measurement: the standardization of meaning as a pre-condition for measurement

James C. Riley takes issue with the generalization that the inflation of morbidity during the decline of mortality was primarily a cultural phenomenon. As Riley points out, most recent research, including his own, has been 'directed toward identifying non-cultural forces and assessing their effect'. He correctly raises issues related to measurement, and to the possibility of estimating to what extent non-cultural changes, largely of a biological nature, are responsible for the rise in reported morbidity levels which appears to accompany the decline of mortality during development.

But in pointing out my apparent failure to measure morbidity, Riley seems to have overlooked the major concern of my essay, the conceptual preconditions that make meaningful measurement possible. 'Morbidity', or 'sickness' as it is currently used in the literature is too vague a term to permit the meaningful measurement of historical trends. Apparently, I should have made it clearer that in discussing the concept 'morbidity', and at such length, I was making a point often made in the philosophy of science: standardizing a definition for research purposes constitutes a conceptual problem which must be solved within a research community before scientifically meaningful measurement can proceed, and experts can begin to converge on the most likely explanation for a trend.

My essay was an attempt to explain how historians or health scientists could move towards meaningful measurement by reducing the natural vagueness which attaches to the concept 'morbidity' (or sickness) by explicitly addressing its multi-dimensional complexity, and separately considering how cultural and biological influences impinge on each dimension over time. From my perspective, Riley wishes to continue treating the history of morbidity trends as if that history were no more conceptually problematic than the history of mortality trends or heights. But since being sick or healthy is not as biologically clear-cut as being dead or alive, or being five feet eleven inches rather than six feet tall, morbidity history is inherently conceptually confusing in a way that mortality history and height history are not.

Without reaching some agreement on how to convert vague vernacular concepts into technical terms, historians of health will continue to talk past one another, as James Riley and I seem to be doing. Therefore I am grateful to Riley for raising questions that once again take us back to the most fundamental issue in health history: why it is so difficult to define morbidity or sickness in a way that makes it simple to measure, or to interpret quantitative data, in either the present or the past.

Judging by Riley's reply, I did not get my main points across. I will therefore try to make them in another way, and then conclude by demonstrating how cultural influences on morbidity data could be quantified by treating them demographically, as if they involved various forms of mental migration which could be measured historically.

### **Measuring trends in height, mortality and morbidity**

Consider the following three generalizations about the historical experience of Western Europe, and its offshoots: (1) the level of mortality fell as development proceeded; (2) the mean height of mature males has risen over the last two hundred years; (3) the level of morbidity/sickness has risen (or fallen) in the last two hundred years during the course of economic development.

The first two are informative but vague generalizations about trends over time; well established conventions exist for converting these implicitly quantitative generalizations into a more precise form for analytical or historical purposes. In the case of the third generalization, about a possible trend in morbidity levels, it is much more difficult to go from a level of colloquial vagueness to a precise form permitting quantification, because 'morbidity' is a complex, multi-dimensional concept, that has several possible, quite different meanings. No metric encompasses all of these meanings.

When historians do height history, for example, it is relatively easy for them to establish whether or not mean heights of various populations have risen or fallen over time, in one country or in several, because the ordinary-language meaning of 'height' is already one-dimensional.<sup>1</sup> Since 'height' has only one dimension, it is immaterial whether we use inches or centimetres to measure it, so long as both are standardized units which can be converted into one another. Moreover, contemporary historians

---

<sup>1</sup> Floud, Wachter and Gregory (1990) discuss the problems of measuring height using the various available sources. Most of the problems are technical in nature and have little or nothing to do with cultural influences.

share the same meaning of height and use the same rulers as the people who actually gathered and reported height data in the past. Thus the height data gathered by past army recruiters in one country can be analysed by present social scientists who live in another. Height historians do not need to figure out which of several possible physiological characteristics the army recruiters in country X were trying to measure when they recorded data on the height of potential recruits. Historians do not need to cope with the fact that centimetres and inches may have been of different lengths in different countries, nor do they need to consider the extent to which the real length of an inch stayed the same over time, even in one country. Moreover recruits were never asked to say how tall they felt, or thought they looked. So long as the recruiters used normal rulers and had no incentive to misrepresent the heights of potential recruits (if they did, then cultural influences on height reporting would have to be considered) measuring trends in the mean height of recruits is a relatively straightforward social-scientific undertaking in which the role of cultural factors can be consigned to the margins of research. (This assumes that the data are published in a precise form, that vague categories like 'short', 'medium' and 'tall' are not used, and that data about recruits are age-specific, and representative of the total population of adult males.)

Similar observations could be made about mortality and its history. Like height, mortality is a one-dimensional concept. Once a system of vital registration has matured sufficiently to encompass most deaths, the cultural barriers to measurement are minimal. Throughout human history the concept 'dead' did not change over time, and virtually everyone everywhere could tell the difference between a living body and a dead one.<sup>2</sup>

But when historians want to measure trends in morbidity over time, they are forced to start by using a very fuzzy colloquial term which has more than one meaning. These meanings can be conceptualized as the separate dimensions of sickness. Epidemiologists have succeeded in standardizing the measurement of each dimension to a certain extent, but they have not and cannot standardize the concepts involved in each dimension because everything associated with health and sickness is so intrinsically cultural.

Cultural considerations are about the learned forms of human behaviour associated with the biological states that are classified as disease. Cultural considerations determine how much disease or sickness is perceived as illness (particularly in its non-extreme forms), and how much perceived illness is reported. Because we live with others from whom we expect help, we must learn to tell the difference between being sick and being well in a culturally acceptable manner which is variable over time and space. Learning to identify sickness is also linked to community-level (or government-level) incentive systems. Various institutions embody those incentives through the distribution of rewards and punishments related to sickness. Incentives are culturally constructed from the specific standpoint of discouraging or rewarding the decision to perceive oneself as sick, to report oneself as sick, or to report that someone else is sick and needs treatment. Therefore reported morbidity, which is what we observe directly in all the morbidity data we have, and which is not based on something like a physician's examination or a laboratory report, has to reflect, first and foremost, how people were educated to identify and report disease, and the extent to which they are responding to the incentives associated with sickness. The reporting process is linked to the biology of disease, but not in a way that is as clear as being dead and being reported as dead, or being tall and being measured as having a height of six feet.

---

<sup>2</sup> As noted in Johansson (1991) the progress of modern medical technology has begun to blur the formerly crisp boundary between life and death, and thus to make the definition of death into a problem which requires a culturally negotiated solution, since not everyone agrees that brain death should be preferred to other forms of biological failure.

The weak links between biology and behaviour in morbidity reporting force health historians to become cultural historians to an extent that height and mortality historians need not be. I apparently did not convey this message clearly in my previous essay. Nevertheless, the obvious point remains: cultural considerations are fundamental to the production of the morbidity data that health historians use. In a very real sense they take precedence over biology in the interpretation of data. There is simply no way to explain observed trends in sickness without first understanding the reporting system that produces morbidity data and how it reflects the way people and doctors are trained to perceive, detect and report sickness, and the differential incentive systems (social and economic) that shape reporting behaviour over time at the individual or institutional level.

Therefore, when an historian talks about sickness as if it were a biologically obvious state (like being dead) and proceeds to assume that some form of morbidity data can be used to measure biological sickness (as if measuring the biological extent of sickness over time were just like measuring height) and then attributes an observed trend to biological factors, he or she cannot be doing health history in a meaningful social-science sense. Since all the available quantitative data on sickness are defective in some way and come in a number of inconsistent and incomplete forms that chronically confuse the biology of disease with the sociology of illness, the facts, even the quantitative ones, can be made to say anything the historian wants them to say, once he or she has fallen into the habit of treating morbidity as a simple, one-dimensional concept (like mortality or height) and the associated data as a direct reflection of the biology of disease.

The social-scientific history of health begins with confronting the problems created by the cultural complexity built into concepts like 'health', 'morbidity', 'sickness' and 'illness'. These problems precede meaningful measurement or the interpretation of quantitative data.

### **Barriers to measurement: measuring the separate dimensions of morbidity during the decline of mortality, given that the available data are defective or incomplete**

Epidemiologists have long recognized that the study of diseases, as biological phenomena, is complicated by individual and institutionalized influences on the identification and reporting of disease. Thus it is common to distinguish the study of the epidemiology of sickness, as the biology of disease, from the sociological study of sickness as illness, or perceived and reported disease. Nevertheless, even when it seems legitimate to assume that some type of morbidity data is accurately reflecting real biological changes, quantifying those biological changes remains a complex demographic undertaking.

Approaching sickness and disease as purely biological phenomena which can be measured demographically begins with understanding that four separate dimensions must be measured, namely, frequency, duration, severity and depth (or co-morbidity). The demographic biology of disease is concerned with trying to measure (1) how frequently individuals get sick; (2) how long sick individuals stay sick; (3) how sick they are; and (4) how many identifiable diseases they are sick with at one time, always assuming we can hold the cultural meaning of sick constant while using a particular data set.

By extension, if morbidity history is to be approached as the history of disease and not just the history of reported illness, the historian has to reconstruct simultaneously its four separate dimensions, and thus produce four separate grand trends. Epidemiologists have devised standard measures for each of the dimensions of disease. Frequency is usually measured with cross-sectional incidence rates in some form; duration with measurement of person-days spent in the sick state during some time period; severity with time-specific case-fatality rates; and depth with cross-sectional co-morbidity rates.

At the beginning of the mortality transition no one anticipated that a multidimensional morbidity transition would accompany declining death rates, and no country therefore institutionalized morbidity

reporting to the extent that mortality registration was institutionalized.<sup>3</sup> Morbidity data have been gathered fitfully for various biological and cultural purposes, by various institutions none of which coordinated their efforts.<sup>4</sup> Thus we have no historical equivalents in the history of morbidity that are as good as our estimates for death rates, or as continuous, nor do we have a continual series of physician-generated assessments of the health status of a European population during the health transition based on detailed examinations supplemented with laboratory reports. Most extant data concern reported illness, and roughly approximate the biology of disease in a certain time and place. Morbidity data must therefore be approached by historians as the product of the cultural forces that affect reporting behaviour. The data might also tell us about the underlying biology of disease, if we carefully consider the four separate dimensions.

All historians of health can do is to use whatever defective or incomplete data happen to have been generated, for whatever purposes, in order to estimate the most likely historical trends in each of the separate dimensions of morbidity. Inevitably, various assumptions about missing data must be made. Riley's critique of my essay presents as disconcerting or problematic that, in discussing the history of morbidity, I made such assumptions. The only difference between me and other historians of health is that I made explicit assumptions. When historians do not fully specify their own assumptions, they may conceal from themselves, and others, their naiveté about the defective data that history forces them to use.

Because reported morbidity data do not primarily represent the changing biology of disease, both of the following historical generalizations are true: (1) morbidity levels have risen over the past two hundred years; (2) morbidity levels have fallen over the past two hundred years.

The first generalization will be true if by 'morbidity' the generalizing historian is referring to data which pertain to the frequency and duration of all *reported* forms of morbidity. The second generalization will be true if the historian is implicitly referring to the diminishing extensiveness of the very *severe* infectious diseases, exposure to which significantly raised the probability of dying over a short time period.

Obviously, a rise in the frequency with which non-severe forms of illness have been reported can produce an extremely rapid inflation of morbidity even as death rates continue to decline, or while they remain stable. Such inflation has been observed in all or most of the developing countries in the past few decades. But earlier, as mortality declined, there was clearly a parallel decline in the severity of the most frequently reported diseases. It is because exposure to the most severe infectious disease was reduced that mortality decline occurred.

Changes in the level of co-morbidity, particularly concerning severe diseases, have not been measured at all during the mortality transition. Historians can do little but make assumptions about this dimension of sickness history since even now we know so little about the extent of co-morbidity in modern societies, or how to standardize its measurement. For example, I assumed that as the rapidly lethal infectious diseases were brought under control, the extent to which the average person was simultaneously assaulted by two or more severe diseases also declined. Thus the underlying extent of severe co-morbidity must have declined. The desire to do more than make assumptions about the

---

<sup>3</sup> Mercer (1990) reviews the evolving attempts made by various European countries to gather more, and more complete data for specific infectious diseases in the course of the eighteenth and nineteenth centuries. Mercer is careful to distinguish real trends in disease from reported trends based on changes in the completeness of reporting.

<sup>4</sup> Even now, cross-national coordination remains only an imagined possibility, and cross-national comparisons using morbidity data are hazardous.

history of European co-morbidity set me to using data from other places that still have relatively high levels of mortality. At present such societies are non-Western.

Although Riley objected to my using material from African societies, what is biologically true of African bodies when mortality remains high is likely to have been biologically true of European bodies when mortality used to be high. This does not mean that the same diseases would have ravaged both sets of bodies, or that both populations would have had the same genetic adaptations to each specific disease. What is similar between the two sets of bodies is that when mortality is high, the ordinary person is likely to have more than one form of disease at once, especially more than one form of relatively severe disease. That is the most obvious, if unobserved, biological foundation for observed high mortality wherever it exists.

As discussed in my earlier article, ordinary people living in high-mortality regimes who have been examined by modern physicians do suffer from high levels of co-morbidity including relatively severe diseases which would be considered disabling in a modern society. Moreover, anthropologists who observe the same non-Western populations have also noted that people who would be sick by modern standards continue to work, or otherwise live normal lives, even though, if they lived in a modern Western culture they would disrupt their ordinary routines in response to such conditions. We have no biological reason to assume that the same was not true of peasant populations, or even high-mortality upper-class populations, in Europe at the beginning of the mortality transition.

It was interesting to note that, having criticized my use of biological data about human bodies from other places to reconstruct the probable history of co-morbidity levels in Europe, Riley used some qualitative 'data' produced by privileged Europeans in the eighteenth century to generalize about European illness-reporting behaviour as found in diaries and journals. Eighteenth-century literate Europeans who kept journals and diaries were members of high-income, upper-class or middle-class families. Because they already enjoyed incomes and lifestyles that placed few constraints on the translation of pain or discomfort into socially supported forms of illness, they were as free to complain about their health or disrupt their routines as modern people who live in contemporary high-income countries. Sociologically, exceptionally privileged eighteenth century Europeans had more in common with ordinary modern Europeans than they did with the poor struggling peasants and landless workers who constituted 80 to 90 per cent of Europeans alive in the eighteenth century.

To use class-specific specific qualitative data, produced by literate, middle or upper class eighteenth-century Europeans, in order to generalize about illness behaviour among all Europeans, is to make an assumption so ahistorical as to deny the past and strip the European health transition of its transitional character. In effect, Riley is using a cultural source to try to hold culture constant, a move that implicitly takes the history out of human behaviour. If we could hold culture constant during the health transition, then it would be legitimate to interpret morbidity data as if it described changes in the biology of disease. I see this assumption as pervading Riley's work on the history of morbidity, whereas I assume that both culture and biology co-evolved as the health transition unfolded in Europe.

If anything is obvious or factually certain about the long health transition that occurred in Europe and its offshoots it is that the meanings, norms, conventional behaviour patterns, and policies associated with illness and reporting illness have changed dramatically over time, and that these changes are intrinsic to studying quantitative data. This salient fact should prevent historians of health from equating real sickness with reported illness during the health transition, or from writing as if cultural changes could be marginalized. The co-evolution of culture and biology means that the various measures devised by epidemiologists to study disease involve measuring rates with a rubber ruler, at least whenever we deal with time periods longer than a decade in the twentieth century, and several decades in the nineteenth century when scientific progress was not so rapid.

To study the co-evolution of both culture and biology during Europe's health transition, historians must be more than narrow empiricists who are prepared to observe and measure only what can be observed or measured using the limited range of data for the time or place in which they specialize. As John Caldwell (1990) has been trying to demonstrate, the health transition is a world-wide process which has certain similarities, despite its country-level and regional differences. Everywhere, fundamental cultural problems must be solved before the health transition succeeds in eradicating or reducing the severe infectious diseases which are so closely linked to high levels of mortality. These cultural problems involve changing how people are trained to conceptualize disease, and what kind of health practitioner they prefer to consult.

Modern health scientists, who think of science as quantitative, everywhere resist acknowledging the important role that such cultural changes have in actually changing the biological extent of disease, just as they resist the idea that cultural influences are fundamental to the production of morbidity data. Epidemiologists are happy when they tackle complex measurement issues related to the various biological dimensions of disease, because these dimensions are, in principle, measurable. But changes of a sociological nature, based on socially constructed meanings, norms, conventionally acceptable behaviour patterns that differ by age, sex, social class, ethnic group, religious group, and government policies, look unmeasurable, and hence are shunned.

So, as a sort of last try to make demographers, epidemiologists or any quantitatively oriented health historian more comfortable with the importance of cultural influences on morbidity history, let me briefly explore the extent to which cultural changes during the health transition in Europe can be approached demographically, and measured, as if they were about mental migration patterns from one sector of a population to another. From this demographic perspective, the health transition concerns the individual's relocation from a perceived state of health into a perceived state of illness, given that there are culturally constructed rules that control the issuing of visas in the form of the rules governing the allocation of health resources. These visas, which can be measured in various ways using statistics about the economics of health, permit the outmigrant to leave the free and productive state of health and to enter the costly state of illness for a given time period at the expense of others.

### **Overcoming barriers to measurement: treating the culturally driven inflation of morbidity as a form of mental migration**

At least two centuries ago, at the beginning of the health and mortality transition in modern Europe, almost everyone (let us estimate that as meaning 85 to 95 per cent of the total population) lived in the sector of the population that did not report being sick to anyone prepared to record that report and link it to other comparable reports that could be counted, summed and stored. When ordinary people felt sick they complained about it to their friends, families or diaries (if they could write and afford to buy a diary, which most people could not) or they took to their beds and medicated themselves, or they simply carried on in various degrees of pain and with impaired levels of function. If they could not manage to perform their usual routines at all they consulted a European folk practitioner who did what he or she could. Nothing in the informal institutions associated with European folk medicine encouraged keeping morbidity data in written form. So much the worse for the health historian who wants to quantify the real level of disease at the beginning of the health transition! Most real disease went unmeasured and most illness behaviour went unreported.

Even in the eighteenth century, at least 5 to 15 per cent of the European population came into regular contact with doctors who already considered themselves scientifically trained. (Whether or not they were by modern standards is another story.) Most of the physicians already had a professional commitment to reducing the real level of disease in their societies, and Riley (1987) has written very

informatively on that subject. Towards that end they began to gather and report some morbidity data for their communities, the hospitals they were connected with, various private insurance companies or employers, or various governments who were trying to improve the public's health; but such data encompassed only a fraction of the total population. Gradually, physicians began actively to sponsor a form of mental migration from the folk-medicine sector of the population, where traditional theories of disease still governed the explanation of causation, and the use of traditional folk remedies still dominated treatment. This migration took less educated, uneducated and illiterate Europeans out of the traditional mental sector, in which disease was a punishment from God or the result of witchcraft, into the modern mental sector of the population in which disease had scientific causes and remedies.

If we had data on this mental migration from one sector to another we could observe a gradual, but not necessarily regular, rise in the proportion of Europeans who mentally resided in the medically modernized sector of the population, and who were thus at an increasing risk of having their illness behaviour reported or recorded in some way.

After two hundred years, at least 95 per cent of the population had been transferred to this mentally modernized sector. Among other things, this sector continues to be responsible for producing morbidity data, and continues to produce more and more of it. Holding biological factors constant, this culturally-sponsored mental migration would have greatly inflated the incidence of reported morbidity given that the entire population was already being enumerated. This inflation of reported morbidity is demographically similar to urbanization, which was a physical transfer of the majority of the population from one geographic sector to another. Medically modernizing a population which is already completely enumerated will inflate the apparent extent of reported morbidity.

But of course we cannot hold everything else equal during the mental modernization of the European population. As the health transition proceeded, the scientifically trained doctors who populated the modern health sector became more and more scientifically sophisticated. They were constantly retrained to detect and report new diseases, and institutionally required to engage in more and more diverse forms of illness reporting. If we could count the number of named diseases that scientifically trained doctors were prepared to report at the beginning of the health transition, and prepared to teach their patients about, I would guess that at the very least this list of named diseases doubled in size over two hundred years.

This increase in the population of named diseases contributes to both the incidence-related inflation of reported morbidity, and the decline in the severity of the most frequently reported diseases. For the most part the diseases formerly reported were quick killers, that is they had a high level of severity, that took the lives of many people. The newly discovered diseases afflicted fewer people, and killed them more slowly. The last development contributes to the biological rise of any measure of prevalence, but, once again, this rise is also a function of better knowledge (most of the new diseases existed previously but were undiscovered), more diligent diagnosis and society's willingness to pay for the treatment of extended forms of disease. But this increasing population of reported diseases was not such as to interfere with the fact that a small but select list of reported diseases was being brought under public control. Declining incidence rates for a few severe infectious diseases were linked to the modern decline of death rates. The slow addition of new, less severe diseases to the list of reportable disease kept up the pressure for a gradual inflation of reported morbidity, while gradually divorcing that rise from trends in mortality.

Perhaps the most profound change which took place in the modern medical sector was the increasing extent to which its ordinary residents were institutionally encouraged, indeed invited, to migrate from well to sick, and generously rewarded for staying that way for longer and longer periods. Imagine going back to the beginning of the health transition and calculating the economic cost of each

reported illness (as the amount of money it cost to pay a doctor, keep someone in a hospital, or pay for time off work and so on). Because income was so much lower than now societies could not afford to sponsor much migration to the state of illness, and the scarcity of resources strictly limited the number of visas that permitted lengthy stays in the state of sickness. Even people who had mentally migrated to the medically modernized sector of the population had to limit their costly visits to the doctor, hospital stays, or their expensive time off work. By so doing they limited the amount of reported morbidity and the apparent duration of sickness associated with each individually reported incidence of sickness, however sick they really felt and for however long.

One of the biggest changes in human welfare over the past century has been the increasing propensity of developed countries to allocate resources to the support of those afflicted with either infectious or chronic diseases. Society no longer limits care to the terminal phases of any disease but increasingly tries to cure or care for the afflicted person from first diagnosis to recovery or death, while detecting disease in earlier and earlier phases. Few controls on the length of care or the cost of trying to cure disease have ever been institutionalized.

Doctors, who benefit economically from society's generosity to the sick, have no professional incentives to control the cost of paying for curing or caring, and thus no incentives to limit the amount of routine migration from wellness to sickness, or the number of visas they issue that permit people to reside in the sick state for long periods. As long as the payments for being sick approach the ordinary wage, people who do not feel in perfect health can use sick leave to leave the state of health whenever they feel mildly under the weather. To people who feel well, generous sick-leave policies constitute a free visa for vacations in the reported state of illness.

Even when we are dealing with the retired population, the generosity of society encourages more frequent and longer departures from the state of wellness. Ageing has always been associated biologically with increasing sickness and disability. In former times society simply refused to pay for this real, age-related increase in sickness and disability and older people were forced to tolerate increasing levels of pain, and their families or villages to cope with caring for them as best they could. Home care went unreported and the cost of it was privately borne. But who would doubt that there was once a huge reservoir of socially unreported, unsupported disease among the aged?

Society's recent and commendable generosity with respect to the welfare of elderly people gradually drains this reservoir of unreported sickness, and sponsors the transformation of unreported forms of real sickness, into reported forms of illness. Even if real sickness were declining among the aged, especially among younger old people, reported levels of illness would rise as society issued the economic visas which permitted the old to migrate to the state of illness and reside there.

The cost of paying for morbidity among the elderly is certainly influenced by the increasing percentage of the population over age 65, but if only demographic changes were at work the costs of paying for disease among the elderly would rise only as fast as their representation in the total population. Everyone knows that the cost of paying for medical care for the elderly is soaring much faster than would be demographically justified if everything else stayed constant. Determining the extent to which this rise is biologically real or sociologically induced is not productive. The biological problems of the elderly have always been real. The main question is, to what extent was sickness among the elderly underreported, even two decades ago? After that question is answered we can begin to estimate to what extent the elderly may or may not be more biologically frail than their counterparts at an earlier date. My guess is that the cultural propensity to support frailty is probably dominating trends, not an increase in real biological frailty. But this is a hypothesis. It cannot be proved or disproved by assuming that morbidity data are a straightforward reflection of the biology of disease among the elderly.

So long as little or no institutionalized controls are imposed on the collection of generous benefits for not being in perfect health, people of all ages will readily migrate to the expensive state of being sick, and stay there as long as it is comfortable, or advantageous, to do so. Limiting this process of migration means culturally negotiating how to draw an ethically acceptable line between being sick and being well on a health or sickness continuum that, as explained earlier, has no naturally identifiable, biologically obvious, break point. Limiting the migration to the socially expensive state of illness is a problem in applied ethics which a confusion of the biology of disease with the sociology of reported illness will do nothing to clarify.

### **Achieving consensus about the history of morbidity**

Riley asks: is culture everything in sickness? Epidemiologically, sickness is about the biology of disease not the sociology of illness-reporting behaviour. But most morbidity data from both the present and the past continually confound illness and sickness. Our present ability to estimate the extent of real disease, by distinguishing it from non-biological changes in the propensity to identify and report illness, is still imperfect. Distinguishing between the two in the past is even more difficult. If quantifying historians want to disentangle the two, they must first understand the cultural mechanics of illness reporting during the health transition. That was the main point of my previous article. Such an understanding is prior to meaningful measurement.

Moreover to quantify the health transition in Europe the historian must use both biological theory and data from comparable cultures that are in the early stages of the worldwide health transition. Our historical data cannot tell us much about the biology of disease, with the possible exception of those data series involving the accurately recorded number of cases of severe infectious diseases which attracted so much attention during the mortality transition.<sup>5</sup>

Although measurement-minded morbidity historians cannot escape from the frustrations of having to deal with rubber rulers, their efforts to measure the biological aspects of disease will be most successful if it is assumed from the start that cultural forces are responsible for most forms of rising levels of reported morbidity during the health transition. If death rates are declining as morbidity rises, some form of culturally influenced reporting is most likely to be inflating morbidity. Only when both morbidity and mortality rise together, and to approximately the same extent, is it safe to assume that the biology of disease is driving the reported rise of morbidity, and that changes in the mechanics of reporting can be analytically marginalized.

In those developing countries currently in the middle of the health transition, issues pertaining to the measurement of morbidity are particularly pressing since they are so closely tied to health policy. Today, the health transition can proceed so rapidly that nothing can be held constant from year to year – neither the proportion of people in the medically modernized sector of the population, nor the diseases which most doctors are prepared to report, nor the kinds of sickness that institutions are prepared to support. In the midst of so much change there is one rule of thumb. If mortality levels are falling as development proceeds, but reported morbidity levels are rising, check whether the reported diseases are increasingly less severe, that is, that they have a lower probability of causing death in a short period of time than reported diseases used to do. If so, the historical experience of Europe is probably being reenacted, and the cultural inflation of morbidity has begun. This is a good thing from a welfare perspective. If, as appears in some Central European countries, both disease-specific incidence rates

---

<sup>5</sup> Szreter (1988) considers the extent to which the probable overreporting of tuberculosis as a cause of death in the middle of the nineteenth century accounts for much of its apparent decline in the late nineteenth century. Most health historians have assumed that the observed decline was real, and not the result of reporting practices related to the changing level of expertise of those reporting causes of death.

and mortality rates are rising together, then some form of biological deterioration is likely to be at work, especially if the apparent severity of the most frequently reported diseases is rising. The historical experience of Europe, properly conceptualized, can illuminate present-day illness-reporting behaviour, just as the modern history of developing countries can be used to illuminate Europe's dim biological past.

### References

- Caldwell, John C. 1990. Introductory thoughts on the health transition. Pp.xi–xiii in *What We Know about Health Transition: The Cultural, Social and Behavioural Determinants of Health*, ed. John Caldwell, Sally Findley, Pat Caldwell, Gigi Santow, Wendy Cosford, Jennifer Braid and Daphne Broers-Freeman. Canberra: Health Transition Centre, The Australian National University.
- Floud, R., K. Wachter and A. Gregory. 1990. *Height Health and History*. Cambridge: Cambridge University Press.
- Johansson, S.R. 1991. The health transition: the cultural inflation of morbidity during the decline of mortality. *Health Transition Review*. 1:39–68.
- Mercer, A. 1990. *Disease Mortality and Population*. Leicester: Leicester University Press.
- Riley, James C. 1987. *The 18th-Century Campaign to Avoid Disease*. London: Macmillan.
- Szreter, S. 1988. The importance of social intervention in Britain's mortality decline c. 1850–1914. *Social History of Medicine* 1:1–38.