

Measuring the burden of disease

John Broome

1. *Introduction*

A important part of the WHO's work has been to develop and publish summary measures of the burden of disease for different nations. It aims to measure both the overall burden that disease and ill-health impose on a nation, and the burden imposed by individual diseases. Its measures are intended to provide general indicators of how a nation's health is progressing, and also to guide the setting of priorities in health care. Much of the thinking that underlies the WHO's measures is set out in Christopher Murray's article 'Rethinking DALYs'. To a large extent, this present paper is a response to some of the arguments contained in Murray's important article.

One claim I shall make is that a measure of the burden of disease should not strictly be identified with a measure of health. We should concentrate on the burden of disease, and no genuine measure of health can be found. However, I have no objection to the general idea of measuring health, and it is so well established that it would be futile to reject it at the start. So I shall set out from this general idea. Only when the idea has become more refined and specific will it emerge that the burden of disease is a better notion to concentrate on. This will begin to emerge in section 5.

The idea of a summary measure of health arose originally out of the process of collecting statistics on health. A summary measure is meant to summarize the mass of statistics that describe a nation's health in detail, so as to pull out of them a single indicator of the overall healthiness of the nation. As a result, the development of summary measures has historically been led by the statistics available. Given the statistics, the question has been to find a good way of summarizing them. But I am going to approach the measurement of health from a different direction. I shall start by asking what we should try to measure, unconstrained by

the available statistics. Once we have the answer to that question, we may find we can measure it with our existing statistics, or we may find we need new and different statistics.

We need to settle one thing at the start. Should our summary measure be strictly a measure of the quantity of the nation's health, or alternatively should it measure how good the nation's health is? Should it be a descriptive or an evaluative measure?

This may not be a real choice at all. Health is such an integral part of our wellbeing that we may have no conception of the quantity of health that is separate from its goodness. We can make the distinction with some other sorts of good. For example, we can plausibly measure how much education a person has, and then separately ask how good her education is for her – what benefit she receives from it. We can also make the distinction with various specific components of health. For example, we can measure the length of a person's life in years, and separately ask how good it is for her to live as long as she does. But when we put together all the different components of health to measure a person's overall healthiness, there seems to be nothing for us to measure apart from how good her state of health is for her.¹

At any rate, I shall assume that our task is to measure the goodness of a nation's health. It is therefore a matter of evaluation rather than strictly of measurement. When I speak of health, I shall mean strictly the goodness of health, in so far as this is different from the quantity of health. By a 'measure of health' I shall mean a measure of the goodness of health. Conversely a 'measure of the burden of disease' is a measure of the badness of the burden of disease.

Once we recognize that we are valuing health rather than measuring it, our work takes on a quite different aspect. It no longer appears as a technical matter of manipulating statistics. It needs to be founded on a proper theory of value. We shall need to ask fundamental, philosophical questions. What, exactly, is good about health? How far is health valuable in

itself, and how far as a means towards other good things? How much good does it do someone to extend the length of her life? And so on.

A further reason to aim at measuring the goodness of health is that it is the way to make our measure most immediately practical. An important aim of health policy should be to improve the nation's health. That is to say: to make it better. So a measure of the goodness of health provides an appropriate objective to aim at. It should not be the only objective – we shall see that fairness should be another – but it is perhaps the main one.

This too puts the question of measuring health into a different light. If our measure is to constitute an objective to be aimed at, the question of how we should measure health converges with the question of how we should direct our resources towards improving health. What should be the priorities of the health system? What principles should guide the rationing of our limited resources? And so on.

2. *Comparing distributions of wellbeing*

The measurement we need to do is fundamentally a matter of making comparisons. For practical decision making, we need to compare together the various options that are open, to see which is the best. We may also have other purposes for a measure of health, but these too are matters of comparison. We may want to see how a nation's health is progressing over time, or how one nation is doing compared with its neighbours. Then we need to compare a nation's health at one time with its health at another time, or health in one nation with health in another. Our comparisons may need to be matters of degree: when comparing two alternatives, we may need to know both which is better than the other, and also how much better it is.

It may seem that we are also interested in the goodness of a nation's health in a sort of absolute sense: how much harm is the nation suffering because of ill-health. This is the

burden of disease born by the nation. But it too is a matter of comparison. To calculate the burden of disease, we need to compare the nation's health with an ideal, healthy state, to see how much worse it is.

Given that our problem of measurement is one of comparison, I can illustrate it in a schematic diagram. I have chosen to illustrate the effects of a single epidemic as an example, but other comparisons could be put into a similar diagram. To give the example a practical aspect, imagine the epidemic is expected in the future, and that it might be possible to prevent it by some action to improve public health. To help decide what effort should be put into preventing the epidemic, we want to measure how much harm it would do. Conversely – this comes to the same thing – we want to measure the benefit that would be achieved by preventing the epidemic.

Figure 1 has two halves. The right half shows what will happen if the epidemic occurs; the left what will happen if it is prevented. Each half is a multiple graph. Time is measured in a horizontal direction. A vertical dotted line marks the date when the epidemic occurs, if it does occur. Each horizontal dotted line marks out the life of a single person; corresponding lines in the two halves of the diagram belong to the same person. Each line is the horizontal axis of a little graph that shows the course of the person's life. A person's graph begins at the time she is born and ends when she dies. During her life, the height of the graph above the axis represents the person's *wellbeing* – how well her life is going.

I use 'wellbeing' in a comprehensive way. A person's wellbeing includes everything that is good for her, with one exception. I think it is good for a person to be treated fairly, and bad for her to be treated unfairly. But fairness works in such a different way from other goods that it is inconvenient to include it within a person's wellbeing.² So as I use the term in this paper, wellbeing includes everything that is good for a person, apart from fairness.

To speak more exactly, the height of the graph represents the person's *temporal* wellbeing: how well her life is going at each particular time within the life. Temporal wellbeing needs to be distinguished from the person's *lifetime* wellbeing: how well her life goes as a whole.

One line in the right hand half of the diagram has no graph on it; this indicates that this particular person will not live at all if the epidemic occurs.

In drawing the diagram, I have implicitly assumed each person's temporal wellbeing can be measured on a cardinal scale. Indeed, I have assumed more than this: the zero of wellbeing is significant, which means I have assumed a ratio scale. Furthermore, I have assumed that the scale of wellbeing at each time is fully comparable with the scales of that person's wellbeing at other times, and everybody else's wellbeing at any time. All these scales need to be properly developed. I believe that can be done,³ but I shall not try and explain how in this paper. I shall simply take the scale of wellbeing for granted. (The definition of the zero is mentioned in section 6.)

Comment [1]: Reference to these terms

Figure 1 shows two 'distributions' of wellbeing. In each half of the diagram, temporal wellbeing is distributed across people and across time. I shall assume that the distributions shown in the diagram contain everything that is relevant to measuring the harm done by the epidemic. This in itself raises several questions. One is why the diagram shows wellbeing at all, rather than the level of the people's health. After all, we are trying to measure the goodness of health specifically, not wellbeing as a whole. I shall answer this question in the last part of this paper, starting in section 6. The principal purpose of this paper is to argue that we should concern ourselves with the effect of disease on wellbeing as a whole. We should aim to measure the whole loss of wellbeing that is caused by disease. But that is for later.

Another question is whether there might be aspects of the harm done by the epidemic that are left out of the diagram. I am assuming there are not; is that a justifiable assumption? I am assuming that the left-hand half of the diagram contains enough information to determine the

overall good of the nation if there is no epidemic, and the right-hand half its overall good if there is an epidemic. If that is correct, the harm done by the epidemic is the difference between these amounts, and that is how we can measure it. So does each half of the diagram contain enough information? I think it does, but I am not in a position to explain why till section 3. For now, I shall simply assume it. The diagram shows each person's times of birth and death, and her wellbeing at each time she is alive. I shall assume this is enough to determine the nation's overall good. Let us call this the assumption of 'distributed good'.

Given that, we can see our measurement problem as a problem of aggregating wellbeing. (By 'aggregating' I mean putting together all the wellbeing that is distributed across the diagram, to arrive at an overall judgement. We need not necessarily do this by adding up; 'aggregating' is a more general term than 'adding'.) The diagram shows wellbeing distributed across people and across time. We need to aggregate it across people and across time to determine how well off is the population as a whole. To calculate the harm done by the epidemic, we need to aggregate wellbeing in both distributions, and compare the two aggregates.

Figure 1 illustrates some of difficult problems we shall need to contend with in the course of aggregation. Each of the following effects is caused by the epidemic in the diagram. One person is killed at the time of the epidemic. One person is disabled by the epidemic, but her life is not shortened. One person's life is shortened by the epidemic, but she does not die till some time later. One person who would have been born is not born as a result of the epidemic (perhaps because one of her parents is killed). One person born later, who would have been healthy, is born disabled (perhaps because of genetic damage to one of her parents).

Each of these effects is typical of disease, and our measure of health will have to be founded on a theory of value that is able to encompass them all. Each poses a theoretical problem, and some pose very difficult problems indeed. The hardest is to know what to do

about the changes in population that are caused by disease. I am thinking of *timeless* population, by which I mean all the people who live at some time or other. Whenever disease kills an existing person, it reduces by one the *temporal* population – the people who are alive at a single time – but it does not immediately reduce the timeless population. However, by preventing the existence of people who would otherwise have existed, disease does also reduce timeless population. It can do this by killing a person who would later have a child, for example. Conversely, reducing disease increases timeless population. Indeed, from the point of view of history, this is its biggest effect by far. Whereas improvements in health in the last two centuries have, say, doubled a typical lifetime, they have multiplied the world's population, say, tenfold. It is extremely hard to know how to evaluate changes in population, and so far all attempts to measure the benefits of reducing disease have simply ignored them.

I am not going to tackle any of these specific problems in this paper; I am going to concentrate on more general principles of evaluation.

3. *Aggregating wellbeing*

How is wellbeing to be aggregated across a two-dimensional distribution like the ones in figure 1? One idea is to proceed in two stages. First, we might try to arrive at a value for each time separately. For each time, we would take all the people alive at each time, and aggregate the temporal wellbeing each is enjoying at that time. This would give us a value for the distribution of wellbeing at that time. Call it a 'snapshot' value. Then, at the second stage, we might take all the snapshot values we have obtained at the first stage, one for each time, and aggregate all of those together to arrive at the value of the distribution as a whole.

This procedure aggregates wellbeing time by time. It will work only under a special assumption: the snapshot values have to make good sense. It has to be the case that we can properly value the distribution at each time, on the basis only of people's temporal wellbeing

at that time, and independently of what the distribution is like at any other time. Moreover, the snapshot values for each time must together fully determine the overall value of the distribution. This double assumption is technically known as *separability of times*.

I shall explain in section 4 that actually times are not separable, so aggregating time by time is not a correct route to aggregation. To say that times are not separable means we simply cannot evaluate a distribution at individual times. To put it graphically, there is no such thing as the health of a nation at a particular time, when health is understood as the goodness of health.

That does not prevent us from evaluating the harm done by the epidemic. It simply means we have to evaluate it in a different way. We must aggregate wellbeing by a different route. The route I recommend is a different two-stage procedure. Instead of aggregating time by time, we aggregate person by person. We start by evaluating each person's lifetime wellbeing. This is a matter of aggregating each person's temporal wellbeing across all times, to arrive at an overall value for her life. Having done that, we aggregate together the lifetime wellbeings of all the different people.

This alternative procedure relies on the assumption that people are separable, which is symmetric to separability of times. It has to be the case that we can properly evaluate each person's life, to arrive at her lifetime wellbeing, on the basis only of her temporal wellbeing at all times, and independently of how other people's lives go. Moreover, people's lifetime wellbeings must together fully determine the overall good of the nation.

The difference is that, whereas times are not separable, it is reasonable to assume that people are. The inseparability of times comes in section 4, but I shall give some arguments for the separability of people now.

The separability of people breaks down into two smaller assumptions. One is that the good of the nation as a whole is fully determined by the lifetime wellbeing of the people who make

up the nation. The second is that the lifetime wellbeing of each person is fully determined by her temporal wellbeing at all times. I think these are both reasonable assumptions to make.

The first is something I call 'the principle of personal good'. It says the overall good of the nation is determined only by the lifetime wellbeing of the people. One can certainly raise doubts about this assumption. For example, the wellbeing of animals should no doubt also count, as well as the wellbeing of people. But when we are specifically interested in people's health, I think the principle of personal good is so plausible that I do not need to defend it further in this paper.⁴

The second assumption is that each person's lifetime wellbeing depends only on her temporal wellbeing at all times. What does this rule out? It does not rule out 'pattern goods' as I call them.⁵ These are goods that depend on the pattern of a person's temporal wellbeing, as it progresses over time. For example, some people think it is better for a person if her life improves over time rather than declines, even if the total of her temporal wellbeing is the same in either case. If this is so, then an improving pattern of temporal wellbeing is a pattern good. The assumption does not exclude pattern goods, since the pattern of temporal wellbeing, as it develops over life, is determined by the person's wellbeing at all times.

On the other hand, the assumption does rule out anything that is good for a person, but that does not show up either in her temporal wellbeing at some particular time in her life, or in the pattern of temporal wellbeing over her life. A possible example is the value that James Griffin calls 'accomplishment'.⁶ To accomplish something in one's life takes a long time, and perhaps the accomplishment that results cannot properly be assigned to any particular time or times in the life. Nor is it a pattern good. I am not entirely convinced by this example; I think it might be possible to assign the good of accomplishment to particular times. Perhaps it might be assigned to all the time the person is working towards the accomplishment, for

example. But in any case, I think it is reasonable to ignore undated goods like accomplishment when we are measuring the goodness of health.

So I think it is reasonable to assume separability of people, and we can evaluate the good of the nation through person-by-person aggregation. Separability of people also gives us retrospective grounds for the assumption of distributed good that I made in section 2. It tells us that the nation's good can be determined by aggregating across a distribution of wellbeing like the ones in figure 1. It therefore implies that this distribution fully determine's the nation's good.

4. *Inseparability of times*

So we have a means of evaluating a distribution of wellbeing. This is what we need for practical decision making. When deciding on a course of action, we need to evaluate the different distributions that will result from each of the things we might do. We can make our evaluations by the person-by-person method. We cannot evaluate them time-by-time, since times are not separable. But that is no problem for decision making.

However, the inseparability of times does cause a different sort of difficulty for the measurement of health. We would often like to know how healthy the nation's population is at a particular time. For example, in order to see how the nation's health is progressing, we may want to compare its health at one time with its health at an earlier time. Yet if times are not separable, this cannot be done.

Existing summary measures of health routinely come with dates attached. This is because of their history as summaries of health statistics. Statistics are collected for particular dates, and the statistics for a particular date are summarized in a summary measure for that date. However, since times are not separable, no compilation of the statistics for a particular date

can tell us the healthiness of the nation at that date, when healthiness is understood as the goodness of health.

Before this comes to seem too paradoxical, I need to explain why times are not separable. The root of the problem is that disease often shortens people's lives. This is one of its bad effects that are illustrated in figure 1. Since it is bad, it must be accounted for in our measure of the goodness of health. However, it cannot be assigned a date.

Why not? It is tempting to give a quick answer by quoting Epicurus, and I shall succumb to the temptation to the extent of quoting him. Unfortunately, the actual answer is longer.

Epicurus said:

Death, the most terrifying of ills, is nothing to us, since so long as we exist death is not with us; but when death comes, then we do not exist. It does not then concern either the living or the dead, since for the former it is not, and the latter are no more.⁷

Death cannot harm you before you die, because before you die death has not arrived on the scene. And death cannot harm you after you die, because by that time you are not around to be harmed. So there is no time when you suffer harm from your death. That is Epicurus's claim, and in this I think he is right. (He himself seems to have drawn the conclusion that death does you no harm, but in that he is not right; death does you harm of a sort that does not occur at a time.)

Epicurus's point looks like a quick explanation of why times are not separable.

Separability requires all the good and bad aspects of health to be assigned to a time. Since the harm of death cannot be assigned to a time, it seems to make separability impossible. But actually it does not. Epicurus's correct point does not refute separability.

The reason is that separability is concerned with the dating of good and bad, not the dating of benefits and harms. Benefits and harms are *differences* in good and bad. Think first of the overall benefit or harm done you by an event. An event benefits you if it causes you to lead a

better life than you would otherwise have done, and it harms you if it causes you to lead a worse life than you would otherwise have done. So we assess harm or benefit by comparing one possible life with another: the life you do lead with the life you would have led had the event not occurred.

It may be possible to break down the overall harm or benefit into a number of particular harms or benefits. To do this we shall have to set up some sort of correspondence between the good and bad features of one possible life and the good and bad features of the other. For example, suppose you suffer some illness on a holiday, which makes your holiday less fun than it would have been. We could say that a particular harm caused by the illness is that it spoils your holiday. In saying this, we are making the good that occurs in your actual holiday correspond to the good that would have occurred at the same time, had the illness not struck. We can assign a time to this harm that your illness does: the time of your holiday.

But take a different example. Suppose an illness forces you to delay a holiday you had planned. Suppose it also makes the holiday less fun than it would have been, because it now takes place in the rainy season. In this case too, we could say that one particular harm caused by the illness is that it spoils your holiday. In saying this, we are making the goodness of your actual holiday correspond with the goodness of the holiday you would have taken at an earlier time, had the illness not struck. Because our correspondence crosses time, it is not clear in this example how we ought to assign a time to this particular harm. Nevertheless, all the good and bad in your life may be assignable to a time. Your holiday is good to some extent, and this goodness can be assigned to the time of your holiday. It is only the harm that is difficult to date, because it involves an intertemporal comparison.

In other cases, it may not even be possible to set up a sensible correspondence between events in one life and events in another. One example is when an event sets your life on an entirely different course. Another is when an event shortens or lengthens your life. For an

event like this, we shall not be able break down its overall benefit or harm into particular benefits or harms. All the good and bad in your life may be assignable to a time nevertheless.

Take this analogy. Imagine that, just before it was published, a particular book had its final chapter excised. The book was shortened by six thousand words, and the published version emerged 187 pages long. Which pages were the six thousand words removed from? Not from any of the first 187, since those pages contain all the words they would have contained had the book not been shortened. Nor from any pages beyond 187, since there are no such pages. So there is no page from which words were removed. Nevertheless, every word in the book appears on a particular page. Furthermore, if the longer version had been published, every word in that version would also have appeared on a particular page. All words are paged, but excisions of words are not paged.

Similarly all goods in a life may be dated, even though the harm caused by shortening the life is not dated. Epicurus demonstrated that the harm caused by shortening life is not dated. But separability requires that the goods in a life are all dated. For all that Epicurus said, therefore, times might be separable.

Nevertheless, times are not in fact separable. The true reason is more subtle than Epicurus's point. It appears in figure 2, which has the same basic design as figure 1. It shows two alternative distributions of temporal wellbeing. In each, people live for a while at a particular level of wellbeing. For simplicity, I have made their temporal wellbeing constant throughout their lives, and I have made it the same for everyone. In the left-hand distribution people live for twice as long as people live for in the right-hand distribution. But twice as many people live in the right-hand distribution. I mean, the timeless population is twice as big. In fact, at each individual time, both distributions have the same temporal population, namely two people. Furthermore, in both distributions, the people living at each time are equally well off.

If times were separable, we would be able to evaluate both distributions in snapshot fashion, time by time. Pick any particular time, and evaluate the two distributions at that time. At that time, they both appear exactly the same. Each has two people living at that time, and each has them living at the same level of wellbeing. At any particular time, the two distributions are indistinguishable, therefore.

Do not think they can be distinguished at some particular times, namely the moments when a person dies or is born. When someone dies in my example, I assume she is alive at every moment up to and including midnight, and at every moment after midnight she is dead. Conversely, when someone is born, she is alive at every moment after midnight, and at every moment up to and including midnight, she is not yet alive. These assumptions mean that two people are alive at every time, in both distributions.

The two distributions are indistinguishable at each individual time. Therefore, if times were separable, at each time they would be equally as good as each other, and it would follow that they would be equally good overall. But they are not equally good overall. They are the same at each time, taking a snapshot view. But there is an important difference between them nevertheless: people live for twice as long in the left-hand distribution. Longevity certainly has value. So the left-hand distribution is better. Separability of times implies it is not better. Therefore, separability of times is false.

I think there is little doubt that the left-hand distribution is really better. Figure 2 represents in stylized form one of the main benefits we have gained from improvements in people's health: we live longer. If the left-hand distribution were no better than the right, we would have gained no benefit from the historical lengthening of people's lives. Compare two possibilities. In one, a person lives her life for eighty years. In the other, she dies after forty years, but is magically replaced by someone else who lives out the remainder of her life. The

first possibility is better than the second. One eighty-year life is better than two forty-year lives. The comparison between left and right in the diagram is exactly like this.

Longevity is valuable, but we miss its value if we try to value a distribution time by time. Longevity is a value that only appears when we look at many different times together, not when we look at single times individually. That is why times are not separable. The good of longevity cannot be assigned a date. Its opposite, the harm of shortening life, therefore also cannot be assigned a date. Since disease shortens life, we cannot ignore the inseparability of times when we measure the burden of disease.

Figure 2 takes account of people's coming into existence as well as their going out of existence: births as well as deaths. It is this that makes the failing of separability obvious. Measures of health typically take account of deaths, but ignore births. This may partly explain why the inseparability of times has not been noticed by those who measure health. But ignoring births could not actually justify us in assuming separability. Births do in fact occur.

To summarize: longevity is good, and a part of health. But it is a good that has no date. It follows that there is no such thing as the health of a nation at a particular date. It is futile to try and measure it.

5. *Causal dating*

However, we may still be able to attach dates to measures of health, in a different way. Take the epidemic represented in figure 1. It causes some definite amount of harm. Some of this harm is spread across time, and some has no date, because it consists in shortening people's lives. Nevertheless, it is plain when all the bad effects are *caused*. They are caused by the epidemic, at the time of the epidemic.

An event that occurs at one date may cause harm at a different, later date. If radiation escapes from a power station, it may cause people to suffer from cancer much later. An event that occurs at one date may also cause undated harm, for example if it kills people. But even if we cannot date the harm, we may be able to date the cause of the harm.

This gives us a different way of attaching a date to a measure of health. Let us call it 'causal dating'. As a handy piece of terminology, let us use the term 'disease' for the cause and 'ill-health' for the effect that disease causes. (Not all the bad effects of disease consist in ill-health – a point that will become important in section 7.) We can assign to the year 2000, say, all the ill-health that results from disease that occurs in 2000. Much of this ill-health will not itself occur in 2000, but later, and some will not occur at any date at all.

From section 6 onwards, I shall be arguing that we need to measure the burden of disease, rather than health, conceived in some other way. Causal dating is appropriate for the burden of disease. The burden of, say, polio in 2000 is appropriately seen as the harm done by the polio that occurs in 2000. If we adopt causal dating, in countries where polio is eradicated, thereafter there will no longer be any burden from polio. However, for many years afterwards, people will still be suffering from ill-health – specifically disabilities – caused by polio.

To adopt causal dating is not to reject the principle I announced at the beginning of this paper: that our measure of health is to be evaluative rather than purely descriptive. In causal dating, the dating is not evaluative; we date by the occurrence of disease. However, what we date is evaluative; it is the badness of the disease. We take the total harm that disease does, and we divide it up by time, according to the time when the harm was caused.

Causal dating is a way of overcoming the problem of inseparability, but it has problems of its own. In my example in figure 1, we are interested in only a single cause – the epidemic – that operates over a short period. This makes it easy to assign the harm to the date of the epidemic. But in practice, many causes operate simultaneously, and many of them continue to

operate over a long period. As a result, it may not be easy to attribute particular quantities of harm to particular causes, and so to particular dates. Well-known problems arise from overdetermination and from joint causation.

Here is one example. Suppose an epidemic in one year weakens the population, with the result that a second epidemic in the following year kills more people than it would otherwise have done. To which year should we assign the harm of those extra deaths? Which epidemic is to blame for them? No doubt the blame should be divided between them, but on what principles?

These are not questions of value. It is for the theory of causation to determine what caused the deaths. I do not know the answers to these questions, but I assume they can be answered. They are not at all like the questions of temporal separability I dealt with earlier. Causal dating has very different problems.

The idea of causal dating casts light on one issue that has surfaced in the discussion on summary measures of health. Should a summary measure be based on incidence statistics or prevalence statistics?⁸ The issue appears in this form because of the way summary measures are typically conceived. Each year, incidence and prevalence statistics are collected, and the summary measure is supposed to be constructed out of them in some way. Should it be constructed out of incidence statistics or prevalence statistics?

The incidence statistics for a particular year show the number of people who are struck down by, say, polio during the year. The prevalence statistics show the number who are, that year, suffering from the effects of polio. If we chose to base the summary measure for that year on prevalence statistics, that would be an attempt to measure the ill-health that afflicts people in that year as a result of polio. Some of this ill-health will result from an occurrence of the disease many years previously. I have explained that this attempt to measure ill-health in a year must fail, because times are not separable.

On the other hand, the incidence statistics for a year indicate the amount of disease that occurs in that year. If they are coupled with an appropriate estimate of the amount of ill-health that generally results from an occurrence, they could be used to construct a satisfactory causally-dated measure. Since I believe a measure can only properly be dated causally, I believe a dated measure can only properly be constructed out of incidence statistics.

Christopher Murray makes a similar point more picturesquely. He says, ‘There are no calculated measures of the prevalence of the dead’.⁹ His point is that, if we want to measure the ill-health that is being suffered at a particular time as a result of a disease, we ought not to forget all the people who have already died of the disease. They were harmed by it, but their harm is not registered in the prevalence statistics. So prevalence statistics are the wrong ones to use. Murray’s remark recalls Epicurus. We never find people around who are suffering from death. Yet dying is a definite harm.

But my point is slightly different. The problem with using prevalence statistics is that they attempt to measure the ill-health at a time, and this cannot be done because times are not separable. The reason they are not separable is not the one identified by Epicurus and repeated by Murray. It is more subtle. Murray’s objection to prevalence statistics could be overcome by inventing a figure for the prevalence of dead people. It would include all the people who would be alive had disease not killed them early. But this would still not make times separable. So even these augmented prevalence statistics could not give a proper measure of ill-health at a particular time.

6. *We cannot measure specifically health*

I now come to the question of why my diagrams show people’s temporal wellbeing, rather than their state of health. From the start, I have taken it for granted that our aim is to measure, or more strictly value, the effects of disease. Figure 1 showed the effects of an epidemic. It

showed the effects on all aspects of people's wellbeing, not just on their health. Although up to now I have used the term 'health', I have actually been discussing the aggregation of wellbeing rather than health.

However, a summary measure of health, of the sort propagated by the WHO, is not supposed to measure all the effects of disease on wellbeing. It is supposed to measure only the effects on health, and health is supposed to be only a part of wellbeing. In this section, I shall consider whether we can indeed produce a measure of specifically health. In section 7, I shall consider whether we should want to.

A person's temporal wellbeing at a particular time will be determined by many factors. Some are: how comfortably the person is housed, her opportunities for entertainment, how much she works and the sort of work she does, and so on. Other factors are all the ones that constitute the state of her health at the time: the quality of her eyesight, the functioning of her limbs, whether she is in pain, and so on. We could write her temporal wellbeing as a function of all these things:

$$w = w(h_1, h_2, \dots, h_m, d_1, d_2, \dots, d_n).$$

w is the person's temporal wellbeing at some particular time. h_1, h_2 and so on are all the individual factors that constitute her health. d_1, d_2 and so on are all the other factors that help to determine her level of wellbeing.

Suppose we could write the function that determines w in the special form:

$$w = v(h(h_1, h_2, \dots, h_m), d_1, d_2, \dots, d_n).$$

If the function could take this form, all the health factors could be evaluated together in a separate function $h(X)$ of their own, and then the value of this function would contribute to determining overall wellbeing w . We could evaluate the health factors together, independently of the other factors in wellbeing. In that case, the health factors would

technically be said to be *separable* from the other factors – this is the same notion of separability as we have come across before, in a new context.

If the health factors were indeed separable, we could take the function $h(X)$ to measure the person's health. For our purposes, it would be a highly satisfactory measure in two respects. First of all, it would measure health as a component of wellbeing, because health defined this way is one of the arguments in the function $v(X)$ that determines wellbeing w . For this reason, it would be a measure of the goodness of health – an evaluative measure as we require. Secondly, it would very clearly be a measure of specifically health, since it would be a function of only the health factors in wellbeing.

So it would suit us well if the health factors were separable. However, unfortunately they are not. Obviously, the influence on a person's wellbeing of the various elements of her health depends a great deal on other features of her life. For example, asthma is less bad if you are well housed, mental handicap less bad in supportive communities, deafness less bad if you have access to the internet. Conversely, features of a person's health affect the value of other things: radios are no good to the deaf, nor running shoes to the lame. The interaction between health and other features of a person's life is so intimate that health cannot be treated as separable.

It will be useful to have a simple example. Concentrate on just two health factors: the quality s of a person's sight and the quality t of her hearing. (Suppose the measures s and t are arbitrarily set to range between 0 and 1.) And concentrate on just two non-health contributors to the person's wellbeing: the number b of books she has, and the number c of CDs. Suppose her wellbeing is given by:

$$w = sb + tc.$$

Granted the stylized nature of the example, this function is plausible. The better you can see, the more benefit you get from books, and the better you can hear the more benefit you get

from CDs. Conversely, the more books you have, the more benefit you get from good sight, and the more CDs you have the more benefit you get from hearing. In this function, the two health factors are obviously not separable. The interaction between the health factors of wellbeing and the non-health factors is too intimate.

To be sure, we could define some function $H(h_1, h_2, \dots, h_m)$ of the health factors in a person's wellbeing to be a measure of the person's health. In the example, we might choose $H(X)$ to be $(s + t)$, say. A well designed measure of this sort might be useful for some purposes. For example, it might be useful in a causal investigation of the effects of a health programme. But it would not tell us how much health contributes to wellbeing. From the beginning, we decided to aim at an evaluative measure of health, so this is not what we are looking for.

Since health factors are not separable, must we abandon the idea of measuring health as a component of wellbeing? Not yet. The function $h(X)$ would have been an ideal measure of health, because it was a function of the health factors only. But we may still be able to find a less ideal measure, provided we do not expect it to be independent of other factors in wellbeing. A measure could still count as a measure of specifically health, if the other factors in wellbeing were held constant during the measurement. A measure of this sort can be called an *index* of health. An appropriate index would compare a person's actual wellbeing, given her state of health, with what her wellbeing would be if she were in good health, but if the non-health factors in her wellbeing were unchanged.

One way to form an index is to express the former as a fraction of the latter. Let H_1, H_2 and so on be the values of the health factors that represent good health. A person's actual wellbeing is $w = w(h_1, h_2, \dots, h_m, d_1, d_2, \dots, d_n)$. If she were in good health but other factors in wellbeing were unchanged, her wellbeing would be $w_H = w(H_1, H_2, \dots, H_m, d_1, d_2, \dots, d_n)$. We can measure her health by the fraction w/w_H . This fraction is her health index; let us call

it $*$. The person's actual wellbeing is $*w_H$, so $*$ is the factor by which ill-health reduces the person's wellbeing, other factors remaining constant. Because the health factors in wellbeing are not separable, we know that $*$ cannot depend only on the health factors.

The health index $*$ is not properly defined until we have fixed a zero for temporal wellbeing. I have been assuming a fixed zero since section 2, but so far I have said nothing about how it should be defined. Since it has now become crucial, I should put that right. The zero of temporal wellbeing is best set at the level of wellbeing that would be equally as good as dying. Compare a shorter life with a longer life that is just like the shorter one, but has a period added at the end, lived at the zero level of wellbeing. The zero level is defined to be such that these two lives are equally good.

In my simple example, the person's actual wellbeing is $(sb + tc)$. If she were in good health, both s and t would be 1, because they are scaled that way. So, if she were in good health but in other respects the same as she actually is, her wellbeing would be $(b + c)$. For her,

$$* = (sb + tc)/(b + c).$$

Here is a way to think of the index. Ill-health is bad for you in two ways. First, it is bad in itself; pain is bad in itself, for example. Second, it is instrumentally bad; it causes other bad effects. For example, ill-health can prevent you from getting out to see friends, and from earning a good income. This instrumental effect works by altering the non-health factors in your wellbeing. The index $*$ is an attempt to measure the badness of ill health itself, separated from its instrumental badness. So to set up $*$, we hold constant the non-health factors, and measure how much better off you would be if you were in good health, with those same non-health factors.

It is difficult to separate the instrumental from the non-instrumental badness of health, both in theory and in practice. Pain is bad in itself, and it also reduces your ability to enjoy music. Is the latter an instrumental effect? It depends whether we count your ability to enjoy music as a health or non-health factor in your wellbeing. I am not sure where to draw the boundaries of health.

But this theoretical problem does not matter much in practice, because in practice it is not possible in any case to hold constant all non-health factors in order to construct an index. The reason is that a measure for degrees of ill-health can only be constructed from people's preferences about different states of health, or from their judgements about how bad it is to be in a particular state of health. In practice, preferences or judgements have to be elicited from groups of people, through a questionnaire or in some other way.¹⁰ For example, to assign a measure to deafness, people are asked to indicate in some way how bad it is to be deaf, compared with having good hearing. If the result is to be the health index $*$, the subjects must hold constant in their imagination all non-health factors. They should certainly hold constant such plainly non-health factors such as a person's employment and her participation in social events. But they cannot be expected to do this reliably, because so much of what is bad about deafness is its instrumental effects on a person's employment opportunities and her participation in society. So in practice, we cannot hope to construct a proper health index $*$, keeping non-health factors constant.

Consequently, practical measures of health do not try. They do not try to hold constant the non-health factors in wellbeing. They simply take a person's wellbeing in her actual state of health, and compare it with the wellbeing she would have were she in good health. They take the former as a fraction of the latter, and use that fraction as the measure of her health.

A person's actual wellbeing is $w = w(h_1, h_2, \dots, h_m, d_1, d_2, \dots, d_n)$. Suppose that, if she were in good health, the values of her non-health factors would be D_1, D_2 , and so on. Then if she

were in good health, her wellbeing would be $W = w(H_1, H_2, \dots, H_m, D_1, D_2, \dots, D_n)$. The measure of her health is $q = w/W$. This measure takes account of the full effect that ill health has on the person's wellbeing, including all the indirect effects it causes by changing her non-health factors.

I shall call q the 'conventional measure of health' for a person at a single time. It is the type of measure that is used in practice. In qalys (quality-adjusted life years), the quality-adjustment factor for each particular state of health is a measure of this sort. In the WHO's studies of the burden of disease, as described by Christopher Murray, the 'disability weight' for a particular health-condition is also a measure of this sort.

In Murray's words, disability weights are evaluated for 'the average individual with the condition described taking into account the average social response or milieu'.¹¹ In this remark, the word 'response' indicates that Murray does not think of holding non-health factors constant. Instead, he takes them to change as, on average, they actually do change as a consequence of ill health.

The conventional measure q is how health is conventionally measured in practice. But since the non-health components in wellbeing are not kept constant, q is not really a measure of health specifically. It is simply a measure of wellbeing. True, it is scaled in a particular way. It is a person's wellbeing expressed as a fraction of what her wellbeing would be were she in good health. That is to say, it is her wellbeing w , expressed as a fraction w/W of her good-health wellbeing W . This scaling ensures that the value of q for a healthy person is always one. Nevertheless, q is nothing other than a measure of wellbeing that happens to have this particular scaling. It is not a measure of health specifically. Still, since it is conventionally called a measure of health, I shall continue to call q the 'conventional measure of health'.

To be sure, although q is not truly a measure of health, it has a solid connection with health. It measures how much a person's wellbeing is reduced by her ill-health, working as a cause. Our aetiological scheme now goes like this: disease causes ill-health, as well as other effects; ill-health is partly bad in itself, but it also causes other bad effects on wellbeing. We are taking it for granted that we aim to measure the badness of the *effects* of disease. But the question I have been exploring in this section is whether we can measure its effects specifically on health. It now emerges that we cannot, and that the measures we use in practice do not really even try to. They measure the reduction in wellbeing that is caused by ill health.

7. *We should not want to measure specifically health*

It turns out, then, that we cannot measure specifically health. But we should not want to, in any case.

In my illustrative example of figure 1, I included all the harm that is caused by an epidemic, by any causal process. The harms caused by disease are multifarious. For example, a person who is disabled by a disease may be unable to earn a living as a result, and for that reason suffer a loss of wellbeing. A disease may orphan some children, who will have less good lives as a result. All effects were supposed to be included in my diagram. All these effects should be counted into the burden of disease. The burden of disease should be conceived of as all the harm that is done by disease.

This is what should concern us when we make decisions in matters of health. It would be wrong for a health service to ignore those effects of its actions that are not effects on health specifically. Take an example. Suppose a particular disease particularly afflicts young women, and kills many of them. An important part of the harm done by this disease will be to leave many children motherless. This is not itself an effect on anyone's health, but it would

be wrong for the health service to ignore it. It should give an extra priority to this disease, compared with one that kills indiscriminately.

The aim of the health service is to control disease. That is to say, it is concerned with one particular cause of wellbeing, or of damage to wellbeing. Other causes are education, employment and so on, and other governmental institutions are concerned with those. The purview of each institution is picked out by the particular cause of wellbeing it is concerned with. It does not follow that different institutions should have different objectives. They should all be concerned with promoting wellbeing by means of whatever causes they are in charge of. (Wellbeing is not be all they should be concerned with; they should also be concerned with fairness, as I shall explain in section 9.) The role of the health service should be picked out by causes, not effects.

The idea of measuring the burden of disease is generally identified with the idea of measuring health. But, this now turns out to be a mistake. The burden of disease includes all the damage to people's wellbeing that results from disease. This sort of damage is illustrated in figure 1, and it is the true subject of this paper. From all this damage to wellbeing, we cannot separate out a component that is damage to health specifically. So we cannot truly have a measure of health, and we do not need one. It is the burden of disease that matters.

8. *Aggregate measures and fairness*

In section 6, I defined a measure q , and called it the 'conventional measure of health'. It is supposed to measure the health of a person at a time. But it is also intended to be aggregated across times in a person's life, and across people, to arrive at an aggregate measure of health. Since it is built on the conventional measure q , which is in truth a measure of wellbeing rather than specifically health, an aggregate measure formed like this is not in truth a measure of specifically health. However, since measures like this are *intended* to measure health and

since I need a name for them, I shall call them 'conventional aggregate measures of health'.

Qalys and dalys are examples.

The unaggregated conventional measure q is a measure of wellbeing. So aggregating it across time and across people means weighing together wellbeing that comes at different times and to different people. The scaling of q implies a particular basis for this weighing. As it happens, q is scaled in such a way that its value for a person who is in good health is one. This means that every period of life lived in good health is given the same value as every other, whoever lives it and whenever it is lived. Also, any period of life lived in some less good state of health is given the same value as any other period lived in that same less good state. Equally healthy years count as equal.

However, every healthy year is actually not equally as good as every other healthy year. Some healthy people are better off than others. Moreover, a healthy person may be better off at some times in her life than she is at others. Also, some people in less than good health are better off than others who are in the same less than good state of health. But the conventional aggregate measures count equally healthy years as equal. So they are not measures of aggregate wellbeing.

To reach this conclusion, I oversimplified. I presumed that a conventional measure aggregates across time and across people by simply adding up. It is simple addition that counts equally healthy years as equal. In practice, the aggregation is often more complicated. Often later times are discounted compared with earlier times; the WHO weighs periods of life according to the age of the person who lives them;¹² and so on. These adjustments have various purposes, but none is intended to make the aggregate measure any nearer to aggregate wellbeing. The conclusion remains: an aggregate measure arrived at by aggregating the conventional measure of health q is not a measure of aggregate wellbeing. Qalys and dalys are not measures of wellbeing, for example.

If we take one of these conventional aggregate measure of health as an objective for the health service, we shall not achieve the maximum of aggregate wellbeing. We shall make decisions that lead to less wellbeing than could have been achieved with the same resources. Within a single person's life, suppose we take as our objective to maximize a person's qalys, constrained by the resources available. Then we shall not maximize the person's lifetime wellbeing. If we take a conventional aggregate measure across people as our objective, we shall not maximize aggregate wellbeing across people. On the face of it, this seems a good reason not to take a conventional aggregate measure as an objective.

If we do, we shall maximize whatever the measure is a measure of. What is that? Conventional measures do not measure wellbeing, and that may explain why they are conventionally considered measures of health. But we know already that they do not truly measure health. In the rest of this paper, I shall argue that they measure nothing at all worth measuring.

I make one qualification to this claim. A conventional measure of health may constitute an acceptable approximate measure of wellbeing. We urgently need to measure the benefits of health-care. But we have no idea how to measure wellbeing as a whole, and we cannot expect to find out soon. There are deep difficulties over how wellbeing should even be conceived, and we cannot wait for these difficulties to be resolved. On the other hand, we do have some conventional health measures of the sort I have described. To be sure, they are subject to controversy in detail,¹³ but at least we are nearer to a consensus about them than we are to a consensus about measuring wellbeing. For the sake of getting on with what we need to do, we might adopt the heroic assumption that they actually measure wellbeing approximately. This implies that everyone in some particular state of health is equally as well off as everyone else in that state: the poor are equally as well off as the wealthy, provided they are in the same state of health; the materialistic are equally as well off as the spiritual; and so on. This is

false, but it may be an acceptable approximation for large-scale measurement in our pressing circumstances. According to this approximation, when we aggregate our health measure across time and across many people, we end up with an acceptable approximate measure of wellbeing. So the aggregate may serve as a reasonable objective for the health service. I find this a plausible defence of aggregate measures of health.

But there is a further, more principled defence. Several authors treat it as a positive merit of their measures that they ignore aspects of people's wellbeing besides health. Christopher Murray is among them. He offers it as a fundamental principle that 'the non-health characteristics of the individual affected by a health outcome that should be considered in calculating the associated burden of disease should be restricted to age and sex'.¹⁴ All equally healthy people of the same age and sex should be counted equally, whatever their level of wellbeing. This is implicit in the conventional measures of health, and Murray supports it as a matter of principle. He believes the health service should not aim to maximize wellbeing; it should take a conventional measure as its objective. He thinks fairness requires this. This principled defence of the conventional measures is a defence on grounds of fairness.

Murray argues by means of an example.¹⁵ Two patients are each in a coma caused by meningitis. Only one can be treated, and the other will die. Each has the same length of life ahead of her if she is saved. One is richer than the other. Suppose that the richer one will have greater wellbeing if she is saved than the other will, just because she is richer. Then more good would be done by saving this person than by saving the poor one. If we measure the benefit of saving each patient on the basis of its contribution to wellbeing, we shall find that saving the richer will be more beneficial.

Should we decide to save the richer patient on that account? Murray thinks not. He believes that doing so would be unfair to the poorer patient. For this reason he thinks our measurement of the benefit of treatment should be blind to the person's wealth. If we take

account of wealth, we shall act unfairly in cases like this. Fairness requires us to ignore such factors as wealth in a person's wellbeing. We should therefore treat the benefit of saving one patient as equal to the benefit of saving the other. This can be achieved by evaluating benefit by means of a conventional measure of health, rather than by a measure of wellbeing. A conventional measure gives the same value to a year of life belonging to either of the two patients. For this reason, Murray thinks we should adopt a conventional measure.

The rest of this paper is devoted to answering this argument from fairness.

9. *Fairness versus goodness: conventional measures*

I agree with Murray that we ought not to give priority to the richer patient in his example. It would be unfair to discriminate against a patient on grounds of wealth. But I do not agree that this should lead us to distort our measurement of benefit. The truth is that, if one patient will live a better life than the other, more good would be done by saving her than by saving the other. We should not hide from this truth.

Instead, we should learn something different from the example. We should learn that doing the most good – maximizing wellbeing – should not be our only objective in making decisions of this sort. We also need to take account of fairness. In this case, fairness determines that we should be impartial between the patients, even though the objective of maximizing benefit would lead us to discriminate between them. Once we see that fairness is a consideration that is distinct from goodness, we shall have no need to try and incorporate considerations of fairness into our measure of the goodness of health.

Put roughly, my answer to the argument from fairness is that fairness has no place in a measure of the burden of disease. I said in section 1 that we should aim to measure the goodness of health. After section 7, we know we must, more accurately, aim to measure the

badness of the burden of disease. But fairness is a distinct consideration from goodness or badness, and it must be accounted for separately.

That is oversimplified. As I said in section 2, I think fairness is actually a good thing. Consequently, a comprehensive measure of goodness would have to take fairness into account. But the good of fairness is achieved in a quite different way from other goods, and it is analytically most convenient to keep it separate. That is why in section 2 I did not include fairness as part of people's wellbeing. For the same reason, we should treat it separately from our measure of the burden of disease.

Why is fairness different?¹⁶ To begin with, it is a second-order value. It comes into play when some first-order good is to be distributed amongst people. In Murray's example, the first-order good is life-saving, but fairness is also involved in the distribution of many other first-order goods. Once fairness is in play, what does it demand? Although it is concerned with the distribution of some first-order good, it is not concerned with how much of this good people get, absolutely. It is concerned with how much they get in comparison to other people. It requires that each person receives a quantity of the good that is in proportion to her claims to it. In this paper, I shall consider only special examples, in which we may take it for granted that people's claims are equal. Given that, fairness requires the people to get equal shares of the good that is to be distributed. Even if they all get very little or none at all, they have been treated fairly.

Briefly, then, fairness requires equality or proportionality in shares of a first-order good, rather than maximizing. This is what makes it difficult to incorporate with other goods.

Now I have introduced a brief account of fairness, I need to interpolate a comment that refers back to the argument of section 7. Fairness is concerned with the distribution of a first-order good. It requires the good to be distributed to people in proportion to their claims to it. It is plausible that people's claims to one sort of good are often independent of their claims to

another sort. For instance, their claims to life-saving may be independent of their claims to income. If so, we can say there are separate 'spheres of fairness', and fairness in one sphere can be treated independently of fairness in another.

This has no tendency to suggest that, when we are concerned with promoting other sorts of good besides fairness, we can consider them independently from each other. For example, it does not suggest we can isolate health in a separate sphere from other factors in wellbeing, and aim to maximize health separately. If there are separate spheres of fairness, it does not follow that there are separate spheres for maximization. I have been arguing that there is no separate sphere for maximizing health. If there a separate sphere of fairness in matters of health, that has no tendency to oppose my argument.¹⁷

That is the end of the interpolated comment; now back to my line of argument. I believe fairness must be separated from other sorts of goodness. But in this paper I shall not try to argue for this claim generally. I shall only argue for the more limited claim that one particular approach to combining fairness with goodness is bound to fail. This is the approach that tries to incorporate fairness into a measure of the goodness of health by applying weights to wellbeing. It attaches weights to people's wellbeing, and then aggregates across people. The weights are designed to capture the demands of fairness. The idea is that we can achieve fairness by maximizing the weighted aggregate.

What sort of a weighted aggregate? I shall mention a scheme of variable weights in section 10, but I shall start with the simple case of fixed weights. Suppose a person lives for a number T of years, and suppose her temporal wellbeing in these years is successively w_1, w_2, \dots, w_T . The first thing we do is apply weights to these amounts of wellbeing, and add up, to arrive at a weighted total for the person, which I shall call x :

$$x = a_1w_1 + a_2w_2 + \dots + a_Tw_T.$$

The weights a_1 , a_2 and so on are fixed in some way or other; I shall mention one possible way in a moment. We find a weighted total similarly for each person, giving us weighted totals x_1 , x_2 , and so on. Then we add these weighted totals across people to get an overall aggregate:

$$X = x_1 + x_2 + \dots + x_N$$

(N is the number of people.) There is no need to apply weights at this stage, because any weighting that is needed can be incorporated into the first stage.

X is a function of the temporal wellbeings of all the people. But neither the individual x s nor the total X is a measure of wellbeing. The idea is that, if the weights are chosen properly, maximizing X will achieve fairness.

The conventional aggregate measure of health that I described in section 8 is a weighted function of this simple sort. To see this, start with a single person. We take her temporal wellbeings w_1 , w_2 and so on, and scale them in the way described in section 6. That is to say, we express them as fractions of what the person's temporal wellbeing would be were she in good health. So we divide them by W_1 , W_2 and so on, where these W s are what the person's levels of wellbeing would be in each year of her life, were she in good health. This gives us w_1/W_1 , w_2/W_2 and so on. In section 6, I defined a fraction like this as the conventional measures of health q . We now have one for each year in the person's life; let us write them q_1 , q_2 and so on. If we write $a_1 = 1/W_1$, $a_2 = 1/W_2$ and so on, we get that $q_1 = a_1w_1$, $q_2 = a_2w_2$ and so on. Then x for this person, as defined above, is

$$x = a_1w_1 + a_2w_2 + \dots + a_Tw_T = q_1 + q_2 + \dots + q_T.$$

This is the conventional aggregate measure of health for the person, over her lifetime. It is her quality-adjusted life years, in fact.

Adding these x s across people gives us an aggregate conventional measure for the people together. We can now see it is a weighted total of temporal wellbeings.

Murray calls on this aggregate measure to handle his example. The two patients in the example will have different levels of wellbeing if they are saved. But the conventional measure of health gives the same value to a year of the poorer patient's life as it gives to a year of the richer patient's, since both will be in good health if they are saved. The measure is blind to the people's wealth. If we take it as our objective, we shall not favour the richer patient over the poorer one. So in Murray's example we shall act fairly.

However, calling on a conventional measure of health cannot be the right way to achieve fairness in general. It happens to work in Murray's own example, but it will not give the right answer if we slightly modify the example. Suppose the two coma patients are equally rich, but one is disabled and the other is not. Say that one has lost an arm. If the disabled patient is saved, her wellbeing will be less than the wellbeing of the uninjured patient, if she is saved. Should the uninjured patient be saved on that account?

Conventional measures of health are blind to wealth, but we cannot make a measure of health blind to disability. Disability is one of the features of health that we are trying to measure. In the terminology of section 6, it is a health factor in wellbeing. So our measure cannot avoid giving more value to saving the uninjured patient.

Yet it would be unfair to save the uninjured patient for that reason. Murray himself appears at one point to be ambivalent about a case of this sort,¹⁸ but we should not be ambivalent. If it is unfair to discriminate between patients on grounds of their wealth, it is certainly unfair to discriminate on grounds of disability. It would plainly be unfair to let the disabled person die because she has already lost an arm. This modified example shows that the particular weighting scheme incorporated into a conventional measure of health cannot achieve fairness properly. If we take a conventional measure as our objective, and decide our actions on that basis, we shall not act fairly.¹⁹

10. *Fairness versus goodness: fixed and variable weights*

But the problem goes deeper. It is not merely that this particular weighting scheme is unable to capture the demands of fairness. No weighting scheme can.

To see why, think about this further modification of Murray's example. First, let me add some more detail. Each of the two patients needs the serum to save her life. If saved, each will live on for twenty years before finally dying. I shall assume that each person's wellbeing will be in proportion to the number of years of life she lives. (If this is not so – if length of life brings diminishing marginal benefits, for example – I need only to modify the figures below.) One patient is already disabled in some separate way. There is enough serum to save only one patient. I have already said that in these circumstances it would be unfair to give priority to the patient who is not disabled. Since the two patients have equal claims, fairness requires the two to be treated equally.

The only way of treating them equally is to save neither of them. I believe this would be perfectly fair, but it would sacrifice a life that could be saved. For the sake of saving a life, we should surely be willing to accept some unfairness. The only other thing we can do is save one of the two patients and allow the other to die. This would treat them unequally, so it would certainly not be completely fair. But there is a way to mitigate the unfairness. We can at least give each patient an equal chance of being saved, by tossing a coin to decide which to save. This will not achieve perfect equality, but it will achieve a sort of surrogate equality. Hence it will be fair to a lesser extent. For the sake of saving a life, this is surely what we should do.

Now I shall modify the example some more, and suppose the serum can be divided. This allows us to do better than tossing a coin: it gives us a way to be perfectly fair. We can divide the serum between the patients, so that each one gains the same number of years. However,

suppose that dividing the serum reduces its effectiveness a little. Undivided, it can give one of the patients twenty years of life; divided, it can give each of them nine years.

This sacrifice in the total is surely small enough to be worth accepting for the sake of fairness. So it is surely now plausible that the serum should be divided. At any rate, I shall assume this is the right thing to do. The figures do not matter, so long as it is worth making some sacrifice of the total number of years saved, for the sake of fairness. If fairness has any value, some small sacrifice is certainly worthwhile. If necessary, I could change the 9 years to 9.9, or any number less than 10.

Could maximizing a weighted total of wellbeing bring us to this conclusion that the serum should be divided? Can we find weights that will have that consequence? Let us try. I am assuming that the disabled patient will have a less good life than the other. So twenty years for her amount to a lesser quantity of wellbeing than twenty years for the other patient. Since fairness forbids us to give priority to the uninjured patient on this basis, we must apply corrective weights to wellbeing, so as to cancel out this difference. These will not be the weights that appear in the conventional measures of health. They must ensure that a year of each patient's life counts equally, even though the two patients will not be equally healthy if they are saved. In the formula above, x for each patient must be just the number of years she lives for. X will be the total of years lived by the two of them, added together. That is to say, we must choose weights to ensure that we maximize the total of life years.

In the example, we have a choice between twenty years to one patient, twenty years to the other, or eighteen years divided between the two. Faced with this choice, maximizing life years tells us to save one patient or the other, but not to divide the serum. Choosing appropriate weights can make us value each patient equally, but that is not the same as valuing equality between the patients. Fairness requires the latter as well as the former. So weights do not correctly capture the demands of fairness.

That is not the end of the argument, however. Actually, maximizing weighted life years can be made to recommend dividing the serum, if we use *variable* weights of a special sort. Rather than applying a weight to a person's total of life years, we might apply a different weight to each increment in her life years, and make this weight decrease with every increase in the number of life years a person already has. We give a lot of weight to extending the life of someone who has a few years only, and less to extending the life of someone who already has many. We give a person's years a diminishing marginal weight, as an economist would say.

The effect is that the overall value we give to years of a person's life is not proportional to the number of years. Twenty years are not worth twice as much as ten years. They might be worth twice as much as eight years, say. In that case, it is better for two patients in the example to get nine years each than for one of them to get twenty years. So in the example, the serum should be divided. Provided the weight diminishes fast enough, this will be the result.

The way to formulate diminishing marginal weight mathematically is to change the formula for our aggregate objective X . Instead of making it the simple sum of individual x s, as I did above, we first *transform* the x s before we add them up. This gives us the formula:

$$X = f(x_1) + f(x_2) + \dots + f(x_N)$$

The function $f(X)$ is the transformation. It must be an increasing function, and strictly concave. This means that its graph slopes upwards but bends downwards.²⁰

Diminishing marginal weight, expressed in this formula, represents the *priority view*. This is the view that we should give priority to worse-off people. In our context, we should give priority to people whose lives are shorter. The *priority view* is explained and discussed elsewhere in this volume. My own contribution to that discussion, 'Equality versus priority', argues that the formula above is the correct way to represent it.

In my most recent version of Murray's example, I showed that the priority view recommends dividing the serum. So it is doing well so far, as an attempt to capture the demands of fairness. Yet it, too, fails in the end to give a proper account of what fairness requires. To show this I need a yet more elaborate example.

Suppose we have the same two patients requiring treatment, but the options available are more complicated. We have a choice between two courses of action. Action *A* has a one-half chance of saving both patients, and a one-half chance of saving neither. Action *B* will for sure save one of the patients and for sure allow the other to die; each will have a one-half chance of being the one who is saved. In either case, any patient who is saved will live for twenty years.²¹

Which is the right action to choose? The answer is *A*, because it guarantees total fairness. Whatever happens, the two people will be treated equally under *A*, and in this example fairness demands equality. On the other hand, *B* guarantees unfairness, because one person will be saved and the other will die. To be sure, it offers some surrogate fairness instead, because each person has the same chance of being saved. But that is less than the full fairness certainly achieved by *A*.

However, the prioritarian formula above will never be able to value *A* above *B*. Let us value the alternative actions according to the formula. *A* has a one-half chance of saving both patients. If it does so, the value of *X* achieved will be $f(20) + f(20)$. *A* also has a one-half chance of saving neither patient; if that happens the value of *X* achieved will be $f(0) + f(0)$. So the *expected* value of *X* brought about by *A* is

$$\frac{1}{2}\{f(20) + f(20)\} + \frac{1}{2}\{f(0) + f(0)\}.$$

The corresponding calculation for *B* gives

$$\frac{1}{2}\{f(20) + f(0)\} + \frac{1}{2}\{f(0) + f(20)\}.$$

These are the same. So according to the prioritarian formula, *A* and *B* must have the same value: they are equally good. Yet fairness demands *A*.

One assumption I made implicitly in the last paragraph is open to doubt. I assumed we should evaluate the alternative actions by the expected value of their results, and that can be doubted. However, my 'Equality versus priority' justifies this assumption. The conclusion is sound: a system of variable weights cannot capture what fairness requires in this example.

In section 8, I explained that fairness can be offered as a principled reason to use a conventional measure of health. A conventional measure applies a particular weighting scheme to people's wellbeing. The idea is that this weighting scheme captures the demands of fairness. But the conclusion that emerges from the long sequence of examples and analysis in sections 9 and 10 is that a weighting scheme cannot possibly capture the demands of fairness. Conventional measures of health failed at an early stage. But I examined other, more sophisticated weighting schemes, even as far as the system of variable weights that appears in prioritarianism. All failed. We have to conclude it is hopeless to try and incorporate the requirements of fairness into a measure of goodness by means of a weighting scheme. Fairness matters, but it has to be accounted for separately.

11. *Conclusion*

There remains only one possible justification for using a conventional measures of health as an objective to be aimed at: the measure may be an acceptable approximation to a measure of wellbeing. It cannot be justified by some other principle.

In controlling disease, we should aim to promote people's wellbeing. We should also aim to be fair. To achieve these aims, we need first a measure of aggregate wellbeing, and second a measure of fairness. Ultimately, we shall need to put these two measures together to

determine an overall objective. But we should not distort our measure of wellbeing by trying to incorporate fairness into it; if we do, we shall fail.

This paper has principally been about the first of the two aims. I have examined, in a preliminary way, how we should measure aggregate wellbeing for the purposes of the health service. From the perspective of a health service, and reading the measure negatively, a measure of aggregate wellbeing is a measure of the burden of disease. I have examined how we should set about constructing one.

When we have a measure of the burden of disease, it will not be strictly correct to call it a measure of health. Even if we wanted to, we could not separate out the damage disease does to people's health specifically, from all the rest of the damage it does to people's wellbeing. But we should not want to. The burden of disease consists of all the damage that disease does. That is what we should measure, and that is what we should aim to reduce.

References

- Broome, John, 'Equality versus priority: a useful distinction', this volume.
 Broome, John, *Weighing Goods: Equality, Uncertainty and Time*, Blackwell, 1991.
 Broome, John, *Weighing Lives*, Oxford University Press, forthcoming.
 Culyer, A. J., 'Commodities, characteristics of commodities, characteristics of people, utilities, and the quality of life', in *Quality of Life*, edited by Sally Baldwin, Christine Godfrey and Carol Propper, Routledge, 1990, pp. 9–27.
 Epicurus, 'Letter to Menoeceus', in *Epicurus: The Extant Remains*, translated and edited by Cyril Bailey, Oxford University Press, 1926, pp. 83–93.
 Griffin, James, *Value Judgement*, Oxford University Press, 1997.
 Harris, John, 'QALYfying the value of life', *Journal of Medical Ethics*, 13 (1987), pp. 117–23.
 Kolm, Serge-Christophe, 'On health and justice', this volume.
 Murray, Christopher J. L., 'Rethinking DALYs', in *The Global Burden of Disease*, edited by Christopher J. L. Murray and Alan D. Lopez, Harvard University Press, 1996.
 Nord, Erik, *Cost-Value Analysis in Health Care*, Cambridge University Press, 1999.
 Wagstaff, Adam, 'QALYs and the equity–efficiency trade-off', *Journal of Health Economics*, 10 (1991), pp. 21–41.

Notes

I am very grateful for the comments I received on this paper from Dan Brock, Frances Kamm and Larry Temkin.

1. A possible use for a non-evaluative measure is mentioned in section 6 of this paper.
2. See sections 9 and 10 of this paper.
3. See my *Weighing Lives*, chapters 3 and 4.
4. There is a comprehensive discussion of it in my *Weighing Goods*, chapters 8 and 9.
5. *Weighing Lives*, chapter 2.
6. In his *Value Judgement*.
7. 'Letter to Menoecus', p. 85.
8. See Christopher Murray, 'Rethinking DALYs', pp. 8–9.
9. 'Rethinking DALYs', p. 9.
10. In 'Rethinking DALYs', p. 38–41, Christopher Murray describes the consultation process employed in the WHO's study of the global burden of disease.
11. Murray, 'Rethinking DALYs', p. 38.
12. Murray, 'Rethinking DALYs', p. 60.
13. An important discussion is Erik Nord's *Cost-Value Analysis in Health Care*.
14. 'Rethinking DALYs', p. 6. A. J. Culyer offers a similar view in his 'Commodities, characteristics of commodities, characteristics of people, utilities, and the quality of life'.
15. 'Rethinking DALYs', p. 7.
16. A fuller account of fairness appears in my *Weighing Goods*, chapter 9.
17. In 'On health and justice' in this volume, Serge Kolm considers the question of whether there is a separate 'sphere of justice' for health, and tentatively suggests there is. By 'justice' I think he means what I mean by 'fairness'. If he is right, it does not follow that we should be separately concerned to promote health, rather than wellbeing.
18. 'Rethinking DALYs', pp. 31–2.
19. For many years, this has been a standard objection to maximizing qalys as an objective. See John Harris's 'QALYfying the value of life', for instance.
20. This method of giving value to equality is very common in economic theory; see my 'Equality versus priority'. So far as I know, it was not used in health economics till recommended by Adam Wagstaff in his 'QALYs and the equity–efficiency trade-off'.
21. This example corresponds to the example of G and H in my 'Equality versus priority'.