

## **Does Cost Effectiveness Analysis Unfairly Discriminate against People with Disabilities?**

---

**GREG BOGNAR**

**ABSTRACT** *Cost effectiveness analysis is a tool for evaluating the aggregate benefits of medical treatments, health care services, and public health programs. Its opponents often claim that its use leads to unfair discrimination against people with disabilities. My aim in this paper is to clarify the conditions under which this might be so. I present some ways in which the use of cost effectiveness analysis can lead to discrimination and suggest why these forms of discrimination may be unfair. I also discuss some proposals for avoiding discrimination without rejecting cost effectiveness analysis altogether. I argue that none of these proposals is ultimately convincing. I describe a different approach to the problem and conclude by answering the question in the title.*

### **1. Introduction**

Cost effectiveness analysis is a tool for evaluating the aggregate benefits of medical treatments, health care services, and public health programs.<sup>1</sup> It combines the costs associated with different treatments with the health benefits obtained from them. On the basis of these data, it is possible to rank alternative treatments according to how many ‘units’ of health benefit they can realize for some level of expenditure. Given the costs, the more benefit a treatment can realize, the more effective it is.

The role of cost effectiveness analysis in health care resource allocation is often limited. Benefit maximization — for which cost effectiveness data are indispensable — is almost never used as the *only* distributive principle, and it seldom plays a part as a major principle. Despite this, cost effectiveness analysis has been the subject of many attacks both in theoretical debates and in public life. Its opponents claim that its use leads to unfair discrimination against people with disabilities or chronic health conditions. This is because cost effectiveness analysis seems to imply that treating such people is less valuable than treating others. In reply, its defenders point out that disregarding the differential benefits of treatments would lead to inefficient and wasteful allocation of scarce health care resources. If there is discrimination, it is not unfair.

My aim in this paper is to clarify the conditions under which the use of cost effectiveness analysis can lead to unfair discrimination. I first need to explain both cost effectiveness analysis and the discrimination objection in more detail. I present some ways in which the use of cost effectiveness analysis can lead to discrimination and suggest why these forms of discrimination may be unfair. Then I discuss some proposals for avoiding discrimination without rejecting the use of cost effectiveness analysis altogether. I will argue that none of these proposals is ultimately convincing. After that, I describe a different approach to the problem. I conclude by answering the question in the title.

## 2. Cost Effectiveness and Disability Discrimination

Cost effectiveness calculations set up a ratio of costs and benefits for treatments. In the numerator, there are monetary values; in the denominator, there is a measure of health benefit. The lower the value of the ratio, the more cost effective the treatment is.

Although many measures of health benefit can be used in cost effectiveness calculations, the ones most frequently used share a common feature: they are evaluative measures of health, reflecting the *value* of health. They do this by assigning weights to different health states according to their impact on the quality of life of the people who are in those health states. These weights are combined with the time in which a person is in the health state. Thus, a year spent with worse quality of life has a lower value than a year spent with better quality of life, and more years at a given quality of life level have higher value than fewer years.

The most familiar evaluative measures of health are quality adjusted life years (QALYs) and disability adjusted life years (DALYs). On both of these measures, health states are given weights between 0 and 1, where 1 represents full health (for QALYs) or a health state that is just as bad as death (for DALYs). Thus, one year in full health has a QALY value of 1 and a DALY value of 0. Decreasing values between 1 and 0 represent increasingly worse health states (for QALYs) or disabilities whose burden is increasingly less (for DALYs). The scales are inverted because QALYs are interpreted as a measure of health benefit, and DALYs are interpreted as a measure of disease burden.

QALYs are used by the National Institute for Health and Clinical Excellence (NICE) in the United Kingdom for technology assessment and the evaluation of the effectiveness of therapeutic procedures and the management of health conditions. NICE's cost effectiveness calculations inform prioritization decisions within the National Health Service (NHS). DALYs are used by the World Health Organization (WHO) to measure the global burden of disease and by various global health initiatives for assessing the success of their programs.

Despite these applications, cost effectiveness analysis remains controversial. Perhaps the best illustration is provided by the case of an aborted Medicaid reform in Oregon. In the late 1980s, the state government attempted to extend Medicaid coverage to all potentially eligible citizens by limiting the number of services that would be covered. The coverage decisions were to be made on the basis of cost effectiveness analyses for different condition-treatment pairs, using QALYs as a measure of health benefit. Treatments and services were ranked according to their cost/QALY ratio. The plan was that the most cost effective items on this priority list would be covered for all eligible citizens, but further down the list there would be a point below which treatments would be deemed not worth the costs and therefore not provided. The threshold was to depend on resource constraints, allowing both the extension of coverage and the maximization of health benefits.

After several rounds of revision of the priority list — during which cost effectiveness considerations gradually faded into the background — the Bush administration rejected the plan on the basis that it violates the Americans with Disabilities Act (ADA) by discriminating against people with disabilities.

When the plan was later resubmitted to the Clinton government, quality of life was no longer a component in the evaluation of the effects of treatments on health outcomes. It was hoped that the new plan would not be in violation of the ADA. The only factors that

were taken into account for any condition-treatment pair were the cost of avoiding death, the probability of death from the condition, and the probability of returning to an asymptomatic state after treatment. Nevertheless, the Clinton administration required Oregon to drop the third factor for the same reason as before: it argued that this factor could lead to unfair discrimination against people with disabilities.<sup>2</sup>

Quality of life enters into cost effectiveness calculations through the use of evaluative measures in the denominator of the cost/effectiveness ratio. By excluding all factors except cost and mortality risk, the Clinton government in effect excluded all quality of life considerations from Oregon's plan. Cost effectiveness analysis became largely irrelevant.

What opponents of cost effectiveness analysis find problematic is the use of *relative* quality of life information. Cost effectiveness calculations reflect the change in quality of life or life expectancy that a treatment provides; they are insensitive to absolute levels of quality of life before or after treatment. Thus, a measure like the QALY reflects *capacity to benefit*: the relative improvement that a treatment can provide for a patient. But some groups of patients, including people with disabilities or chronic health conditions, often (though not always) have a diminished capacity to benefit. For instance, a prior or pre-existing disability may act as a comorbidity, lowering the patient's prospects for improvement. Treatments for people with disabilities may also be more expensive, complex, or uncertain.<sup>3</sup>

Nevertheless, at first glance it is puzzling why the use of cost effectiveness information would discriminate against anyone. Cost effectiveness analysis ranks *treatments*. It provides information that can help select the most effective treatment for a patient or group of patients. NICE uses it to evaluate therapeutic procedures and technologies. The Oregon plan used it for ranking condition-treatment pairs. How can that be discriminatory?

Discrimination may ensue *indirectly*. Suppose people with disabilities need a particular treatment more often than others. (Some rehabilitation therapies may be examples.) If the treatment is given low priority because of its high cost or limited effectiveness, the health care needs of this group of patients will be unmet. This sort of discrimination, however, is not inherent in the use of cost effectiveness analysis. There need not be any intention to discriminate: the discrimination is merely a consequence of how costs and benefits work out in practice.

In such cases, the discrimination is a matter of contingent facts, and both defenders and opponents of cost effectiveness analysis can agree that the resultant disadvantages need to be addressed. They can agree that benefit maximization must be supplemented by other distributive principles. The possibility of indirect discrimination, therefore, does not provide an argument for denying cost effectiveness analysis a role in health care resource allocation. It just shows that it must be used together with other principles.

But discrimination may also arise *directly*. Cost effectiveness analysis can be used to divide patients into separate groups according to their capacity to benefit from a given treatment. I call this *partitioning*. A treatment can be made more cost effective by giving lower priority to those who have a diminished capacity to benefit from it. In this case, there is an intention to discriminate: rather than ranking treatments, cost effectiveness calculations are used to rank patients. Opponents of cost effectiveness analysis find partitioning particularly egregious. They believe it is unfair.

Defenders of cost effectiveness analysis point out that if capacity to benefit cannot be used in priority setting, then resource allocation will be inefficient and wasteful. Sometimes it might be desirable to consider separately the effectiveness of the same treatment for different patient groups. If a health care system is prohibited from distinguishing between those who can only be provided with marginal benefits and those who can be provided with substantial benefits, many people will end up worse off than they could be. And that could be unfair too.

So there is a dilemma: either it is permissible to take capacity to benefit into account in health care resource allocation or not. On the first horn of the dilemma, cost effectiveness analysis can be used to partition people according to their capacity to benefit, and many think this is unfair. On the second horn, effectiveness is largely ignored, which will lead to inefficient and wasteful allocation of scarce health care resources. Neither of these alternatives seems palatable.

In fact, we know from empirical studies that people's intuitions are somewhere between these two extremes. They agree that it would be unfair if some groups were systematically discriminated against in the health care system because of their diminished capacity to benefit; but they also agree that capacity to benefit can be taken into account in resource allocation to avoid gross inefficiencies. Fairness matters, but so does efficiency.<sup>4</sup>

### **3. Fairness and Efficiency**

Can the conflict between fairness and efficiency be resolved? Can we avoid unfair discrimination while remaining sensitive to efficiency considerations?

Perhaps we can't. Fairness and goodness are different moral considerations. Fairness is concerned with satisfying people's claims, and goodness is concerned with promoting benefits. In our context, fairness requires that health care resources be distributed in a way that avoids unfair discrimination; goodness requires that health benefits be allocated in the most efficient manner. These two requirements often conflict. Hence ultimately our task is to find the least unacceptable trade-offs.<sup>5</sup>

Undoubtedly, this is often the case. It is unfair if some people's claims are disregarded, or the claims of some people are treated as more important when everyone's claims should be equal. But unfairness can also arise in a different way.

Suppose you are concerned with goodness, or benefit; suppose also that a person's claim is simply a function of how much benefit she would get from the resource (that is, there is no conflict between fairness and goodness). So what you want to know in order to make your decision is simply whether giving the resource to this person or that person would do more good. But unfairness can also arise if your evaluation is based on the consideration of the wrong sorts of benefit. Even though one of the candidates would benefit if you make your decision on this basis, the sort of benefit that she would receive is morally irrelevant in the context. Letting these benefits tip the balance would be unfair.

As it happens, health care resource allocation offers a good example of the possibility of this sort of unfairness. In order to allocate health care resources efficiently, you could include economic benefits. This way, cost effectiveness calculations would reflect the social and economic benefits that some groups of patients would provide if they were

given priority in health care resource allocation. (Perhaps they could create more wealth or the livelihood of others depends on them.) But normally we do not give priority to more productive people in health care; we think this would be unfair. The unfairness arises not from disregarding some people's claims on the resource, but from considering the wrong sort of benefit — benefits which should be morally irrelevant in the given context. With respect to evaluating different health care resource allocations, economic benefits are irrelevant. These costs and benefits are in 'separate spheres'.<sup>6</sup>

People with disabilities often have diminished capacity to benefit, hence their health care needs will be given lower priority if effectiveness considerations are taken into account. According to the opponents of cost effectiveness analysis, this is unfair, since capacity to benefit is morally irrelevant — at least when it comes to setting priorities among patients. It should not play a role when health care resources must be allocated between different patient groups. This is then the question that is at issue between defenders and opponents of cost effectiveness analysis: is a person's capacity to benefit morally irrelevant? If not, under what conditions is it permissible to take it into account?

#### 4. Rationing and Partitioning

The use of cost effectiveness analysis can lead to discrimination against people with disabilities directly or indirectly. The latter depends on how costs and benefits work out in practice. If people with disabilities end up worse off this way, their disadvantage can be addressed by other principles. But cost effectiveness analysis can also lead to discrimination directly if it is used to partition patients into groups by their capacity to benefit. This form of discrimination raises questions about the fairness of the use of cost effectiveness analysis itself.

Consider the following example. A few years ago, several local trusts of the National Health Service in the UK announced that people with a body mass index (BMI) over 30 would be ineligible for hip and knee replacements, even if they were in serious pain. (The body mass index is used as a rough indicator for healthy body weight; a person with a BMI of 30 or above is regarded clinically obese.) The motivation was, as reports put it, 'unashamedly financial': the trusts were under government pressure to eliminate their deficits. Obesity reduces the capacity to benefit from hip replacements. The reduced benefits were deemed not worth the costs. The policy was not uncontroversial, but apparently it had some public support — including the support of local doctors, who emphasized the increased risks of operating on obese patients. The risks further reduce these patients' capacity to benefit.<sup>7</sup>

In the absence of surveys, it is hard to know what reasons the supporters of the policy would have given. But supporting the policy was not *obviously* wrong. In the face of severe resource constraints, it is not obviously unacceptable to give lower priority to those who have diminished capacity to benefit from hip replacements due to their obesity, even though it may turn out to be wrong all things considered.

I will now introduce some fancy additions to the example. Let us assume that the degree to which patients with a BMI over 30 would benefit from hip replacements is similar to that of some other groups: for instance, it is equal to the benefit that blind people would get from hip replacements. This might be because even with reasonable accommodation, blindness limits mobility and makes many everyday activities suffi-

ciently difficult so that hip replacements bring a relatively small quality of life improvement for the blind. Thus, I assume that the same savings can be achieved by denying hip replacements either to the obese or to the blind, while the aggregate health benefits remain equal.

Suppose then that a proposal was made that in the face of the resource constraints, hip replacements should be denied to blind people.

I think most people would find this proposal outrageous — even those who would not object to taking into account the impact of obesity on capacity to benefit. There is something about this proposal that is *obviously* wrong. It seems intuitively clear that denying hip replacements to blind patients would be a form of unfair discrimination.

This suggests that some ways of partitioning patients according to their capacity to benefit are more objectionable than others. But what is the moral difference? What makes the proposal about blindness a non-starter?

Perhaps there is no moral difference.<sup>8</sup> Perhaps the different reactions to the obesity and the blindness cases simply reflect a growing prejudice against people who are obese. Despite our intuitions, we should not treat the two cases differently. Perhaps denying hip replacements to people who are obese and to people who are blind is not unfair at all. Since it is not cost effective to provide the treatment to either of these patient groups, there is no unfairness.

Or perhaps it is equally unfair to select either of these groups. Suppose that, as in the example of the NHS trusts, there is severe resource scarcity. In order to maintain health care services, we have to cut costs, and the only way to do that is to deny hip replacements to some group of patients. We recognize this will be unfair to these patients, but taking the resource constraints into account, the unfairness is outweighed by the benefit of maintaining other services. If our reaction to the obesity case were based on mere prejudice, then there would be no difference with respect to the degree of unfairness when we select patients who are obese and patients who are blind.

It seems to me, however, that most people would agree that it remains less unfair to deny hip replacements to the obese than to the blind even if the dangers of prejudice are pointed out. So there remains a moral difference between the two cases. Nevertheless, perhaps the obesity case seems less objectionable for a different reason. For instance, even though no suggestion was made that personal responsibility for health justified the proposed policy of the local NHS trusts, some people might have supported it because they believed that people who are obese are responsible for their condition. If only they had made appropriate lifestyle choices, they could have avoided the need for a hip replacement. Since they brought their misfortune upon themselves, it is not unfair if they are given lower priority or denied the operation altogether.

Empirical studies show that people do consider personal responsibility for health a relevant factor in priority setting decisions. In several studies, a considerable number of respondents (although not the majority) were willing to give lower priority to smokers, substance users, and those who have an unhealthy diet.<sup>9</sup> But there is a worry that these judgments are more informed by perceptions of blame than genuine concern for personal responsibility. For suppose, once again, that the only way to cut costs is to deny hip replacements to some group of patients. There are two candidate groups: obese patients and former marathon runners. Members of the former group need new hips because of their obesity, and members of the latter group need new hips because of the strain they put on their bodies through years of long distance running. Since both of these groups

are equally responsible (we are assuming) for their health needs, whether or not personal responsibility matters, either it is not unfair to distinguish patients by their capacity to benefit, and then it is not unfair at all to deny hip replacements to members of either of these groups; or it would be equally unfair to deny the treatment to obese people and former marathon runners. But it seems not unreasonable to hold that it remains less unfair to deny hip replacements to the obese than to former marathon runners. So, whether or not personal responsibility matters, there remains a moral difference between the two cases, and this moral difference does not depend on beliefs about responsibility for health.

Notice that those who share the judgment that it would be more unfair to deny hip replacements to former marathon runners than to people who are obese appear willing to discriminate against a group of people with disabilities even if the other candidate group consists of people who have no prior disability at all. For being a former marathon runner is no disability.

So what might be the moral difference between a case like obesity and a case like blindness? In the next section, I will argue that there are more and less objectionable cases of partitioning patient groups by disability status because of the difference between a genuine comorbidity and an unrelated disability. In the example that I began with, clinical obesity seems to be a comorbidity and blindness is an unrelated disability. Our reactions to the two proposals differ because one of them disadvantages some group solely on the basis of an unrelated disability. Partitioning is permissible only if it avoids treating unrelated disabilities as comorbidities.

Before I present this argument, I look at some proposals for avoiding unfair discrimination against people with disabilities in health care resource allocation. Can they distinguish between comorbidities and unrelated disabilities? I briefly consider two alternatives.

One proposal for avoiding discrimination is to take into account *potential for health*.<sup>10</sup> For a simple illustration, suppose there are two patients at the same level of quality of life, suffering from the same illness. Treatment can return one of them to full health; however, because she has a prior disability, the other patient can be returned to a lower level of quality of life only. Since she has a diminished capacity to benefit, the potentials for health of these two patients are unequal.

A patient's potential for health is determined by the ratio of the actual improvement that can be realized from a treatment and the improvement that would return her to full health. Thus, if a treatment can return a patient to full health, the ratio will be 1; otherwise, it will be smaller. In order to offset the effect of the patient's diminished capacity to benefit, her potential improvement is weighted: the smaller her potential for health, the greater the weight. Empirical studies can be carried out to learn about the weights people give to realization of potential — that is, about their aversion to discriminating against those who have a diminished capacity to benefit. Thus, cost effectiveness calculations that include potential weights for health improvement can take into account both efficiency and the concern for enabling people to achieve the highest level of functioning that is possible for them. If treating a disabled person can substantially improve that person's functioning given what is possible for her, it may be desirable to give higher priority to treating her.

Potential weights are irrelevant when the choice is between different treatments for the *same* patient or group of patients, since in that case no issue of discrimination arises. But

they can help avoid unfairness when cost effectiveness analysis would lead to discrimination *indirectly*. If potential weights are applied to the improvement of those who have diminished capacity to benefit, they are likely to reduce the disadvantage that arises from the fact that the treatments that these patients need are less cost effective.

Partitioning is a different issue, however. Consider again the example of setting limits on the provision of hip replacements because of financial constraints. Compared to others, people with a body mass index over 30 have diminished capacity to benefit from the operation. What sort of priority ranking the potential for health proposal would establish for the obese and other groups depends on the size of their health benefits and the weights given to those benefits. It might be that the benefits of hip replacements for those who are not obese are sufficiently great to counterbalance the value of the realization of the health potentials of the obese. If so, it is not unfair to give lower priority to this group.

But, by assumption, people who are blind have similarly diminished capacity to benefit. The potential for health proposal would evaluate their situation the exact same way. On the possibility that I have been exploring — that it is more objectionable to give lower priority to people on the basis of certain kinds of disability than to do so on the basis of other disabilities — it matters what sort of disability one has. It is more objectionable to limit access to hip replacements for blind people than for those who are clinically obese. Potential weights do not reflect this difference.<sup>11</sup>

Proposals that assign different weights to the benefits of people who have a smaller potential for health or who are worse off does not reflect the distinction between disabilities that we intuitively make in cases like those of obesity and blindness. What is needed is a more direct approach for separating the cases in which partitioning is objectionable from those in which it isn't.

One place to look is work in non-consequentialist ethics. Frances Kamm, for instance, has written extensively about disability discrimination.<sup>12</sup> In her view, it would be unfair to let a *disability itself* influence resource allocation decisions, but it might be permissible to take into account its effects. As she puts it, 'we should, if we wish not to discriminate invidiously, abstract from the disability a person brings to the treatment situation as a *component of his life*, not from its *causal* role.'<sup>13</sup>

Thus, she argues that life expectancy can be taken into account: it is not unfair to prefer to give a scarce resource to a non-disabled person rather than a disabled person if the disability reduces the disabled person's life expectancy. The disability has an effect that is relevant to the allocation decision. In this case, you are not holding the disability against the disabled person in your choice — you are only taking into account its effects. Kamm also puts the distinction in terms of the effects of a disability as opposed to its 'intrinsic properties'. For instance, it is permissible to take into account the effects of haemophilia on life expectancy, but it is not permissible to take into account its intrinsic properties, like being painful.<sup>14</sup>

But it is not entirely clear how this distinction works. Perhaps Kamm means that being painful is intrinsic in the sense that it is experienced, while the reduction of life expectancy is not. (If your life is cut short, you won't experience its reduction; the lost time is not part of your life.) So you cannot take into account the effects of a disability on the quality of life, but you can take into account the effects on its *quantity*. But this just denies that any *quality* information can be taken into account. Perhaps Kamm means instead that some features of a disability may be essential in ways that its other features

are not. But then life expectancy is not a good example. Some disabilities — like cystic fibrosis — reduce life expectancy in a way that seems ‘intrinsic’ to them.

The problem is that it is difficult to distinguish a disability from its effects. The reason is that disabilities are functional limitations which affect major life activities (to borrow from the language used by the ADA). Since they are *defined* in functional terms, it is unclear which of their features are intrinsic and which are extrinsic. In the absence of an account of that, this proposal is unhelpful.

Neither does it help if we try to distinguish between disabilities in terms of whether they are a ‘component of a person’s life’. A disability should have a significant impact on life activities — a condition that does not limit functioning, or causes only trivial limitations, is usually not regarded a disability. In this sense, a disability seems to *always* be a component of life.

On Kamm’s view, what matters is the role the disability plays in the person’s life. However this may be further elaborated, what Kamm seems to be focusing on is the relation of the person to her disability. But this approach is unlikely to be able to capture what makes the obesity and the blindness cases different.

## 5. Which Disabilities are Relevant?

The proposals I have discussed focus on some characteristic of persons: their potential for improvement, the severity of their condition, or whether the disability is a component of their lives. Neither of them is able to account for the difference between the obesity case and the blindness case: that is, none of them can distinguish between genuine comorbidity and unrelated disability. For that, we need an approach that considers disabilities directly.

Consider the following distinction. Sometimes the condition associated with the treatment whose cost effectiveness is under consideration — called the *index condition* — is unrelated to, or *independent* of, the prior disability. Suppose there is a blind patient who needs a hip replacement. Because of her reduced mobility, this patient has a diminished capacity to benefit from the operation. Her hip problem, however, has nothing to do with her blindness — in the absence of the prior disability, the treatment would restore her to full health. In this example, the index condition and the prior disability are independent.

In other cases, however, the index condition and the prior disability are *dependent*. In the way I am using the term, for a condition to be dependent on a prior disability, it is both necessary that the disability is causally responsible for the condition and that its presence diminishes the patient’s capacity to benefit (for otherwise the discrimination problem would not arise). For instance, suppose that another patient also needs a hip replacement. This patient has a body mass index over 30. Moreover, his obesity is causally responsible for his need of the surgery. His capacity to benefit is diminished because of his weight: perhaps because his improvement in mobility can be only modest or perhaps because the duration for which his quality of life can be improved is limited. In this example, the index condition and the prior disability are dependent.

On the basis of this distinction, the following proposal can be formulated: *if the condition associated with the treatment whose cost effectiveness is under consideration is independent of the prior disability, and that disability diminishes the person’s capacity to benefit, then it would lead to unfair discrimination to take the impact of the disability into account.* The

presence of the prior disability should not be a relevant consideration for cost effectiveness analysis. On the other hand, if the index condition is dependent on the prior disability, its impact on capacity to benefit may be taken into account, and the discrimination that results from the use of cost effectiveness analysis might be not unfair. (It might be, however, unfair for other reasons, as I explained in Section 3.) For the purposes of health care resource allocation, the prior disability is not irrelevant.

Of course, it would be desirable that all of those who need a hip replacement be able to get one. It is important to keep in mind that the question we are trying to answer is conditional: *if* partitioning is inevitable because of resource constraints, what are the conditions under which disability status can be taken into account? What are the conditions under which partitioning by disability avoids (or at least minimizes) unfairness?

Consider the example of hip replacements again. Should we take into account the diminished capacity to benefit of the groups with prior disabilities for calculating the cost effectiveness of hip replacements? Is the partitioning of patients on this basis permissible? As I have argued, there seems to be a difference with respect to the kind of disability that we are considering. Most people are likely to find the suggestion that a different cost effectiveness ratio should be used for blind patients unquestionably unfair. I suggest this intuition can be explained by the fact that the medical need of these patients is independent of their disability.

Things might be different if the suggestion is that we should do the same thing only for the group of patients whose body mass index is over 30. Even though some people may continue to think that this would be wrong, others may have a less firm reaction to this case. I think many people would agree that this suggestion is not obviously unreasonable, even if all things considered it should be rejected. On the proposal I am exploring here, this judgment can be explained by the fact that the medical need is not independent of the prior disability. There is a moral difference between the two cases.

Consider also my earlier example of former marathon runners. These patients have no prior disability that diminishes their capacity to benefit. I assumed that they can be returned to full health by the hip replacement. On the basis of the distinction we are considering, it may be not unfair to give priority to them on account of their greater capacity to benefit as opposed to patients who are obese, given that the capacity to benefit of the latter group is diminished due to a genuine comorbidity. It would, however, be unfair to give priority to former marathon runners on account of their greater capacity to benefit when they stand in competition with blind patients. Blindness is an unrelated disability whose impact on capacity to benefit is morally irrelevant for the prioritization of hip replacements.

To be sure, someone might accept the empirical distinction but deny its moral significance. They might argue that it provides no basis for distinguishing between different cases. For suppose that the former marathon runners are now standing in competition with patients who have brittle bones. Having brittle bones is a prior disability that reduces these patients' capacity to benefit and is also responsible for their need for a hip replacement. Since it is a genuine comorbidity, it seems that it is not unfair to give priority to former marathon runners in the face of resource constraints. But this might strike some as unfair indeed.

Just as with the judgment about obesity, we should be careful in locating the source of this judgment. Perhaps what strikes some as unfair in this case has nothing to do with

whether the medical need is independent of the prior disability. Denying care to some patients while providing it to others may in a sense always be unfair given that all patients have equal claims to be treated regardless of their disability status — whether they are blind, obese, have brittle bones or some other disability. As I explained in Section 3, unfairness can arise when equal claims are not treated equally. In the comparisons of various disabled patient groups and former marathon runners what we have is a conflict between fairness and efficiency. Sometimes the benefits are sufficiently great to outweigh the unfairness of denying treatment to some patients. This does not remove the unfairness, even though it justifies it. This might be the case with the example of brittle bones as well.

Unfairness can also arise, however, if a choice is based on considering the *wrong sorts* of benefit. My proposal tries to explain this sort of unfairness. Two candidates for a health care resource may be equal with respect to health benefits, but differ with respect to economic benefits; that is, although they would derive the same benefit from the resource with respect to health, one of them would derive additional benefits. But these are different respects in which costs and benefits should be evaluated (or ‘separate spheres’), and in the context of health care resource allocation only capacity to benefit with respect to health is relevant.

But there is no reason to hold that this sort of unfairness cannot arise *within* the health care sphere as well. Cost effectiveness analysis is concerned with the evaluation of the costs and benefits of treatments and health programs with reference to some particular health condition. Hence we should treat the cost effectiveness calculation of a given treatment as a particular ‘respect’ for which costs and benefits are calculated. In order to be able to avoid discrimination, we have to make more fine-grained evaluations than we might have initially thought.

Thus, on my proposal, prior disabilities of which a condition is independent should be considered morally irrelevant, whereas disabilities on which a condition is dependent could be considered relevant. (At the cost of some imprecision, from now on I shall simply say dependent and independent disabilities for the sake of brevity.) Thus, the impact of an independent prior disability on a person’s capacity to benefit should not be taken into account in cost effectiveness analysis, whereas the impact of a dependent disability could be. The distinction between dependent and independent disabilities is important for the evaluation of the particular treatment whose effectiveness we are considering. Any condition-treatment pair is a ‘separate sphere’, and independent disabilities are not within its sphere.

Of course, even if someone accepts the moral significance of this sort of unfairness, they might still wonder how the distinction between dependent and independent disabilities can be formulated in a more precise way. Let me briefly address this issue.

One of the main difficulties with determining the relevance of prior disabilities is due to the complexity of determining whether an index condition is independent of a disability. Comorbidities are present in many cases, and it is often unclear whether one causes the other, or they have a common cause, or they are unrelated. In other cases, the index condition has multiple causes, and the prior disability is only one of them. What’s more, even if their independence can be established, identifying the impact of individual disabilities presents additional difficulties, especially if their impact is synergistic. Hence it may not always be possible to determine conclusively whether a prior disability is relevant.<sup>15</sup>

But the empirical difficulties may to some extent help with the conceptual difficulties. On the one hand, the *less* difficult it is to establish independence, the more confident we can be that taking the effects of the prior disability into account could lead to unfair discrimination, and hence it should be considered irrelevant. On the other hand, the *more* difficult it is to establish independence, the more confident we might be that the prior disability is relevant to determining a person's capacity to benefit.

Still, it might be objected that there is a deeper problem here. In the foregoing discussion, the distinction between the dependence and independence of index conditions and prior disabilities has been drawn in somewhat intuitive terms. In the biomedical literature, an independent condition is one that appears purely by chance; if, on the other hand, a disability contributes to the risk of the occurrence of another condition, then the condition that develops is dependent. This rough way of drawing the distinction, however, is too broad and ambiguous.<sup>16</sup> Even if it can be established that a prior disability is a risk factor that is responsible for the condition, it must be shown that it is responsible in the *causally appropriate way* — that is, through the right sort of causal mechanism. If our blind patient needs a hip replacement because of accidents that would not have occurred had she been able to see, does this make her present condition dependent on the prior disability? Is her blindness causally responsible for the condition in the appropriate way? In order to answer such questions, we need an account of causation in biomedicine.<sup>17</sup>

I do not think, however, that there is only one such account. It is more likely that several accounts are possible that draw the distinction at different places depending on how far back they look on the causal chain, under what conditions they treat a condition as a significant enough risk factor, and what they consider an appropriate causal mechanism. Ultimately, we have to use our considered judgments to select an account for doing the work of distinguishing relevant and irrelevant disabilities. The best we can hope for is some sort of reflective equilibrium between our considered judgments and the best account we have of causation in biomedicine. In this respect, the proposal made on these pages is more appropriate as a framework for thinking about the question of cost effectiveness and disability discrimination, rather than a determinate answer. But the indeterminacy is not all bad news, for part of what needs to be explained is why reasonable disagreements remain about the permissibility of partitioning by different disabilities.<sup>18</sup>

## 6. Does Cost Effectiveness Analysis Unfairly Discriminate against People with Disabilities?

According to the disability discrimination objection, the use of cost effectiveness analysis can lead to unfair discrimination. The unfairness arises because cost effectiveness analysis relies on quality of life measures that reflect a person's capacity to benefit, and people with disabilities tend to have diminished capacity to benefit. The problem, then, is whether the impact of a prior disability on a person's capacity to benefit can be taken into account. I proposed that independent prior disabilities should be considered irrelevant for cost effectiveness calculations; dependent disabilities, however, remain relevant. On this view, cost effectiveness analysis may indeed lead to unfair discrimination if the impacts of independent prior disabilities are taken into account.

In practice, however, quality adjusted measures seldom take prior disabilities into account. Consider disability adjusted life years, for example. The disability weights used for DALYs were determined on the basis of individual conditions only. In measuring the burden of disease, the burden of co-disabilities are, with a few exceptions, simply the sum of the burdens of the individual disabilities. When the measure was developed, there was no systematic attempt to distinguish between dependent and independent disabilities and to establish disability weights for comorbid conditions.

In addition to emphasizing the methodological complications that this would introduce, the developers of DALYs justified their choice by arguing that if the composite disability weights were smaller than the sum of the weights of individual disabilities, co-disabilities would contribute *less* to the burden of disease — that is, they would be regarded as less bad. People with co-disabilities would be disadvantaged, since treating two patients with two separate disabilities would lead to a greater reduction of the disease burden than treating one patient who has both disabilities. The developers of the DALY measure argued that this would lead to unfair discrimination, and chose to err on the side of caution to avoid the ethical problems.<sup>19</sup>

When cost effectiveness analyses use QALYs as their outcome measure — as in the Oregon Medicaid reform or by the National Institute for Health and Clinical Excellence — cost effectiveness ratios are usually calculated for specific treatment-condition pairs. These ratios can be ranked from most effective to least effective, and a cut-off point may be established below which treatments are not considered cost effective. The health-related quality of life measures on which quality adjustment factors are based encompass the consequences of a condition for different domains or aspects of a person's life, but these measures seldom include more complex health states that can be characterized by the presence of multiple conditions. Co-disabilities are usually ignored in cost effectiveness calculations.<sup>20</sup>

As I have argued, the way in which cost effectiveness analysis can be directly discriminatory is partitioning: dividing patients into separate groups according to disability status and calculating different cost effectiveness ratios for groups with and without a disability. On the proposal I have discussed, the partitioning would avoid unfair discrimination for any condition-treatment pair only if the disability and the condition were dependent.

In practice, however, partitioning is very rare. Real-world attempts at priority setting and burden of disease measurement tend not to take into account the impact of co-disabilities at all. Thus, generally speaking, cost effectiveness analysis discriminates less against people with disabilities than it might have seemed.

*Greg Bognar, Center for Bioethics, New York University, 285 Mercer Street, New York, NY 10003, USA. greg.bognar@nyu.edu*

### **Acknowledgements**

I would like to thank audiences at New York University, Queen's University, and at the third annual conference of the Harvard University Program in Ethics & Health, as well as Dan Brock, John Broome, Norman Daniels, Frances Kamm, Sam Kerstein, Cameron McIntosh, Josh Salomon, Markus Schrenk, Carlos Soto, and an anonymous referee for comments on an earlier version of this paper.

NOTES

- 1 For the sake of brevity, I will talk of treatments in this paper. As far as I can see, however, what I have to say below applies to any activity in health care and public health for which cost effectiveness analysis is applicable.
- 2 For details, see James F. Blumstein, 'The Oregon experiment: The role of cost-benefit analysis in the allocation of Medicaid funds', *Social Science and Medicine* 45 (1997): 545–554, and David C. Hadorn, 'The Oregon priority-setting exercise: Quality of life and public policy', *Hastings Center Report* 21 (1991): 11–16.
- 3 Henceforth, I will use the notion of disability inclusively for any disabling chronic health condition. This is in line with the ADA's definition of disability that includes any chronic medical condition that substantially impairs major life activities.
- 4 Numerous empirical studies have shown this; for an overview, see Erik Nord, Jose Luis Pinto, Jeff Richardson, Paul Menzel & Peter Ubel, 'Incorporating societal concerns for fairness in numerical valuations of health programmes', *Health Economics* 8 (1999): 25–39. For a different classification of the ways the use of cost effectiveness analysis can lead to discrimination, see Dan W. Brock, 'Cost-effectiveness and disability discrimination', *Economics and Philosophy* 25 (2009): 27–47.
- 5 On fairness, see John Broome, 'Fairness' in *Ethics Out of Economics* (Cambridge: Cambridge University Press, 1999), pp. 111–122; for the argument in this context, see John Broome, 'Fairness versus doing the most good', *Hastings Center Report* 24 (1994): 36–39.
- 6 The idea of separate spheres of distribution comes from Michael Walzer, *Spheres of Justice: A Defense of Pluralism and Equality* (New York: Basic Books, 1983), but in the debate on health care resource allocation, it is used in a somewhat different way (see Dan W. Brock, 'Separate spheres and indirect benefits', *Cost Effectiveness and Resource Allocation* 1, 4 (2003): 1–12). Roughly, a sphere is defined by the purpose and the objectives of an activity (e.g. health care), and it determines a 'respect' with regards to which costs and benefits are evaluated.
- 7 See 'NHS cash crisis bars knee and hip replacements for obese', *The Guardian* 23 November 2005.
- 8 I thank an anonymous referee for pressing me to explore the possibilities below.
- 9 For an overview, see, for instance, Jan Abel Olsen, Jeff Richardson, Paul Dolan & Paul Menzel, 'The moral relevance of personal characteristics in setting health care priorities', *Social Science & Medicine* 57 (2003): 1163–1172. Of course, respondents may be unaware of recent findings about the genetic bases of obesity (see Marten Hofker and Cisca Wijmenga, 'A supersized list of obesity genes', *Nature Genetics* 41 (2009): 139–140) and the role of the social determinants of body weight. We know, for instance, that in developed countries there is a strong negative association between socioeconomic status and body mass index (in developing countries, the association is positive, although this is changing). The association is stronger for women than men, and there are also marked differences with respect to race and ethnicity, residential area, and immigration status. (For a study on the population of California, see Emma V. Sánchez-Vaznaugh, Ichiro Kawachi, S. V. Subramanian, Brisa N. Sánchez & Dolores Acevedo-García. 'Do socioeconomic gradients in body mass index vary by race/ethnicity, gender, and birthplace?' *American Journal of Epidemiology* 169 (2009): 1102–1112; for an overview of earlier and international work, see the references therein.)
- 10 See Erik Nord, *Cost-Value Analysis in Health Care: Making Sense out of QALYs* (Cambridge: Cambridge University Press, 1999) and Nord *et al.* *op. cit.*
- 11 A similar proposal is to apply prioritarianism to health care resource allocation. Prioritarianism is the view that benefits have greater moral weight the worse off the people who receive those benefits are. A prioritarian view, just like the potential for health proposal, might help avoid discrimination that arises indirectly, since the groups who are disadvantaged as a consequence of the use of cost effectiveness analysis tend to be among the worse off. But partitioning is once again a different issue. A prioritarian view is sensitive to people's absolute positions and the size of the benefit that they can get, but it does not reflect the difference between more and less objectionable forms of partitioning. For a discussion of the issues raised by applying prioritarianism to health care resource allocation, see Dan W. Brock, 'Priority to the worse off in health-care resource prioritization' in R. Rhodes, M. P. Battin & A. Silvers (eds) *Medicine and Social Justice: Essays on the Distribution of Health Care* (New York: Oxford University Press, 2002), pp. 362–372.
- 12 See Frances M. Kamm, 'Deciding whom to help, health-adjusted life years and disabilities' in S. Anand, F. Peter & A. Sen (eds) *Public Health, Ethics, and Equity* (Oxford: Oxford University Press, 2004), pp. 225–242; 'Aggregation, allocating scarce resources, and the disabled', *Social Philosophy & Policy* 26 (2009):

- 148–197; ‘Disability, discrimination, and irrelevant goods’ in K. Brownlee & A. Cureton (eds) *Disability and Disadvantage* (New York: Oxford University Press, 2009), pp. 260–318.
- 13 Kamm, ‘Aggregation, allocating scarce resources, and the disabled’ op. cit., p. 188, her emphases.
- 14 See Kamm, ‘Deciding whom to help, health-adjusted life years and disabilities’ op. cit., p. 240.
- 15 For background, as well as an empirical study of the effects of combinations of conditions on disability, see Lois M. Verbrugge, James M. Lepkowski & Yuichi Imanaka, ‘Comorbidity and its impact on disability’, *The Milbank Quarterly* 67 (1989): 450–484; for a review of measurement methods of comorbidity, see Vincent de Groot, Heleen Beckerman, Gustaaf J. Lankhorst & Lex M. Bouter, ‘How to measure comorbidity: A critical review of available methods’, *Journal of Clinical Epidemiology* 56 (2003): 221–229. I have also learned from discussing these issues with Cameron McIntosh.
- 16 To complicate matters, co-occurring conditions with a shared etiology are also sometimes referred to as dependent disabilities or comorbidities. (See Gavin Andrews, Kristy Sanderson & John Beard, ‘Burden of disease. Methods of calculating disability from mental disorder’, *British Journal of Psychiatry* 173 (1998): 123–131.) On my approach, they would be independent, given that there might be no causal relation between them at all. Moreover, a common cause may not necessarily be any disability or health condition — for instance, in a case when the disabilities are the result of a genetic predisposition.
- 17 I owe this objection to John Broome.
- 18 There is a parallel debate in health economics on whether the effects of prior disabilities and comorbidities can be taken into account in cost effectiveness analysis, although it is not formulated in terms of discrimination. For an argument in favour, see, for instance, Ryan A. Harris & Robert F. Nease, ‘The importance of patient preferences for comorbidities in cost-effectiveness analysis’, *Journal of Health Economics* 16 (1997): 113–119; for an argument against, see, for instance, Melissa M. Brown, Gary C. Brown & Sanjay Sharma, *Evidence-Based Medicine to Value-Based Medicine* (Chicago, IL: American Medical Association Press, 2005), pp. 161–166. See also Han Bleichrodt, David Crainich & Louis Eeckhoudt, ‘Comorbidities and the willingness to pay for health improvements’, *Journal of Public Economics* 87 (2003): 2399–2406.
- 19 For the details of the argument, see Christopher J. L. Murray, ‘Rethinking DALYs’ in C. J. L. Murray & A. D. Lopez, (eds) *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020* (Cambridge, MA: Harvard School of Public Health on behalf of the WHO and the World Bank, 1996), pp. 1–98, at 38–42. He also suggests, however, that combined disability weights may be appropriate for some dependent co-disabilities and future burden of disease studies might consider introducing them.
- 20 On this point, see also Louise B. Russell, Marthe R. Gold, Joanna E. Siegel, Norman Daniels & Milton C. Weinstein, ‘The role of cost-effectiveness analysis in health and medicine’, *Journal of the American Medical Association* 276 (1996): 1172–1177.