Justice, Health, and Healthcare

Norman Daniels, Tufts University

Healthcare (including public health) is special because it protects normal functioning, which in turn protects the range of opportunities open to individuals. I extend this account in two ways. First, since the distribution of goods other than healthcare affect population health and its distribution, I claim that Rawls’s principles of justice describe a fair distribution of the social determinants of health, giving a partial account of when health inequalities are unjust. Second, I supplement a principled account of justice for health and healthcare with an account of fair process for setting limits or rationing care. This account is provided by three conditions that comprise “accountability for reasonableness.”

Three Questions of Justice

A theory of justice for health and healthcare should help us answer three central questions. First, is healthcare special? Is it morally important in ways that justify (and explain) the fact that many societies distribute healthcare more equally than many other social goods? Second, when are health inequalities unjust? After all, many socially controllable factors besides access to healthcare affect the levels of population health and the degree of health inequalities in a population. Third, how can we meet competing healthcare needs fairly under reasonable resource constraints? General principles of justice that answer the first two questions do not, I argue, answer some important questions about rationing fairly. Is there instead a fair process for making rationing decisions?

About twenty years ago I answered the first question by claiming healthcare was special because of its impact on opportunity (Daniels 1981, 1985). Specifically, the central function of healthcare is to maintain normal functioning. Disease and disability, by impairing normal functioning, restrict the range of opportunities open to individuals. Healthcare thus makes a distinct but limited contribution to the protection of equality of opportunity. Though I construed healthcare broadly to include public health as well as individual preventive, acute, and chronic care, I ignored other factors that have a profound effect on population health. Unfortunately, focusing on just healthcare adds to the popular misconception that our vastly improved health in the last century is primarily the result of healthcare.

During the last twenty years, a major literature has emerged exploring the social determinants of health. We have long known that the richer people are, the longer and healthier their lives. The powerful findings of the last couple of decades, however, have deepened our understanding of the factors at work producing these effects on population health and the distribution of health within populations. It is less tenable to think that it is simply poverty and true deprivation that diminishes the health of some people, for there is growing evidence that race and class effects operate across a broad range of inequalities. Since social policies—not laws of human nature or economic development—are responsible for the social and economic inequalities that produce these health effects, we are forced to look upstream from the point of medical delivery and ask about the fairness of the distribution of these goods. Rawls’s theory of justice as fairness, quite serendipitously, contains principles that give a plausible account of the fair distribution of those determinants, thus providing an answer to the second question (Daniels, Kennedy, and Kawachi 1999, 2000).

During the 1980s I became aware that my account of just healthcare, like other general theories, failed to give specific guidance, or gave implausible answers, to certain questions about rationing (Daniels 1993). Though philosophers may work out middle-level principles that can supplement general accounts of distributive justice and solve these unsolved rationing problems, it is unlikely that there will be consensus on them in the foreseeable future. Distributive issues remain highly contested.

In the absence of consensus on distribution principles, we need a fair process to establish legitimacy for critical resource allocation decisions. My account of a fair process for addressing these distributive issues is called “accountability for reasonableness” (Daniels and Sabin 1997, 1998a, n.d.). It
is an attempt to connect views about deliberative democracy to decision making at various institutional levels, whether public or private, in our complex health systems.

My goal in this essay is to sketch the central ideas behind my approach to all three questions and to suggest how they all fit together. Detailed arguments can be found in the references. By pushing a theory of justice toward providing answers to all three questions, and not simply the first, I hope to give a fuller demonstration that justice is good for our health.

**What is the special moral importance of healthcare?**

The central moral importance, for purposes of justice, of preventing and treating disease and disability with effective healthcare services (construed broadly to include public health and environmental measures, as well as personal medical services) derives from the way in which protecting normal functioning contributes to protecting opportunity. Specifically, by keeping people close to normal functioning, healthcare preserves for people the ability to participate in the political, social, and economic life of their society. It sustains them as fully participating citizens—normal collaborators and competitors—in all spheres of social life.

By maintaining normal functioning, healthcare protects an individual’s fair share of the normal range of opportunities (or plans of life) reasonable people would choose in a given society. This normal opportunity range is societally relative, dependent on various facts about the society’s level of technological development and social organization. Individuals’ fair shares of that societal normal opportunity range are the plans of life it would be reasonable for them to choose were they not ill or disabled and were their talents and skills suitably protected against mis- or underdevelopment as a result of unfair social practices and the consequences of socioeconomic inequalities. Individuals generally choose to develop only some of their talents and skills, effectively narrowing their range of opportunities. Maintaining normal functioning preserves, however, their broader, fair share of the normal opportunity range, giving them the chance to revise their plans of life over time.

This relationship between healthcare and the protection of opportunity suggests that the appropriate principle of distributive justice for regulating the design of a healthcare system is a principle protecting equality of opportunity. Any theory of justice that supports a principle assuring equal opportunity (or giving priority to improving the opportunities of those who have the least opportunity) could thus be extended to healthcare. At the time I proposed this approach, the best defense of such a general principle was to be found in Rawls’s theory of justice as fairness (Rawls 1971). One of the principles Rawls’s social contractors would choose is a principle assuring them *fair equality of opportunity* in access to jobs and offices. This principle not only prohibits discriminatory barriers to access, but requires positive social measures that correct for the negative effects on opportunity, including the underdevelopment of skills and talents, that derive from unfair social practices (e.g., a legacy of gender or race bias) or socioeconomic inequalities. Such positive measures would include among other things the provision of public education and other opportunity-improving early childhood interventions.

Rawls, however, had deliberately simplified the formulation of his general theory of justice by assuming that people are fully functional over a normal life span. His social contractors thus represented people who suffered no disease or disability or premature death. By subsuming the protection of normal functioning under (a suitably adjusted version of) his principle assuring fair equality of opportunity, I showed how to drop that idealization and apply his theory to the real world. (Rawls (1993) supports this approach.) In the last two decades, however, other work on egalitarianism has suggested alternative ways to connect healthcare to opportunity or to positive liberty or capabilities, and I shall comment on them shortly. First, I want to highlight some key elements of my approach.

The fair equality of opportunity account does not use the impact of disease or disability on welfare (desire satisfaction, happiness, or utility) or utility as a basis for thinking about distributive justice. One might have thought, for example, that what was special about healthcare was that good health was important for happiness. But illness and disability may not lead to unhappiness, even if...
they restrict the range of opportunities open to an individual. Intuitively, then, there is something attractive about locating the moral importance of meeting healthcare needs in the more objective impact on opportunity than in the more subjective impact on happiness.

This analysis fits well with and extends Rawls’s (1971) nonwelfarist account of primary social goods. For purposes of justice, Rawls argued, we should not seek to determine what we owe each other by measuring our satisfaction or welfare, but we should measure our levels of well-being by publicly accessible measures. For Rawls this means an index of primary social goods that includes rights and liberties, power and opportunity, income and wealth, and the social bases of self-respect. My account includes the protection of normal functioning within the scope of the primary good of opportunity. Drawing on insights from Scanlon’s (1975) discussion of the “urgency” of meeting some “preferences” to relieve decrements in well-being but not others, my account explains why we believe we have obligations to assist others in meeting healthcare needs but not necessarily to provide them with other things they say they need to make them happier.

Consider an actual issue where the contrast is important. People with long-standing disabilities will often rank their welfare higher than do people who are merely imagining life with such disabilities. Perhaps people with disabilities accommodate by adjusting their goals and expectations. Even if they are more satisfied with their lives than people without disabilities might expect, there is an objective loss in their range of capabilities and opportunities, and that loss is captured by the appeal to a fair share of an opportunity range. The fair equality of opportunity account thus avoids a troubling feature that haunts cost-utility analysis and its treatment of such disabilities.2

Healthcare is of special moral importance, because it helps to preserve our status as fully functioning citizens. By itself, however, this does not distinguish healthcare from food, shelter, and rest, which also meet basic needs of citizens by preserving normal functioning. Since medical needs are more unequally distributed than these other needs and can be catastrophically expensive, they are appropriately seen as the object of private or social insurance schemes. It might be argued that we can finesse the problem of talking about the medical needs we feel we are obliged to meet for one another if we assure people fair income shares from which they can purchase such insurance. We cannot, however, define a minimal but fair income share unless it is capable of meeting such needs (Daniels 1985).

Some economists and philosophers may object that giving special status to health insurance will be “paternalistic” and inefficient since some people prefer to trade income for things other than healthcare.3 Our social obligation, however, is to provide institutions (such as social insurance or subsidies to buy private insurance) that protect opportunity, not to maximize aggregate welfare or achieve efficiency above all else. The principles of justice defended here thus depart from utilitarian goals.

The account sketched here has several implications for the design of our healthcare institutions and for issues of resource allocation. Perhaps most important, the account supports the provision of universal access to appropriate healthcare—including traditional public health and preventive measures—through public or mixed public and private insurance schemes. Healthcare aimed at protecting fair equality of opportunity should not be distributed according to ability to pay, and the burden of payment should not fall disproportionately on the ill (Daniels 1985, 1995; and Daniels, Light, and Caplan 1996).

Properly designed universal coverage health systems will be constrained by reasonable budgets, since healthcare is not the only important good. Reasonable resource constraints will then require judgments about which medical needs are more important to meet than others. Priority setting and rationing is thus a requirement of justice, since meeting healthcare needs should not and need not be a bottomless pit.

The elderly might object that an opportunity based account of just healthcare will leave them out in the cold, for their opportunities might seem

2. Daniels (1996a) discusses the relationship between the equal opportunity account and the rationale for “reasonable accommodation” required of employers under the Americans with Disabilities Act; see also Brock (1995). Brock (1998) discusses the implication of disabilities for cost-effectiveness analysis; see also Buchanan et al. (2000), ch. 7.

3. Economists view an arrangement as efficient or pareto optimal if no one can be made better off without making someone else worse off. If some people trade their access to healthcare through socially provided insurance for other goods, they are not in a pareto optimal situation.
to be in the past. We can avoid this by not biasing our allocations in favor of one stage of life and instead considering the age-relative opportunity range. Still, treating people differently at different stages of life—for example, saving resources from one stage of life for use at another—does not produce inequalities across persons in the way that differential treatment by race or gender does. We all age—though we do not change gender or race. Fairness between age groups in designing a healthcare system is appropriately modeled by the idea of prudent allocation over a life span (Daniels 1988). Under some conditions of scarcity, this implies that “pure” rationing by age (where age is not proxy for other traits) is permissible. 4

Some universal coverage healthcare systems permit a supplementary tier that is purchased by those who are best off in society. For example, the British private insurance sector allows about 10% of the population to have quicker access to services that others must wait longer for in the British National Health Service. Other countries, such as Norway, prohibit a supplementary tier, fearing it will undermine the political solidarity needed to sustain their generous healthcare system. The fair equality of opportunity account constrains, but does not rule out, all tiering (Daniels 1998a; Daniels, Light, and Caplan 1996).

One controversial implication of my approach provides a way to contrast the fair equality of opportunity view with some alternative egalitarian accounts. In aiming at normal functioning, my approach views the prevention and treatment of disease and disability as the primary rationale for what we owe each other by way of assistance in cooperative healthcare schemes (Buchanan et al. 2000). Enhancing otherwise normal conditions—even when they put us at a disadvantage compared to others through no fault of our own—is then viewed as “not medically necessary.” For example, there is support in my view for the common insurance practice of covering treatment for very short children who have growth hormone deficiencies but not covering it for equally short children who are otherwise normal.

The objection to my view is that this coverage policy seems to place too much weight on the presence of disease and disability and too little on what really should matter to an account aiming at protecting opportunity, namely, reducing the disadvantage that extreme shortness brings. This objection might be pressed by those who defend “equal opportunity for welfare or advantage” (Arneson 1988; G. A. Cohen 1989). Their view rests on the claim that anyone who suffers bad “brute luck”—a deficit in welfare or advantage that is no fault of their own—has a claim on others for assistance or compensation. In contrast, bad “option luck,” the result of the choices we make or are responsible for making, does not give rise to claims on others. A disadvantage in talents or skills or even height that is not our fault thus provides a basis for claims on others for compensation or possibly enhancement. I argue (Daniels 1990, 2000a) that this view gives too much centrality to choice or responsibility, a centrality we do not and should not recognize when we want to protect our capabilities as citizens in a democratic society; there are good policy objections to this view as well (Sabin and Daniels 1994).

A similar objection might be raised from a perspective grounded in the importance of positive liberty or freedom, thought of as our capability to do or be what we choose (Sen 1980, 1992, 1999). The claim is that we should not necessarily be focused on a concept such as disease or disability but rather on whether individuals have the appropriate set of capabilities to do or be what they choose. Perhaps the very short child who is otherwise normal still lacks a key trait or capability that we should address.

If we consider more carefully, however, when differences in capabilities give rise to claims on others, support for treating the short but normal child may disappear. Sen (1992) himself notes that many differences in capabilities will be “incommensurable,” since there will be no consensus about whether a person is worse off than others. The short but normal child, for example, may have an excellent temperament or wonderful social or cognitive skills. The cases where there is likely to be agreement that someone is clearly worse off in capabilities are likely to be captured by the catego-

---

4. The age rationing implied by this account is different in rationale from that advocated by Callahan (1987), who thinks the old have a duty to step aside in favor of the young; it is also different from those who argue for a version of the “fair innings” view, which gives priority to the young on the grounds that the old have already had their opportunity to acquire years (see for example, Brock 1989; Williams 1997); it is also different in rationale from Kamm (1993), who argues that the young would be worse off than the old and in that sense “need” years more than the old. The considerable disagreement about what justice permits, even among those who accept some forms of age rationing, argues for the importance of the type of fair process described later in this paper.
ries of (serious) disease and disability. In practice, then, this view converges much more with the view I defend than it may appear at first.5

These alternative views obviously deserve more careful discussion than I can offer here. Still, my answer to the original question, that the special moral importance of healthcare derives from the protection of our opportunities, remains a defensible member of a family of views connecting healthcare to our opportunities and capabilities. Moreover, its practical implications converge more with those of its cousins than is apparent from the family quarrels among them.

**Which health inequalities are unjust?**

Universal access to appropriate healthcare—just healthcare—does not break the link between social status and health that I noted earlier, a point driven home in studies of the effects on health inequality of the British National Health Service (Black et al. 1988; Acheson et al. 1998; Marmot et al. 1998), and confirmed by work in other countries as well (Kawachi, Kennedy, and Wilkinson 1999). Our health is affected not simply by the ease with which we can see a doctor—though that surely matters—but also by our social position and the underlying inequality of our society. We cannot, of course, infer causation from these correlations between social inequality and health inequality (though later I explore some ideas about how the one might lead to the other). Suffice to say that, while the exact processes are not fully understood, the evidence suggests that there are social determinants of health (Marmot 1999).

If social factors play a large role in determining our health, then efforts to ensure greater justice in health outcomes should not focus simply on the traditional health sector. Health is produced not merely by having access to medical prevention and treatment, but also—to a measurably greater extent—by the cumulative experience of social conditions over the course of one’s life. By the time a sixty-year-old heart attack victim arrives at the emergency room, bodily insults have accumulated over a lifetime. For such a person, medical care is, figuratively speaking, “the ambulance waiting at the bottom of the cliff.” Much contemporary discussion about reducing health inequalities by increasing access to medical care misses this point.

5. See Daniels (2000a); also Rawls (1993). The convergence is clearer still when Sen (1999) addresses the ways in which we should focus on our capabilities as citizens—see Anderson (1999).

Of course, we still want that ambulance there, but we should be looking as well to improve social conditions that help to determine the health of societies.

As I noted earlier, Rawls’s theory of justice as fairness was not designed to address issues of healthcare. He assumed a completely healthy population, and argued that a just society must assure people equal basic liberties, guarantee that the right of political participation has roughly equal value for all, provide a robust form of equal opportunity, and limit inequalities to those that benefit the least advantaged. When these requirements of justice are met, Rawls argued, we can have reasonable confidence that others are showing us the respect that is essential to our sense of self-worth. The fair terms of cooperation specified by these principles promote our social and political well-being.

The conjecture I explore is that by establishing equal liberties, robustly equal opportunity, a fair distribution of resources, and support for our self-respect—the basics of Rawlsian opportunity—a fair allocation of resources within the healthcare system. Still, acting to promote social justice is a crucial step toward improving our health because there is this surprising convergence between what is needed for our social and political well-being and for our mental and physical health.

To see the basis for this conjecture about Rawlsian principles, let us review very briefly some of the central findings in the recent literature on the social determinants of health. If we look at cross-national studies, we see that a country’s prosperity is related to its health, as measured, for example, by life expectancy: In richer countries, people tend to live longer. But the relationship between per capita gross domestic product (GDPpc) and life expectancy levels off at around $8,000 to $10,000; beyond this threshold, further economic advance buys virtually no further gains in life expectancy. This leveling effect is most apparent among the advanced industrial economies. Nevertheless, even within this relationship, there are telling variations. Though Cuba and Iraq are equally poor (each has a GDPpc of about $3,100),
life expectancy in Cuba exceeds that in Iraq by 17.2 years. The poor state of Kerala in India, which invested heavily in education, especially female literacy, has health outcomes far superior to the rest of India and more comparable to those in much wealthier countries. The difference between the GDPpc for Costa Rica and the United States is enormous (about $21,000), yet Costa Rica’s life expectancy exceeds that of the United States (76.6 to 76.4).

Taken together, these observations show that the health of nations depends, in part, on factors other than wealth. Culture, social organization, and government policies also help determine population health. Variations in these factors—not fixed laws of economic development—may explain many of the differences in health outcomes among nations.

One especially important factor in explaining the health of a society is the distribution of income: the health of a population depends not just on the size of the economic pie, but also on how the pie is shared. Differences in health outcomes among developed nations cannot be explained simply by the absolute deprivation associated with low economic development—lack of access to the basic material conditions necessary for health, such as clean water, adequate nutrition and housing, and general sanitary living conditions. The degree of relative deprivation within a society also matters.

Numerous studies support this relative-income hypothesis, which states, more precisely, that inequality is strongly associated with population mortality and life expectancy across nations (Wilkinson 1992, 1994, 1996). Rich countries vary in life expectancy, and that variation dovetails with income distribution. In particular, wealthier countries with more equal income distributions, such as Sweden and Japan, have higher life expectancies than the United States, despite having lower per capita GDP. Likewise, countries with low GDPpc but remarkably high life expectancy, such as Costa Rica, tend to have a more equitable distribution of income.

We find a similar pattern when we compare states within the United States. If we control for differences in state wealth, income inequality accounts for about 25% of the between-state variation in age-adjusted mortality rates (Kennedy, Kawachi, and Prothrow-Stith 1996; Kawachi et al. 1997). Furthermore, a recent study across U.S. metropolitan areas found that areas with high income inequality had an excess of death compared to areas with low inequality—a very large excess, equivalent in magnitude to all deaths due to heart disease (Lynch et al. 1998). Longitudinal studies, which look at a single place over time and examine widening income differentials, support similar conclusions.

At the individual level we also find that inequality is important. Numerous studies have documented what has come to be known as the socioeconomic gradient: At each step along the economic ladder, we see improved health outcomes over the rung below (even in societies with universal health insurance). Differences in health outcomes are not confined to the extremes of rich and poor, but are observed across all levels of socioeconomic status.

The slope of the socioeconomic gradient varies substantially across societies. Some societies show a relatively shallow gradient in mortality rates: Being better off confers a health advantage, but not so large an advantage as elsewhere. Others, with comparable or even higher levels of economic development, show much steeper gradients. The slope of the gradient appears to be fixed by the level of income inequality in a society. The more unequal a society is in economic terms, the more unequal it is in health terms. Moreover, middle income groups in a country with high income inequality typically do worse in terms of health than comparable or even poorer groups in a society with less income inequality. We find the same pattern within the United States when we examine state and metropolitan area variations in inequality and health outcomes (Kennedy et al. 1998; Lynch et al. 1998).

Earlier, I cautioned that correlations between inequality and health do not necessarily imply causation. Still, there are enough plausible and identifiable pathways through which social inequalities appear to produce health inequalities to make a reasonable case for causation. In the United States the states with the most unequal income distributions invest less in public education, have larger uninsured populations, and spend less on social safety nets (Kaplan et al. 1996; Kawachi and Kennedy 1997). Studies of educational spending and educational outcomes are especially striking. Controlling for median income, income inequality explains about 40% of the variation between states in the percentage of children in the fourth grade who are below the basic reading level. Similarly strong associations are seen for high school drop-out rates. It is evident from these data that educational opportunities for children in high income inequality
states are quite different from those in states with more egalitarian distributions. These effects on education have an immediate impact on health, increasing the likelihood of premature death during childhood and adolescence (as evidenced by the much higher death rates for infants and children in the high inequality states). Later in life these effects appear in the socioeconomic gradient in health.

When we compare countries, we also find that differential investment in human capital—in particular, education—is a strong predictor of health. Indeed, one of the strongest predictors of life expectancy among developing countries is adult literacy, particularly the disparity between male and female adult literacy, which explains much of the variation in health achievement among these countries after accounting for GDPpc. For example, among the 125 developing countries with GDPpcs of less than $10,000, the difference between male and female literacy accounts for 40% of the variation in life expectancy (after factoring out the effect of GDPpc). In the United States differences among the states in women’s status—measured in terms of economic autonomy and political participation—are strongly correlated with higher female mortality rates.

These societal mechanisms—for example, income inequality leading to educational inequality leading to health inequality—are tightly linked to the political processes that influence government policy. For example, income inequality appears to affect health by undermining civil society. Income inequality erodes social cohesion, as measured by higher levels of social mistrust and reduced participation in civic organizations. Lack of social cohesion leads to lower participation in political activity (such as voting, serving in local government, volunteering for political campaigns). And lower participation, in turn, undermines the responsiveness of government institutions in addressing the needs of the worst off. States with the highest income inequality, and thus lowest levels of social capital and political participation, are less likely to invest in human capital and provide far less generous social safety nets (Kawachi and Kennedy 1999).

Rawls’s principles of justice thus turn out to regulate the key social determinants of health. One principle assures equal basic liberties, and specifically provides for guaranteeing effective rights of political participation. The fair equality opportunity principle assures access to high quality public education, early childhood interventions (including day care) aimed at eliminating class or race disadvantages, and universal coverage for appropriate healthcare. Rawls’s “Difference Principle” permits inequalities in income only if the inequalities work (e.g., through incentives) to make those who are worst off as well off as possible. This principle is not a simple “trickle down” principle that tolerates any inequality so long as there is some benefit that flows down the socioeconomic ladder; it requires a maximal flow downward. It would therefore flatten socioeconomic inequalities in a robust way, assuring far more than a “decent minimum” (Cohen 1989). In addition, the assurances of the value of political participation and fair equality of opportunity would further constrain allowable income inequalities.

The conjecture is that a society complying with these principles of justice would probably flatten the socioeconomic gradient even more than we see in the most egalitarian welfare states of northern Europe. The implication is that we should view health inequalities that derive from social determinants as unjust unless the determinants are distributed in conformity with these robust principles. Because of the detailed attention Rawls’s theory pays to the interaction of these terms of fair cooperation, it provides us—through the findings of social science—with an account of the just distribution of health.

The inequalities in the social determinants that are still permitted by this theory may produce a socioeconomic gradient, albeit a much flatter one than we see today. Should we view these residual health inequalities as unjust and demand further redistribution of the social determinants?

I believe the theory I have described does not give a clear answer. If the Rawlsian theory insists that protecting opportunity takes priority over other matters and cannot be traded for other gains (and Rawls generally adopts this view), then residual health inequalities may be unjust. If health can be traded for other goods—and all of us make such trades when we take chances with our health to pursue other goals—then the account may be more flexible (Daniels, Kennedy, and Kawachi 1999). Still, Rawls’s principles give us more specific guidance in thinking about the distribution of the social determinants than is given by the fair equality of opportunity account of just healthcare alone.

I noted earlier that there is considerable convergence between the opportunity-based view I defend and A. K. Sen’s (1992) appeal to a capabili-
ties-based account (or freedom based account) of the target of justice. The convergence is even more pronounced when Sen (1999) discusses the ways in which health in developing countries is affected by different development strategies and emphasizes the importance of education and the growth of democratic culture and institutions. Rawls’s focus on the “capabilities of free and equal citizens” suggests the convergence works in both directions (Daniels 2000a). Both approaches allow us talk informatively about justice and the distribution of health.

When are limits to healthcare fair?
Justice requires that all societies meet healthcare needs fairly under reasonable resource constraints. Even a wealthy, egalitarian country with a highly efficient healthcare system will have to set limits to the healthcare it guarantees everyone (whether or not it allows supplementary tiers for those who can afford them). Poorer countries have to make even harder choices about priorities and limits. However important, healthcare is not the only important social good. All societies must decide which needs should be given priority and when resources are better spent elsewhere.

How should fair decisions about such limits be made? Under what conditions should we view such decisions as a legitimate exercise of moral authority?

Answering these questions would be much simpler if people could agree on principles of distributive justice that would determine how to set fair limits to healthcare. If societies agreed on such principles, people could simply check social decisions and practices against the principles to see if they conformed with them. Where decisions, practices, and institutions fail to conform, they would be unjust and people should then change them. Disagreements about the fairness of actual distributions would then be either disagreements about the interpretation of the principles or about the facts of the situation. Many societies have well-established and reliable, if imperfect, legal procedures for resolving such disputes about facts and interpretations.

Unfortunately, there is no consensus on such distributive principles for healthcare. Reasonable people, who have diverse moral and religious views about many matters, disagree morally about what constitutes a fair allocation of resources to meet competing healthcare needs—even when they agree on other aspects of the justice of healthcare systems, such as the importance of universal access to whatever services are provided. We should expect, and respect, such diversity in views about rationing healthcare. Nevertheless, we must arrive at acceptable social policies despite our disagreements. This moral controversy raises a distinctive problem of legitimacy: Under what conditions should we accept as legitimate the moral authority of those making rationing decisions?

I shall develop the following argument:

1. We have no consensus on principled solutions to a family of morally controversial rationing problems, and general principles of justice for health and healthcare fail to give specific guidance about how to solve them (Daniels 1993);
2. In the absence of such a consensus, we should rely on a fair process for arriving at solutions to these problem and for establishing the legitimacy of rationing decisions (Rawls 1971); and
3. A fair process that addresses issues of legitimacy will have to meet several constraints that I shall refer to as “accountability for reasonableness” (Daniels and Sabin 1998a); these constraints tie the process to deliberative democratic procedures (Daniels and Sabin 1997, n.d.). This issue of legitimacy and fair process arises in both public and mixed public-private healthcare systems and it must be addressed in countries at all levels of development.

To support the first step of the argument, consider a problem that has been labeled the “priorities problem” (Daniels 1993; Kamm 1993): How much priority should we give to treating the sickest or most disabled patients? To start with, imagine two extreme positions. The Maximin position (“maximize the minimum”) says that we should give complete priority to treating the worst off patients. One might think that Maximin is implied by the fair equality of opportunity account (though I believe my account is only committed to giving some priority to the worst off, placing it in a broad family of views that leave the degree of priority unspecified). The Maximize position says that we should give complete priority to treating the worst off patients. One might think that Maximin is implied by the fair equality of opportunity account (though I believe my account is only committed to giving some priority to the worst off, placing it in a broad family of views that leave the degree of priority unspecified). The Maximize position says that we should give priority to whatever treatment produces the greatest net health benefit (or greatest net health benefit per dollar spent) regardless of which patients we treat.

Suppose comparable resources could be invested in Technologies A or B, but the resources are “lumpy” (we cannot introduce some A and some B), and we can only afford one of A or B in our healthcare budget. The Maximin position
would settle the matter by determining whether patients treated by A are worse off before treatment than patients treated by B. If so, we introduce A; if patients treated by B are worse off, we introduce B. If the two sets of patients are equally badly off, we can break the tie by considering to whom we can provide the most benefit. The Maximize position chooses between A and B solely by reference to which produces greatest net benefit.

In practice, most people are likely to reject both extreme positions (Nord 1995, 1999). If the benefits A and B produce are nearly equal, but patients needing A start off much worse than patients needing B, most people seem to believe we should introduce A. They prefer to provide A even if they know we could produce somewhat more net health benefit by introducing B. But if the net benefit produced by A is very small, or if B produces significantly more net benefit, then most people will overcome their concern to give priority to the worst off and will prefer to introduce B to A. Some people who would give priority to patients needing A temper their preference if those patients end up faring much better than patients needing B. Disagreement persists: A definite but very small minority are inclined to be maximizers and a definite but very small minority are inclined to be maximiners. Most people fall in between, and they vary considerably in how much benefit they are willing to sacrifice in order to give priority to worse off patients.

Two other types of rationing problems also suggest we are not straight maximizers or maximiners, though we lack principled characterizations of acceptable middle-course solutions (Daniels 1993). The Fair Chances/Best Outcomes problem asks, “Should we give all who might benefit some chance at a resource, or should we give the resource to those who get the best outcome?” The Aggregation problem asks, “When do lesser benefits to many outweigh greater benefits to a few?”

Two strategies have been pursued to address these kinds of rationing problems, one philosophical, one empirical. The philosophical approach, brilliantly exemplified by Kamm’s work (1993), examines subtly varied hypothetical cases, seeking to reveal agreement on a complex set of underlying principles that can account for the judgments the philosophical inquirer makes about these cases.6

6. Kamm insists on exploring hypothetical cases or thought experiments rather than real ones, attempting to isolate more clearly in these cases the relevant features that motivate our judgments. She believes that her method will uncover an “internal program” or underlying moral structure to our beliefs. Crucial to this approach is the claim that people will agree on a central range of cases, that is that others will have the same responses Kamm does to them. For doubts about the method, see Daniels (1998b).

This strategy may well help us arrive at middle-level principles for addressing these rationing problems, and it should be pursued by others. Nevertheless, given the subtlety of the method and the likelihood that some disagreements about cases will reflect broader moral disagreements about other matters, I do not believe this method will produce consensus on such principles in the foreseeable future. The insights from this approach are important inputs into a fair, deliberative process of decision making, but they are not a substitute for such a fair process.

The empirical approach has been ingeniously developed by the economist Erik Nord (1999), who also explores hypothetical cases by asking groups of people “person-tradeoff” questions. These questions are a variation on a standard economic approach seeking “indifference” points or curves reflecting when an individual finds two benefits or outcomes equivalent. For example, if we can invest only in treatments A and B, and A is used for people more seriously ill than B, we might ask how many treatments with B someone would trade for some number of treatments with A. Nord hopes this approach can uncover the structure of moral concerns in a population of people. A key risk of the method is that it disguises moral disagreement by talking about a “range” of responses. For this, and other reasons I discuss elsewhere (Daniels 1998c, 2000b), the results of Nord’s work can help inform fair, deliberative decision processes but cannot substitute for them.

If we have persistent disagreements about principles for resolving rationing problems, then we must retreat to a process all can agree is a fair way to resolve disputes. The retreat to procedural justice as a way of determining what is fair when we lack prior agreement on principles is a central feature of Rawls’s account (thus “justice as (procedural) fairness”). Rather than argue for this familiar approach (the second step of my argument above), I shall move directly to characterizing the features of such a fair process.

We would take a giant step toward solving the problems of legitimacy and fairness that face public agencies and private health plans making limit-
setting decisions if the following four conditions were satisfied (Daniels and Sabin 1997): 7

**Publicity Condition:** Decisions regarding coverage for new technologies (and other limit-setting decisions) and their rationales must be publicly accessible.

**Relevance Condition:** The rationales for coverage decisions should aim to provide a reasonable construal of how the organization (or society) should provide “value for money” in meeting the varied health needs of a defined population under reasonable resource constraints. Specifically, a construal will be “reasonable” if it appeals to reasons and principles that are accepted as relevant by people who are disposed to finding terms of cooperation that are mutually justifiable.

**Appeals Condition:** There is a mechanism for challenge and dispute resolution regarding limit-setting decisions, including the opportunity for revising decisions in light of further evidence or arguments.

**Enforcement Condition:** There is either voluntary or public regulation of the process to ensure that conditions 1–3 are met.

The guiding idea behind the four conditions is to convert private health plan or public agency decisions into part of a larger public deliberation about how to use limited resources to protect fairly the health of a population with varied needs. The broader public deliberation envisioned here is not necessarily an organized democratic procedure, though it could include the deliberation underlying public regulation of the healthcare system. Rather, it may take place in various forms in an array of institutions, spilling over into legislative politics only under some circumstances. Meeting these conditions also serves an educative function. The public is made familiar with the need for limits and appropriate ways to reason about them.

The first condition requires that rationales for decisions be publicly accessible to all affected by them. One American health plan, for example, decided to cover growth hormone treatment, but only for children who are growth hormone deficient or who have Turner’s syndrome. It deliberated carefully and clearly about the reasons for its decision. These included the lack of evidence of efficacy or good risk/benefit ratios for other groups of patients, and a commitment to restrict coverage to the treatment of disease and disability (as opposed to enhancements). It did not, however, state these reasons in its medical director’s letter to clinicians or in support materials used in “shared decision making” with patients and families about the procedure. It’s reasons were defensible ones aimed at a public good all can understand and see as relevant, the provision of effective and safe treatment to a defined population under resource constraints. The restriction to treatment rather than enhancement requires a moral argument, however, and remains a point about which reasonable people can disagree, as we saw earlier.

One important effect of making public the reasons for coverage decisions is that, over time, the pattern of such decisions will resemble a type of “case law.” A body of case law establishes the presumption that if some individuals have been treated one way because they fall under a reasonable interpretation of the relevant principles, then similar individuals should be treated the same way in subsequent cases. In effect, the institution generating the case law is saying, “We deliberate carefully about hard cases and have good reasons for doing what we have done, and we continue to stand by our reasons in our commitment to act consistently with past practices.” To rebut this presumption requires showing either that the new case differs in relevant and important ways from the earlier one, justifying different treatment, or that there are good grounds for rejecting the reasons or principles embodied in the earlier case. Case law does not imply past infallibility, but it does imply giving careful consideration to why earlier decision makers made the choices they did. It involves a form of institutional reflective equilibrium, a commitment to both transparency and coherence in the giving of reasons.

The benefits of publicity in the form of case law are both internal and external to the decision-making institution. The quality of decision making improves if reasons must be articulated. Fairness improves over time, both formally, since like cases are treated similarly, and substantively, since there is systematic evaluation of reasons. To the extent that we are then better able to discover flaws in our moral reasoning, we are more likely to arrive at fair decisions.

---

7. The conditions described were developed independently but fit reasonably well with the framework of principles for democratic deliberation developed by Gutmann and Thompson (1996). For some reservations about their approach, see Daniels (1999).
decisions. Over time, people will understand better the moral commitments of the institutions making these decisions.

The Relevance Condition imposes two important constraints on the rationales that are made publicly accessible. Specifically, the rationales for coverage decisions should aim to provide (a) a reasonable construal of (b) how the organization (or society) should provide "value for money" in meeting the varied health needs of a defined population under reasonable resource limits. Both constraints need explanation.

We may think of the goal of meeting the varied needs of the population of patients under reasonable resource constraints as a characterization of the common or public good pursued by all engaged in the enterprise of delivering and receiving this care. Reasoning about that goal must also meet certain conditions. Specifically, a construal of the goal will be "reasonable" only if it appeals to reasons (evidence, values, and principles) that are accepted as relevant by "fair-minded" people. By "fair-minded" I mean people who seek mutually justifiable terms of cooperation. The notion is not mysterious, since we encounter it all the time in sports. Fair-minded people are those who want to play by agreed-upon rules in a sport and prefer rules that are designed to bring out the best in that game. Here we are concerned with the game of delivering healthcare that meets population needs in a fair way.

Recall the restriction on the use of growth hormone treatment to those with growth hormone deficiency. As I noted earlier, some object that a theory that emphasizes protecting equal opportunity, as mine does, should also use medical interventions to eliminate extreme but normal shortness if it is disadvantaging. Still, proponents on both sides of this dispute can recognize that reasonable people might disagree about the specific requirements of a principle protecting opportunity. Both sides of the dispute about the scope of the goals of medicine nevertheless must recognize the relevance and appropriateness of the kind of reason offered by the other, even if they disagree with the interpretation of the principle or the applications to which it is put.

Consider further the implications of the Relevance Condition. "Including this treatment benefits me (and other patients like me)," just like "excluding this treatment disadvantages me (or other patients like me)," is not the kind of reason that meets the constraints on reasons. Because comparative coverage decisions always advantage some and disadvantage others, mere advantage or disadvantage is not a relevant reason in debates about coverage. If, however, a coverage decision disadvantages me compared to other patients similar to me in all relevant ways, then this is reason based on disadvantage that all must think is relevant. Also, if a coverage decision disadvantages someone (and others like him or her) more than anyone need be disadvantaged under alternatives available, then this too is a reason that all must consider relevant.

How should we view the claim that a treatment "costs too much"? First, suppose this is a claim about relative cost-effectiveness or worthiness. People who share in the goal of meeting the varied medical needs of a population covered by limited resources would consider relevant the claim that a particular technology falls below some defensible threshold of cost-effectiveness or relative cost-worthiness. Suppose, however, the claim that something "costs too much" refers to its effects on profits or competitiveness. Supporting this claim often requires providing information that private health plans will not reveal (for good business reasons), often turns on economic and strategic judgments requiring special experience and training, and ultimately depends on a much more fundamental claim about the design of the system, namely, that a system involving competition in this sort of market will produce efficiencies that work to the advantage of all who have medical needs. My point is not that these reasons fail to meet the Relevance Condition, but that providing support for them requires information that is often not available, that is hard to understand when it is available, and that ultimately depends on fundamental moral and political judgments about the feasibility of quite different alternative systems for delivering healthcare. Nevertheless, if for-profit health plans are to comply with the Relevance Condition, they must either be willing to provide information they would ordinarily not make public, or make their decisions on the basis of reasons that they can defend to other relevant stakeholders.

The constraints here imposed on reasons have a bearing on a philosophical debate about the legitimacy of democratic procedures. An aggregative or proceduralist conception of democratic voting sees it as a way of aggregating preferences. Where, however, we are concerned with fundamental differences in values, not mere preferences, an aggregative view seems inadequate. It seems insensitive to how we ideally would like to resolve moral dis-
Debate, namely through argument and deliberation. An alternative “deliberative” view imposes constraints on the kinds of reasons that can play a role in deliberation. Not just any preferences will do. Reasons must reflect the fact that all parties to a decision are viewed as seeking terms of fair cooperation that they accept as reasonable. Even if we have to rely on a majority vote to settle a disagreement where there are serious moral issues involved, if the reasons are constrained to those all must view as relevant, then the minority can at least assure itself that the preference of the majority rests on the kind of reason that even the minority must acknowledge appropriately plays a role in deliberation. The majority does not exercise brute power of preference but is instead constrained by having to seek reasons for its view that are justifiable to all who seek mutually justifiable terms of cooperation.

The Appeals and Enforcement Conditions involve mechanisms that go beyond the publicity requirements of the first two conditions. When patients or clinicians use these procedures to challenge a decision, and the results of the challenge lead effectively to reconsideration of the decision on its merits, the decision-making process is made iterative in a way that broadens the input of information and argument. Parties that were excluded from the decision-making process, and whose views may not have been clearly heard or understood, find a voice, even if after the original fact. The dispute resolution mechanisms do not empower enrollees or clinicians to play a direct, participatory role in the actual decision-making bodies, but that does not happen in many public democratic processes either. Still, it does empower them to play a more effective role in the larger social deliberation about the issues, including in those public institutions that can play a role in regulating private health plans or otherwise constraining their acts. The mechanisms we describe thus play a role in assuring broader accountability of private organizations to those who are affected by limit-setting decisions. The arrangements required by the four conditions provide connective tissue to, not a replacement for, broader democratic processes that ultimately have authority and responsibility for guaranteeing the fairness of limit setting decisions.

Together these conditions hold institutions—public or private—and decision makers in them “accountable for the reasonableness” of the limits they set. All must engage in a process of establishing their credentials for fair decision making about such fundamental matters every time they make such a decision. Whether in public or mixed systems, establishing the accountability of decision makers to those affected by their decisions is the only way to show, over time, that arguably fair decisions are being made and that those making them have established a procedure we should view as legitimate. This is not to say that public participation is an essential ingredient of the process in either public or mixed systems, but the accountability to the public in both cases is necessary to facilitate broader democratic processes that regulate the system.

In many public systems the reasoning that lies behind decisions that affect the length of queues—a rationing device—are inscrutable to the public. They are made in a “black box” of budgetary decisions. Queues may then be adjusted if the public complains too much—there is this kind of accountability to the squeaky wheel. But there is in general too little accountability of the sort demanded by the four conditions I describe (Ham and Packard 1998; Coulter and Ham 2000). Only through such accountability and the way in which it facilitates or enables a broader social deliberation will there be a wider perception that rationing decisions are fair and are made through an exercise of legitimate authority.

One issue facing this “process” approach to rationing seem to be more problematic in public systems than it does in mixed ones. In a mixed system two different insurers or health plans might arrive at different judgments about what limits to set. I have suggested both might be “right” if their decisions are the results of fair procedures (Daniels and Sabin 1998b). The anomaly is that some patients will then have access to services that others will not have, and this might seem to violate a formal constraint on fairness, that society treat like cases similarly. In a mixed system we might see this variation as a price we pay for whatever virtues (if any) the mixed system brings (the variation might ultimately lead us to better decisions over time). In a public system, however, such variation, e.g., between districts, might seem more objectionable if all are governed by the same public legislation and funding. Still, despite such anomalies, fair process may be the best we can do wherever we have no prior consensus on fair outcomes.

**Concluding Remarks**

A comprehensive approach to justice, health and healthcare must address all three questions I have
discussed. My extension of Rawls’s theory of justice to health and healthcare provides a way to link answers to the first and second questions. There are three ways in which Rawls’s theory also provides support for my approach to the third question. First, I propose that we use a fair process to arrive at what is fair in rationing, since we lack prior consensus on the relevant distributive principles. This retreat to procedural justice is at the heart of Rawls’s own invocation of his version of a social contract. Second, Rawls places great emphasis on the importance of publicity as a constraint on theories of justice. Principles of justice and the grounds for them must be publicly acknowledged. This constraint is central to the conditions that establish accountability for reasonableness. Finally, Rawls develops the view that “public reason” must constrain the content of public deliberations and decisions about fundamental matters of justice, avoiding special considerations that might be elements of the comprehensive moral views that people hold (Rawls 1993). Accountability for reasonableness pushes decision makers toward finding reasons all can agree are relevant to the goals of cooperative health delivery schemes. In this way accountability for reasonableness promotes the democratic deliberation that Rawls also advocates.

In pointing out these connections, I am not suggesting that this is the only approach to developing a theory of justice that applies to all aspects of health and healthcare. Indeed, I have pointed to other theories that converge in practice and to some extent in theory with the approach adopted here. I am proposing that concerns about justice and fairness in health policy should look to political philosophy for guidance and that some specific guidance is forthcoming. At the very same time, seeing how we have to modify and refine work in political philosophy if it is to apply to real issues in the world suggests that we should abandon the unidirectional implications of the term “applied ethics” or “applied political philosophy” (Daniels 1999b).

Acknowledgments

This article appears by agreement with and courtesy of Oxford University Press (New York, New York), and will appear in Medicine and Social Justice, edited by Rosamond Rhodes, Margaret Palst Battin, and Anita Silvers in 2002. Bruce Kennedy and Ichiro Kawachi collaborated with me to write the material on which I draw for the section, “Which health inequalities are unjust?” James Sabin collaborated on the research and writing on which I draw for the section, “When are limits to healthcare fair?” Research for this paper was supported by a Robert Wood Johnson Investigator Award and Tufts Sabbatical Leave. I have benefited from discussions with Richard Arneson, Allen Buchanan, Dan Brock, G. A. Cohen, Joshua Cohen, John Rawls, and Dan Wikler about the relationship between my approach and that of other recent work on egalitarianism.

References


———. 1993. Rationing Fairly: Programmatic consider-
inequality bad for our health
Daniels, N., B. Kennedy, I. Kawachi, and A. Sen, 2000.
Is coverage for new technologies. In Deliberative politics:
ed. S. Macedo, Pluralism, fair procedures, and legitimacy.
Institute of Medicine, 58±71. Washington: National Academy Press.


