Equity of Access to Health Care: Some Conceptual and Ethical Issues

NORMAN DANIELS

Department of Philosophy,
Tufts University

The literature on equity of access to health care is complex and confusing. Indeed, there is no consensus on what equitable access involves. There are three central reasons for divergence on this question. First, access is itself a complicated notion, composed of many factors. Consequently, determining what counts as equality of access, let alone equity of access, is a nontrivial problem; moreover, in some cases considerations about equity already play a role in our judgments about equality. Second, health care services are nonhomogeneous. They have many functions, some more important, more basic, or more urgent than others. So it is not possible to settle questions about equity of access until we have made it clear what the access is to. Are we worried about access to all the services offered in our health care system? Or are we worried only about a key set of services, defined by reference to some central or basic function, regardless of their availability within our system? And how can we pick these out in the context of an evolving system and technology? Third, and perhaps most fundamental, divergence on what to count as equitable access derives from divergence on more basic moral questions, specifically questions of distributive justice. There is moral disagreement about the nature of health care as a social
good, about what sort of special importance, if any, attaches to it. Moreover, disagreement on this question is tied to other fundamental disagreements about what distributions of social goods are just.

I shall organize my remarks as follows. In Section I, I shall briefly explain why the problem of defining equality of access is not trivial. This point runs counter to the view that we have a clear, noncontroversial notion of equality of access and that our only disagreements are about which departures from such equality are morally acceptable. In Section II, I shall discuss three approaches to the definition of equity of access. These include the utilization and needs-based account of Aday and Andersen (1980), a more "process"-oriented approach of some of their critics, for example, Sloan and Bentkover (1979), and a distinctively different "market"-based account, variants of which abound in the medical economics and planning literature. (I shall here include Enthoven, 1980.) I shall look at the strengths and weaknesses of each of these approaches and suggest ways in which their fundamental differences depend on broader disagreements about the nature of health care as a social good and about other principles of distributive justice. In Section III, I shall take up the question: to what must we have access? Specifically, I will look at the presuppositions about equity underlying reform proposals such as those advanced by Enthoven. These are important because they force us to consider the view that equitable access is access to a "decent basic minimum" of health care. I shall then sketch an account of a theory of distributive justice which avoids some of the problems facing the approaches to equity of access that we will have considered, though it leaves some issues unresolved. Finally, in Section IV, I shall comment briefly on the relationship between the demand for equity of access and other issues of distributive justice that arise in the assessment of a health care system.

A note on terminology is in order. In this paper, the term "equity" is used in a broad sense and is roughly equivalent to "distributively fair or just." There is also a narrower usage in which equity is an "interstitial" concept, one that raises questions of justice only against a background of institutions whose conformance with principles of justice is not at issue (cf. Barry, 1965:1522ff.). I use the broader notion, though I might prefer to restrict "equity" to the narrower one, because the literature on access to health care seems to employ the broader concept.
I. When Is Access Equal?

*Equal Income versus Equal Access*

It is tempting to think that we can give a completely noncontroversial definition of equal access to health care—much as we can do for equality of income—and reserve all controversy for debates about which departures from equality conform to acceptable principles of justice. Suppose I earn $10,000 less than you. The inequality might be thought *equitable* by some if you work longer or harder than I; by others, if your skills have a higher market value than mine; and by others, if you need more than I do. Here our moral disagreements about appropriate distributive principles show up as disagreements about just or *equitable* income distribution, though there is no controversy about whether or not our incomes are equal. The situation is arguably different for the notion of equal access: to arrive at a notion of equal access, we must already have made various decisions about what kinds of considerations ought to count in judging when access is equal. These decisions reflect our purpose or interest in making the judgment about equality, and some of these discriminations are themselves of a moral nature. Moral considerations, then, are already included in the specification of equality and are not held at bay until we get to decisions about equity.

Consider the problem in a slightly different context. There is a coffee supply in a lounge not far from my office. When is access to the coffee equal among my colleagues? Some cases seem clear: if the lounge is open only to male colleagues, then female colleagues can complain they do not have equal access to the coffee. If the lounge is up a flight of stairs and there is no wheelchair ramp, then my paraplegic colleague may have ground for claiming unequal access to the coffee. After all, he has to ask someone to fetch it, but none of the rest of us do. (Does his complaint disappear if in fact he drinks as much coffee as we do?) Other factors have a less clear impact. Should we worry about the fact that not all offices are equidistant from the pot? Some are thirty feet away, some only ten. Does it matter how the offices were assigned, randomly, by choice, or by seniority? Does it matter if the distance correlates negatively with coffee use? Should we worry that some colleagues use more calories in walking to the pot than others? Suppose the lounge is painted a
brilliant green, a color I so dislike that it takes more psychological effort for me to fetch coffee than it takes my colleagues, who chose the color. Is our access unequal? Suppose that on a previous job, I had unpleasant coffee room experiences and their memory hinders me from getting coffee; my colleagues had pleasant coffee room experiences. Do we have equal access?

I am not suggesting we spend much effort figuring out when access to the coffee is equal, for I do not want to trivialize the problem of access to health care. Still, it is worth seeing what underlies my inclination to say that the access to the coffee is equal despite variations in office distance, use of calories, preference for the color of the lounge, or past experiences that influence preferences. Where we have something that is merely an amenity (this is not an honest statement of my feelings first thing in the morning!), variables that merely affect preference orderings are properly ignored in judging equality of access. Factors like the "male only" lounge or the flight of steps have an impact in a way independent of effects on preference orderings. If I felt differently about the importance of coffee, because I thought it met a basic need (let us ignore habit or addiction), I might be more sensitive to factors that affect preferences and I might want to make finer discriminations, especially where there is an effect on utilization of the coffee pot.

My assessment of the importance of the coffee is connected to other views I have about which variables affecting access or use are relevant to issues of equity. That is, I allow my notion of equality of access to be determined, in part, by prior judgments about equity of access. Lest the case of the coffee pot be suspect because of its triviality, consider the same point in a nontrivial context. Thus, one tradition in our society is content to judge that equality of opportunity obtains if there are no formal—for example, legal or quasi-legal—barriers to persons of different races or sexes competing for a job or office. Others argue that positive steps must be taken to compensate for various natural and social variations between people which arguably confer "unfair," or at least undeserved, advantages. On this view, unless the impact of this "natural lottery" is compensated for, equality of opportunity does not obtain (Rawls, 1971:Sect. 14,17). Clearly, this difference in judgment about what counts as equality of opportunity is itself the product of other moral assessments, and so the notion of equal opportunity is itself controversial. Consequently, it cannot
serve (without further argument) as a noncontroversial baseline which we can use in debating the equity of inequalities in opportunity. The case is the same, I am suggesting, for equal access to health care.

Equal Access: Moral Complaint or Moral Ideal?

A related point is worth making before beginning our discussion of various approaches to defining equitable access. In a broad range of settings, individuals and legislators talk as if there is agreement on what would count as equal access. Thus, we find the remark in the Health Planning Act of 1974 (P.L. 93-641) that it is an important objective of federal policy to provide “equal access to quality care at a reasonable cost.” It is probably fair to say that all that most people have in mind when they talk about equal access is a negative criterion, specifically that certain traditional constraints on access, mainly financial and geographical, should play a minimal role in determining whether people who need health care get it. There may be implicit in this negative characterization a positive ideal: for example, “any two persons of comparable health status who want appropriate care have an equal chance of getting it.” But nothing so schematic may be in anyone’s mind at all; there may only be a moral complaint against a particular inequality. This point might be fruitfully expanded by a careful look at the history of our concern for equal access (cf. Starr, 1981).

II. Three Accounts of Equitable Access

Equity as Utilization for Need

The Aday and Andersen Approach as an Analytic Framework. I turn now to examine the ethical presuppositions and implications of three accounts of equitable access, a use- (or use-per-need) based account, a “process” variable account, and what I shall call a “market” account. The framework provided by Aday (1975), Aday and Andersen (1974, 1980) and Andersen et al. (1975, 1976) is a useful place to start. Not only is their work seminal, but seeing what motivates their
project will allow an economical discussion of the alternative approaches. I stress that my goal here is to point to ethical implications of the work, not to undertake a systematic, empirical assessment of it. Moreover, I must confine myself to the central thrust of each approach and not the subtle ways in which the accounts may be refined, extended, or combined. My central point will be to show the way in which accounts of equitable access are really disguised ways of talking about principles of distributive justice for health care.

Aday and Andersen’s (1980) work has both an analytic and normative importance. To see its analytic function, consider the fact that alternative theories of the behaviors or processes through which people seek health care suggest many factors which could have an effect on the use of health care services. We need a way of testing to find out which factors actually do have a significant effect on access. These potential access factors are of two, and by some classifications, three kinds. Some are structural features of the health care system, such as the availability of physicians or hospitals in different geographical areas, as measured, say, by physician-patient ratios. Others are features of individuals in the population including predisposing factors, such as age, health status, and cultural background, and enabling factors, such as income or insurance coverage levels. A particularly important subgroup of factors is sometimes viewed separately and called “process” factors. They play an important role in the process of seeking health care, at least in some theories of the process. The analytic task is to decide which variations in these many factors influencing potential access are important. Which ones count because they really contribute to differences in access to health care?

Aday and Andersen seek to test the importance of potential access variables by determining their effect on the measure of actual or realized access, the output of the system. Measures of realized access are of two main types: subjective measures concern satisfaction with treatment; objective measures are concerned with various utilization rates. The overall strategy is thus to look at variations between population subgroups on some measure of potential access to see if there is an effect on realized access, here utilization rates. Suppose our working theory of the process through which we seek health care suggests that a certain variable—say waiting time for an appointment—will have a big effect on access to health care. Aday and
Please ___ Enter my subscription
      ___ Renew
Subscriptions for complete
sets only

___ Students, Retired, $15.00
___ Individuals, $20.00
___ Institutions, $42.00
Purchase order # ____________________

Outside USA and Canada please add $4.00
for surface rate or $18.00 for air mail.

NAME ________________________________
ADDRESS ________________________________
CITY______ STATE_______ ZIP_____

PREPAYMENT IS REQUIRED
___ Payment enclosed
___ Bill me

___ Charge to my credit card:
      ___ Master Charge or ___ VISA
             (Bank Americard)

Card # ____________________________
( ) ( ) ( ) Master Charge only
bank digits

______________________ Signature

______________________ Card expiration date
Andersen propose a method for "testing" such claims: Potential access variables are important only if they produce an intergroup effect on realized access (utilization rates). If a process variable—say waiting time for an appointment—does not have an effect on utilization rates for population subgroups, then it is not causally significant according to the criterion being proposed. So the utilization rate test is a way of saying that the proof of the pudding is in the eating; the test of access is use.

Of course, details become important here. A process variable—for example, time spent in a waiting room—may have a significant effect on measures of satisfaction with care, but has relatively little effect on utilization rates. So the choice of objective or subjective measures of realized access may yield different assessments of the importance of a process variable and, ultimately, of the equity of access to health care services. In contrast to Aday and Andersen's primary emphasis on utilization rates, some critics, who are particularly interested in process factors, tend to emphasize satisfaction measures (cf. Sloan and Bentkover, 1979). Even if we are inclined to use an objective measure, however, it matters which one. For example, utilization rates between income groups do not seem to vary significantly following the introduction of Medicare and Medicaid. But if health status of lower income groups is lower than that of higher income groups, there may still be variation in a measure of use per need, say use per disability days.

Aday (1975) and Aday and Andersen (1980) urge use of a measure of utilization per need. Even here, detail matters. Different measures of need, say a more complex measure than disability days, which Aday and Andersen use, might give different utilization-per-need rates and thus a different assessment of which variables are "important" in the sense so far discussed. For example, if some income or cultural groups inflate their disability days, say because of different attitudes toward adopting a sick role or different absenteeism incentives (cf. Sloan and Bentkover, 1979:3), we may find that measure of need to be problematic for the analytic task at hand. Similarly, Davis et al. (1981) suggest that the choice of bed-disability days as a measure of need reveals a difference in use per need between high- and low-income groups which is not apparent with the less severe measure of need, disability days.
It is important to see that the focus on use per need rates stops short of measuring possible differences in the efficacy of the services delivered to different population groups. It is at least conceivable that use-per-need rates could be similar, but the quality—here, efficacy—of the services might vary, so that use per need would not effectively measure impact on health status. This point bears on the relevance of the focus on use-per-need rates and on the ultimate rationale for the definition of equity based on use.

**Normative Use of the Account.** It is crucial to see that the framework Aday and Andersen advocate is advanced primarily for its normative implications, not merely for its relevance to understanding or describing the causal relationships among the phenomena involved in access to health care. Indeed, we find the analytic test for the causal importance of a potential access variable reformulated as a definition of equity of access. Aday and Andersen argue: "The greatest ‘equity’ of access is said to exist when need, rather than structural (for example, availability of physicians), or individual (for example, family income) factors determine who gains entry to the health care system" (1980:26; Andersen et al., 1975:10–11). In other words, access is equitable if the important potential access variables, as operationally defined, are all related to health status in the proper way. If, however, important potential access variables are not related to health status, then an inequity of access obtains. As Aday and Andersen (1980:43) put it:

Inequity in health service distribution occurs when individuals receive services primarily according to their place in the social structure, their enabling characteristics, or the characteristics of the health system instead of according to their need . . . The inequity may be only “apparent,” however, if the lower access levels can be explained by varying age structure (demographic or illness levels). For example, lower use rates for an ethnic group consisting of predominantly younger persons in reasonably good health may be only “apparent inequity.” Excess utilization is indicated if a subgroup has higher utilization rates than the rest of the population. However, like inequity, it may be only “apparent excess.” One example would be higher utilization rates by the elderly which could be attributed to the larger number of symptoms and disability days they experience.

This normative account of equity of access is extremely influential. Davis et al. (1981), though they point to some “inequities” between
subgroups that Aday and Andersen (1980, 1981) fail to find, nevertheless appeal to the underlying Aday and Andersen definition. The difference lies in how the estimate of need is measured or operationalized.

The Argument from Function. The attempt to recast the utilization-rate test for important potential access variables as a normative criterion for equity of access needs justification. Why ought we to look at variations in utilization rates in this way? Here Aday and Andersen draw on a widely held view, namely that an "'equitable distribution' of health care services is one in which illness (as defined by the patient and his family or by health care professionals) is the major determinant of the allocation of resources" (1980:41).

One common way to defend this view involves an argument from function which goes something like this:

The (main) function of health care services is to prevent and cure illness, i.e., to meet health care needs. A distribution of health care services that is not determined by the distribution of health care needs is therefore unreasonable in some important sense. Specifically, it ignores similarities and differences—in health status—between persons that, given the function of health care, ought to be relevant to establishing its reasonable distribution. Ignoring such relevant similarities and differences makes a distribution inequitable.

Something like this argument is clearly foreshadowed in Bernard Williams's now classic discussion of equality in which he concluded that "leaving aside preventive medicine, the proper ground of distribution of medical care is health care; this is a necessary truth" (Williams, 1962:27). In any case, of course, the argument at most establishes a necessary but not a sufficient condition for equitable distributions. Moreover, some would argue (as we shall see) that it is not even a necessary condition for equity that health care be distributed according to needs for health care. For example, some might object that the argument mistakenly presupposes that health care services are homogeneous in function and that people will always have preferences that correspond to their presumed health care "needs."

This underlying view, resting as it seems to do on an argument from function, is a central and powerful one. It is a view with a long history of advocacy. I shall come later to sketch a position which incorporates those aspects of it I think correct. Still, we can see it is not the whole of the story, if we examine more carefully some objections to the Aday and Andersen approach.
Objections to the Utilization for Need Account. One central objection is that a focus on utilization rates ignores at least one other necessary condition for equity of access. Specifically, variations in certain potential access variables, especially process variables, can have equity implications even if they do not show up as important variables on the Aday and Andersen criterion, that is, by reference to their effects on utilization rates. Time spent in a waiting room, or out-of-pocket health care expenditures, if they vary with income group (corrected for health status), are differentially burdensome even if utilization rates (or use-per-need rates) are not affected. More generally, someone may argue that some inequalities in potential access variables between subgroups raise equity questions even if they do not affect the outcomes of health care seeking behavior, as measured by utilization rates. They still affect other outcomes: for example, what else someone has the chance to do with his time or money.

A second objection is that subgroup uniformity in utilization rates (or use-per-need rates) is not even a necessary condition for equitable access. Some subgroup variations that correlate with utilization rate differences will reflect differences in attitudes toward health care. Consequently, these variations may not be inequities in the access to health care at all. A standard example might be the deliberate under-utilizer who, for religious, esthetic, or cultural reasons has a principled aversion to some or all traditional (mainstream) health care services. Another form the problem may take has already been mentioned. Attitudes toward assuming a sick role may differ between cultural or income groups, so that some judge themselves to have a disability day more readily than others. But if one group inflates its needs in this way relative to another, the equity of the distribution is affected. (Cf. Sloan and Bentkover, 1979:3.) Of course, we need to draw finer distinctions, say between those attitudes toward health care based on ignorance, which society has a responsibility to correct, and those based on principled, informed choice. But clearly, some modification of the Aday and Andersen account is needed to accommodate these worries. In any case, they clearly leave room for such a modification.

The issue is even more complex, however. It may be that some variation in utilization rates, even corrected for health status, is not importantly related to health outcomes, that is to health status after treatment. It is at least arguable that only those utilization rate variations are inequitable which reflect significant differences in the preventative, curing, and caring functions of health care services. Aday
and Andersen shy away from looking at health outcomes because so many factors intervene between utilization and outcome; but their simplification here may leave room for a systematic bias. The kind of case of greatest interest is one in which one treatment is more intensive than another but where there is no evidence that it is also more efficacious. The issue arises, for example, in the context of mental health care, where some evidence points to the use of more intensive interventions (therapy) for upper-middle-class groups and low intensity (more drug oriented) treatments for low-income and minority groups (cf. Mollica and Redlich, 1980). If there is no demonstrable difference in efficacy between the two types of treatment, does the inequality in utilization rate by income group constitute an inequity? Finally, one might insist that some differences in utilization rates, even where there are effects on health status outcomes, reflect informed choices about how risk-free one wants to be. They reflect a choice about how important one thinks health care services are compared to other things on which one may want to spend income (cf. Fried, 1978). Such choices may well lead to utilization-rate differences, but they still do not indicate inequity of access.

It is now possible to explain what I am calling the "process" and the "market" accounts of equity of access in terms of the possible objections to the utilization rate account that we have just noted. Minimally a process account seems committed to the view that the utilization rate account captures at most one necessary condition for equitable access and, in any case, fails to capture another necessary component of equity, intergroup equality in process variables. Indeed, the process account may even drop intergroup uniformity in use-per-need rates as a necessary condition for equitable access. In this case, the process account clearly overlaps what I am calling the "market" approach, for the latter insists that uniformity of use per need is not even a necessary condition for achieving equitable access. Of course, the market approach parts company with the process account in that it also rejects the view that intergroup variations in process variables constitute inequity of access.

**Equity as Equality in Process Variables**

*Sloan's and Bentkover's Account.* Consider as an example of the "process" approach Sloan's and Bentkover's (1979) fine study of access to ambulatory care. Their view is not so much that one should ignore
utilization rates, which they admit are an important measure of realized access. Rather, they object to its selection as the sole or even primary criterion in an account of equitable access (1979:2–3). Their concern for the way in which certain process variables—for example, travel or waiting time—may vary with income group, race, or geographical area is a concern for what they refer to as the "humaneness" of the care delivered (1979:4). Presumably, such variations are likely to be captured more by subjective (satisfaction) measures of realized access even when they do not affect objective (utilization rate) measures. Process variable differences, even where they do not affect utilization rates, may reflect differences in the difficulty of seeking care, that is, inequalities in the burdens that attend seeking care. As Sloan and Bentkover put it: "Many, for example, would view the long waits the poor experience in clinics as an injustice, irrespective of the effect patient waiting might have on utilization rates" (1979:24). Gutmann, 1981, seems to suggest a similar view. The basic contention is that access to health care cannot be considered equitable if it is much more difficult for some people to get care than it is for others, even if people make adjustments to the burdensomeness of the process and get the amount of care they need.

Amenities, Quality, and Claims to Access. There is a certain plausibility to this worry about the burdensomeness of the process of seeking care and the claim that considerations of equity are raised by such differences. What is missing, however, is the kind of moral argument we saw was immediately forthcoming in support of the utilization rate approach. What we need to know is why ease of access must be roughly equal for population subgroups even when the "ease" factor does not affect utilization rates. The problem is made more glaring by a formulation of the issue which can be found in Sloan and Bentkover themselves. Specifically, they tend to group the impact of these process variables under the heading of "quality" factors in health care. Indeed, within their economic model, the term "amenities" is used (1979:24–25). The picture that emerges is that health care for some population subgroups may have more "amenities" than for other groups—for example, less travel or waiting time, more physician contact time—though not necessarily higher utilization rates or better health care outcomes. Do these subgroup differences in "amenities" constitute inequities of access? If the "process" account of equitable access can be construed as asking for a more egalitarian distribution
of these amenities, and not just of utilization-per-need rates, what justifies the demand?

The problem here is that only some qualitative aspects of health care services clearly seem directly relevant to worries about equity of access. It is worth noting in this connection an important simplifying assumption that underlies the utilization rate account we have looked at. Utilization rates tell us nothing about the quality of the services rendered, where quality is some measure of net benefits minus harms (cf. Donabedian, 1979 and Daniels, 1981b). Such a measure of quality is largely concerned with the efficacy of services rendered. The simplifying assumption is that when use-per-need rates are roughly equal, we are dealing with qualitatively equal sorts of services, in the sense that their impact on health outcomes is likely to be roughly equal. But the “amenities” involved with many of these process variables, when they are viewed as qualitative features of the health services rendered, are not clearly linked to health status outcomes—only to subjective measures of satisfaction with the treatment. So an argument that grants health care services a very special status because of their primary function of meeting health care needs does not by itself seem powerful enough to justify the concern, present in the process account, that the equal distribution of amenities is also a necessary condition for equity of access.

A more promising line of argument for the process account might go something like this. In order to be sure that variations in use-per-need rates are in fact the result of informed choices or preferences about the use of health care services, we must be sure that decisions to utilize are not made harder for some persons than others because of variations in the process variables. But the force of this argument is not obvious in the face of evidence that use-per-need rates are equitably distributed, because it still must be argued that we have a difference here that makes a difference. Still, variations in process variables are a useful focus of concern when an attempt is made to explain away a variation in use-per-need rates—for example, by saying they are merely the result of differences in preference. Such an argument puts the process account more in the role of an opponent of the market view than an opponent of the use-per-need account.

Of course, strongly egalitarian views about distributive justice in general might be invoked to justify a concern about equality among process variables. If, for example, one were prepared to argue that
only differences in need or preference should be allowed to explain variations in the services used by different groups, then systematic variations in even "amenities" would look like inequities, just as unequal distributions of any goods that cannot be fully accounted for by need or preference differences constitute inequities in distribution. Or if one were to allow inequalities only if they act to maximize the well-being of the worst-off, then some variations in "amenities" might count as inequities. But I am not concerned to discuss such strongly egalitarian views here, largely because I am interested in the special arguments people are inclined to make about equity with regard to health services which they are not inclined to make for many other social goods. Still, more specialized arguments may be invoked here. Dickman (1981) argues that a principle requiring us to show "equal respect for persons"—plus the fact that people in need of health care are especially vulnerable to affronts to their self-respect—requires that we pay special attention to the roughly equal distribution of at least some reasonable set of "amenities." Interesting though the argument is (also cf. Jonsen, 1976), I cannot consider it here.

**Equity as the Market Availability of a Decent Basic Minimum**

*Constraints on the Market.* I should like now to sketch in more detail what I have called the "market" approach to equity of access. In contrast to the utilization rate and process variable approaches, the market approach is not really a position represented in the empirical literature on access. Rather, it is a composite abstracted from views which are common in economic and health planning literature. It is of interest here because of the quite different limits it places on the notion of equitable access and because of its quite different underlying view of health care and distributive justice. Nevertheless, as with the utilization-rate approach in particular, an underlying approach to issues of distributive justice plays a prominent role in defining what counts as an equitable access.

I have already noted that one common line of objection to the utilization-rate approach is that similarity in intergroup utilization-per-need rates is not even a necessary condition for equitable access (or distribution). A view that provides a rationale for such a claim
is the view that health care services are commodities like any others. On this view, there is nothing so “special” about these services that cannot be accommodated by allowing markets to respond to people’s preferences. Thus, equity of access is assured if three main conditions obtain: first, the commodity must be available at something like “true social cost”; second, individuals are capable of making rational informed decisions about using the system; third, income distribution must be approximately equitable. The second condition requires that information about alternatives—for example, therapies or insurance schemes—is available and that people are competent and informed enough to make use of such information. Some access inequities arise when this condition is not met and these must be addressed by public policy. But I will say nothing about them here.

Aside from the problem of subsidies to the poor to guarantee equitable income distribution, the central problems of access are those brought about by departures of the medical market from the ideal of a truly competitive market (cf. Arrow, 1963). In particular, there may be various distortions on the supply side which amount to the market not delivering services at their “true social cost.” For example, some groups—rural populations or inner city minorities—may not be able to get the care they want and can pay for. They may not be able to get it in the desired quantities, or at the desired times, with the characteristics they desire. Viewed in this way, the problem is that the market is unresponsive to consumer preferences on the supply side, and interventions may be needed to correct the problem, generally by addressing structural problems—for example, obstacles to manpower supply, or problems in capital expenditure policy. A central problem here is the way in which the choice of a health insurance plan is tied to features of employment and the unavailability of an adequate range of plans—for example, ones that cover people between jobs. One structural feature of the insurance market is the relative unavailability of prospective per capita, rather than fee-for-service, schemes; this feature is a central focus of criticism by a number of planners (cf. Havighurst, 1971 and Enthoven, 1980). The central issues of access and equity of access are concerned with these supply malfunctions of the market.

The third condition, equitable income distribution, usually requires only that no one fall below the officially defined poverty line. The
assumption is that one can buy a subsistence level of basic social goods (food, housing, health care) and that transfers should make sure every one can. Moreover, it is usually insisted that the sum of cash and aid-in-kind benefits to the poor not produce work disincentives. Where this ceiling is not argued for just on efficiency grounds, it is also claimed that an inequity would result if employed workers were less well-off than the unemployed poor. Much could be said about the adequacy of these views of equitable income distribution, but this is not the occasion (cf. Brown et al., 1981; especially Daniels, 1981c).

One real issue that concerns us here is how to characterize the cash, voucher, or aid-in-kind transfer needed to meet these assumptions. The position that seems to be held in common—either explicitly or implicitly—by many “market” proponents is that the transfer must be adequate to buy a “decency basic minimum” of health care. If the transfer falls short of this, it is agreed we have an inequitable transfer. So the market view I am sketching is not that of the pure libertarian who might reject all such transfers, but rather one that marks an implicit acceptance of some important moral claims that might loosely be characterized as welfare rights. I shall examine the problems involved in characterizing the decent basic minimum in Section III, but in any case it seems clear that there are definite limits to the transfer needed in order to assure that equitable financial access to the medical market place is provided.

Implications of the Market Account. Assuring equitable access in the ways defined by the “market” approach leaves extensive room for all sorts of departures from equitable access as defined by either of the other two approaches we have considered. Surely, there may be variations in the “amenities” that accompany health care services, if that is how we want to look at some process variables. Equal quality in these dimensions is surely not required, just as not everyone “prefers” equal quality in automobiles. Similarly, utilization-per-need rates may vary with “suspect” variables, like income or race, and yet not indicate any inequity of access, contrary to the Aday and Andersen formulation. Rather, the unequal distribution of health care—in quantity and quality—is viewed merely as the expression of different preference curves, just as food budgets might vary among a welfare recipient, a factory worker, and a wealthy industrialist. If we take the underlying income distribution to be morally acceptable, its expression in terms of utilization of health services need indicate no inequity.
Put succinctly, then, the "market" approach I am considering here comes to this: access to health care is equitable if and only if there are no information barriers, financial barriers, or supply anomalies that prevent access to a "reasonable" or "decent basic minimum" of health care services. How plausible such an account is depends on the characterization of such a decent minimum and the moral arguments that provision of such a minimum is all that requirements of equity (justice) demand. The problem facing the "market" proponent thus appears to be the other side of the coin from the problem facing the utilization rate account. One central problem with that account was its simplifying assumption that health care was relatively homogeneous in function and that the proper basis for its distribution must be the realization of that function. If, however, we want to treat health care services as nonhomogeneous in function, and we are willing to ground equity claims only by reference to some features of some of those services, we must present an account of how to draw the lines.

A Pragmatic Remark

Before looking at some of the ethical issues involved in the attempt to define equity of access by reference to some decent minimum, it might be worth commenting on a more pragmatic approach. One might suggest that even if disagreements about equity of access are rooted in fundamental disagreements about distributive justice, there may still be points of agreement as well. Are there points of agreement among the different approaches about inequities in the existing health care system? Indeed, some might argue that matters of public policy in the face of fundamental disagreement must rest on principled compromise, and we should look solely at the points of convergence between them.

A careful empirical survey of the literature should readily isolate major points of convergence. I leave that task to someone more competent in assessing the empirical literature. Indeed, a careful reading of Enthoven (1980) suggests where some of those points of convergence lie.
III. Decent Minimums and the Requirements of Justice

Problems of Characterization

Lists versus Criteria. Earlier I noted that there is a basic question which must be answered before we can understand disagreements about access: access is always access to something, but to what? There is a tendency in the utilization and process variable accounts to assume that the answer must be "access to whatever range of services are available in the system." This answer ignores the nonhomogeneity of function of health care services and systems. In contrast, the market approach circumscribes the demands of equity with regard to access in a way which may avoid this objection. It insists that we are concerned only with access to a "decent basic minimum" of care. I want here to examine some of the problems with this notion.

What is meant by a "decent basic minimum"? There are three ways to elucidate the notion: 1) the provision of a general criterion by reference to which we can tell if services are among the minimum or are above it; 2) the description of a fair procedure for determining the minimum; or 3) simply listing the types of services included. In the market literature, indeed in much of the literature, there is little attempt to give a general criterion or describe an appropriate, fair procedure. What attempts we get are far too vague. Charles Fried (1976:32), for example, suggests the "decent minimum should reflect some conception of what constitutes tolerable life prospects in general. It should speak quite strongly to things like maternal and child health which set the terms under which individuals will compete and develop." There may be the nucleus for a helpful idea here, but it is not developed enough to tell us when prospects are tolerable. As John Arras (1981:32) asks, tolerable to whom?

More specific is the characterization that emerges from Enthoven's (1980) discussion of a Consumer Choice Health Plan. To qualify for tax credits, vouchers, or Medicare payments, an insurance plan would have to meet certain requirements. "A qualified plan would be required to cover, at a minimum, the list of services called 'basic health services' in the Health Maintenance Organization (HMO) Act of 1973 (as amended). This list includes physician services, inpatient and outpatient hospital services, emergency health services, short term out-
patient mental health services (up to twenty visits), treatment and referral for drug and alcohol abuse, laboratory, and X-ray, home health services, and certain preventive health services" (Enthoven, 1980:128). Enthoven adds, however, that "it might make sense to start the program with a less costly list" (1980:128), that, unfortunately, leaves the specification by list indeterminate, a point to which I shall return shortly. Qualified plans would be required to offer a low-option plan, consisting of just the basic services (or an acceptable subset of them), in addition to any higher-option plans they market.

If we recall our earlier analysis, equity of access is guaranteed on the market approach if there are no important information, financial, or structural obstacles to buying into a low-option plan. Enthoven suggests that structural barriers will be reduced if qualified plans are required to have "open enrollment" to all eligible in its service area and "community rating" to avoid division of qualified plans into high and low risk groups. Furthermore, breaking the connection between employment and type of insurance plan available will remove another anomaly of the insurance market and close an important gap in access. Enthoven assumes that using vouchers to enhance the purchasing power of rural areas, which have relatively poor populations, will improve the availability of services. One must ask, however, whether or not open enrollment is sufficient guarantee that plans are available to all in an area; should there be requirements on demographic mix or on option mixes (cf. Havighurst's [1971] idea of a 50% nonvoucher requirement on membership)? Similarly, we must know how well manpower and facilities will be disseminated geographically merely because vouchers increase purchasing power in underserved areas. Whether Enthoven's measures are adequate to eliminate structural problems affecting access I leave to a more empirically focused discussion. Instead, I return to the question of what equitable access is access to.

Low and High Option Plans. Consider again Enthoven's effort to specify the decent minimum by reference to a list. The list is open to emendation; not everything on the 1973 HMO Act list is mandatory. By virtue of what are things on the list in the first place? By virtue of what can we leave them off? Can we leave mental health coverage off? Dental care is already omitted. Why should things be on or off? It should be remembered that we just cannot determine the list by reference to average costs for actuarial categories. These
costs are merely the costs for types of services—for example, physician or hospitalization. When Enthoven (1980) uses the figure of $1350 for a family of four for the voucher available to a family with a maximal total income of $4200, the figure is based only on actuarial costs for physicians and hospitalization. So we have to know what is to be on the "decent basic minimum" list before we calculate the voucher. Unfortunately, we have been offered neither a principle nor a fair procedure for arriving at the list.

Consider now some ways in which low option plans might differ from high option plans. The indeterminacy of what must be included in the low option plan will of course show up as an equity of access question if more comprehensive plans include broad categories not included in basic plans—for example, mental or dental coverage. But comprehensiveness can vary even within categories. What surgical procedures are covered? What mental health therapies? What dental plans?

There are other important ways in which low and high option plans may vary. Suppose the extensiveness of diagnostic services is allowed to vary between lower and higher budget plans. Then we might imagine persons with similar health status being given less or more extensive diagnostic services. One way to conceptualize this variation is as a variation in quality, here taken to be a measure of the net health benefits minus burdens (cf. Donabedian, 1979). If we imagine that, under different budget ceilings, quality can be optimized in different ways, we return to our fundamental problem. Suppose, for example, one quality optimization, available under a high option plan, allows some greater degree of freedom from risk, say of risk of undiscovered cancer, than is available under the low option plan. We can imagine people thus “buying” a degree of freedom from risk, so that people who like to be relatively risk-free would buy the high option plan, while those who want to tolerate a greater risk would buy the low option plan.

Do we now have an equity of access problem? The decent basic minimum approach would, at least implicitly, be setting a degree of freedom from risk against which it is "decent" to be protected. Anything higher must be viewed as a matter of preference—a commodity to be floated in the market. But it is not clear that we have any such clear idea of a decent minimum. No doubt some of the variation here would be curtailed by malpractice litigation, but it is
also not obvious that this litigation is the proper forum in which to decide these matters of health policy. Nor is it obvious that we can appeal to "standards of practice" to resolve this question. Where such standards are based on good studies of efficacy and cost-effectiveness, we may resolve some issues. But where the standards are derived from a clinical practice setting in which reimbursement is generally on a fee-for-service basis, the issue of what standard is acceptable under different budget ceilings has not been squarely faced. Of course, where the differences in "quality" are primarily of the sort Sloan, Bentkover, and others worry about under the heading "amenities," we can expect the "decent minimum" standard to allow some care to be considerably less decent than others.

My criticisms of Enthoven's characterization of the decent basic minimum are not responsive to one defense he might make. He might say that the decent basic minimum must be defined relative to existing practices within the society, specifically those that lead to the average actuarial costs for the items on his list. The decent minimum is defined by reference to the average. How can the worst-off complain if they end up doing as well as the average? Of course, this response ignores the flexibility of the list itself; still, it rests on a healthy pragmatism. To see that an account might aim for a more principled characterization of what justice requires, we must turn to an alternative view.

Toward a Distributive Theory: The Fair Equality of Opportunity Account

Is Health Care "Special"? My suggestion that the notion of a "decent basic minimum" is inadequate to support the moral weight it bears in the "market" approach is best supported by the proposal of an alternative account. It is possible to give a more perspicuous, if still abstract, account of what equitable access should be access to. My account (cf. Daniels, 1981a) is an attempt to answer the question: what is so special or important about health care compared to other social goods? Many people in many societies believe it is especially important, for they often insist health care be more equally, actually equitably, distributed than various other social goods. What might explain this special importance?
We need to back up a bit and consider more carefully the function of health care. Such an analysis is what was missing in the use-per-need account. Suppose we adopt a rather narrow, if not uncontroversial, view of disease: diseases will be departures from normal species functioning. Health care needs, broadly construed, concern things we need to prevent, maintain, restore, or compensate for—departures from normal species functioning. Why are such departures from normal functioning of social importance? One initially plausible answer is that, whatever else we need or want, we need normal functioning—it is a necessary condition for happiness, say. But this answer seems less plausible when we note that happiness or satisfaction in life do not so clearly require normal functioning. Many people “cope” well with significant impairments.

A more plausible answer, I believe, is that normal species functioning is an important component of the opportunity range open to individuals in a society. The opportunity range is the array of life plans that it is reasonable to pursue within the conditions obtaining in a given society. This range is, of course, relative to various social facts about the society—its stage of technological development, material well-being, and so on. Thus, similar impairments of normal species functioning might have different effects on opportunity range in different societies. But within a society, it becomes possible to give at least a crude ranking to the effects of different impairments of normal functioning in terms of their effects on the normal opportunity range. In turn, this gives us a crude ranking of the importance of different health care needs. Moreover, on this account, some uses of health care services—for example, some cosmetic surgery or some kinds of counseling—do not meet health care needs, but only certain other wants and preferences.

I am suggesting that we can account for the special importance ascribed to health care needs by noting the connection between meeting those needs and the opportunity range open to individuals in a given society. This suggests that the principles of justice governing the distribution of health care should derive from our general principles of justice guaranteeing fair equality of opportunity (cf. John Rawls, 1971:Sect. 14). Specifically, health care institutions will be among a variety of basic institutions (for example, educational ones) which are important because they insure that conditions of fair equality of opportunity obtain. I cannot argue here the issues in the general
theory of justice that would support the view that fair equality of opportunity is a requirement of justice. But if I am granted the assumption that it is, we have the foundations for important social obligations in the distribution of health care. Moreover, a concern for fair equality of opportunity—in theory if rarely in practice—has a long historical tradition in this country.

There are, to be sure, worries with my approach. For example, the notion of opportunity has to be age-relativized or it seems to embody a significant age bias—like productivity measures of the value of life-saving technologies (cf. Daniels, 1981d). Similarly, I must show that these requirements of justice do not open a bottomless pit into which we are required to pour endless resources in quest of an unreachable egalitarian goal. But this is not the place to consider even such important details, and I have discussed them elsewhere (cf. Daniels, 1981a).

**Implications for Access.** The fair equality-of-opportunity account of distributive justice for health care has several important implications for the issue of equitable access we have been discussing. First, the account is compatible with, though it does not imply, a multi-tiered health care system. In contrast, the “market” approach requires at least a two-tier system. Thus, my account shares with the market approach the view that health care services serve a variety of functions, only some of which may give rise to social obligations to provide them. The basic tier in my account would include health care services that meet health care needs, or at least important health care needs—as judged by their impact on opportunity range. Other tiers, if they are allowed, might involve uses of health care services to meet less important health care needs or to meet other needs and wants. My account leaves open the possibility that other tiers of the system might also be important enough to be given special precedence over other uses of social resources; but if they are, it will be for reasons different from those which give such precedence to the basic tier.

Second, the fair equality-of-opportunity account provides a way of characterizing the health care services that fall in the socially guaranteed tier. They are the services needed to maintain, restore, or compensate for the loss of normal functioning. In turn, normal functioning constitutes a central component of the opportunity range open to individuals. This account is, to be sure, abstract. It requires moral judgment in its application. Still, it provides a principled basis for
argument about what is included in the basic tier, a basis we found lacking in the notion of a decent basic minimum and in Fried’s gloss on the notion of “tolerable life prospects.”

Third, whichever way the upper tiers of the health care system are to be financed, there should be no obstacles—financial, racial, geographical, and so on—to access to the basic tier. The importance of such equality of access follows, I think, from basic facts about the sociology and epistemology of the determination of health care needs. The “felt needs” of patients are at best only initial indicators of the presence of real health care needs. Structural and other process barriers to initial access—for example, to primary care—compel people to make their own determination of the importance of the symptoms they feel. Of course, every system requires some such assessment, but financial, geographical, and other process barriers (waiting time, for example) impose the burden for such assessment on particular groups of persons. Indeed, where it is felt that sociological and cultural barriers exist preventing people from utilizing services, positive steps are needed (in the schools or through relevant community organizations) to make certain that decisions are informed.

The Aday and Andersen approach may be helpful here. Their utilization-per-need criterion, or a refinement of it, gives us a way of telling when a potential access factor is likely to be affecting opportunity through its impact on utilization rates. Moreover, whereas their unqualified assumption about the homogeneity of health care was problematic for the health care system as a whole, it is not problematic in this context. Indeed, my account characterizes that function in a perspicuous way, enabling us to see why it has special moral importance. In addition, my account permits “suspect” variations in utilization-per-need rates to be explained away as informed choice where this is plausible. (Aday and Andersen also leave room for such modifying explanations.) In short, I think the account I offer takes what is reasonable from the argument from function which underlies the utilization-per-need-account and provides a clearer moral rationale for it.

Fourth, the fair equality-of-opportunity account remains silent on what to make of demands for strict equality in process variables (“amenities”), that is, independently of their effect on utilization-per-need rates. It also remains silent on equity of access requirements for the upper tiers, if such there be. It also needs to be carefully applied
if it is to answer the kinds of problems that I raised concerning the market approach with regard to variations in quality—that is, efficacy and protection from risk. These are not issues I am prepared to take a direct stand on here. Still, it is worth characterizing in general terms the kinds of arguments that might be brought to bear. The crude typology of arguments I will suggest at least tells us what kinds of considerations we should avoid conflating.

A Typology of Arguments about Equality. Arguments about equity concerning "suspect" variations in "amenities" or in quality (protection against risk, for example) fall into three main categories. The first kind of argument rejects the inequality on general grounds of distributive justice, independently of the fact that we are concerned with health care. A general distributive argument, for example, might suggest that income inequalities of the sort the market approach tolerates, which do not allow some people to buy extensive amenities or superior quality, are not justifiable. One need not be a strict egalitarian here. Even a principle that constrained inequalities in the way Rawls's "difference principle" does, so that inequalities must act to make the worst-off groups best off, might not allow the kinds of inequalities tolerable to the "market" approach. Though I am inclined to take a rather egalitarian stand on income distribution, I find, for our purposes here, such general distributive arguments are not as interesting as arguments which more specifically address problems about health care.

Arguments that are directly concerned with health care can be divided into two kinds. A primary health care argument is one that asserts all health care services are special in some way and that this specialness forces us to be egalitarian in ways not necessary for many other social goods. For example, as we noted when we discussed possible rationales for the "process" account of equity, someone might argue that there is a special connection between health care and self-respect. Consequently, a society fails to show equal respect for persons if it allows inequality of access to even nonbasic health care services (cf. Dickman, 1981). Or some might argue that much of our health care manpower, facilities, and technology has at one point or another been heavily subsidized by public funds; consequently all citizens deserve equal access to what society has so extensively funded. The latter argument, it should be pointed out, is inadequate to justify equal access to health care in societies that have not subsidized their...
health care systems. My own suspicion is that primary arguments about equal access to all health care service are not likely to succeed. For example, they may assume a homogeneity of function for health care services which is contrary to fact; or the property they pick out—for example, public funding—also characterizes many things or institutions where no comparable argument about equal access is advanced.

Secondary health care arguments may hold more promise, but they are likely to rest on far more complicated and disputed empirical claims. For example, a secondary argument might advance a distinction between basic and nonbasic categories of health care services, or between adequate and above-adequate levels of quality. Still, because the tiers or sectors that deliver both kinds or qualities of service are causally connected to each other, by allowing a market for the nonbasic level, we might threaten the possibility of delivering the basic level equitably. For example, a market approach to nonbasic services might undermine the quality or raise the cost of the basic tier through drains on manpower and competition for resources (cf. McCreadie, 1976). Of course, counter-arguments of the secondary type are possible too. For example, it has been argued that a market tier above the decent minimum promotes innovation, or that prohibiting such a tier will generate a black market (cf. Fried, 1976).

My sketch of a fair equality of opportunity approach thus leaves some issues unresolved. Still, it may point the way toward taking the best from the various approaches that we have seen it is otherwise difficult to reconcile.

IV. Non-Access Issues of Equity in Health Care

My central point throughout this discussion has been that different approaches to defining equity of access to health care ultimately depend on different underlying accounts of the kind of social good health care is, and on appeals to different principles of distributive justice. I briefly sketched an approach that I believe is an improvement over existing accounts, though it too leaves some problems unresolved. It is worth pointing out here, by way of conclusion, that issues of equity arise in other contexts than in disputes about access. We get
a very one-sided picture of what a theory of justice for health care requires if we concentrate solely on issues of access.

My account of health care needs and their connection to fair equality of opportunity has a number of implications for resource-allocation issues (cf. Daniels, 1981a, 1981b, 1982). I have already noted an important distinction between the use of health care services to meet health care needs and their use to meet other wants and preferences. The tie of health care needs to opportunity makes the former use special and important in a way not true of the latter. Moreover, we get a crude criterion—impact on normal opportunity range—for distinguishing the importance of different health care needs, though this falls far short of being a solution to many hard allocation questions. Three further implications are worth noting here.

There has been much debate about whether the United States health care system overemphasizes acute therapeutic services as opposed to preventive and public health measures. Sometimes the argument is focused on the relative efficacy and cost of preventive—as opposed to acute—services. My account suggests there is also an important issue of distributive justice here. Suppose a system is heavily weighted toward acute interventions, yet it provides equal access to its services. Thus anyone with severe respiratory ailments—black lung, brown lung, asbestosis, emphysema, and so on—is given adequate and comprehensive services as needed. Does the system meet the demands of justice? Not if they are determined by the approach of fair equality of opportunity. The point is that people are differentially at risk of contracting such diseases because of work and living conditions. Efficacy aside, preventive measures have distributive implications distinct from acute measures. The opportunity approach requires we attend to both, equity of access as well as equity in the distribution of risk (Daniels, 1981e).

My account points to another inequity in allocation. One important function of health care services, i.e., personal medical services, is to restore handicapping dysfunctions—for example, of vision, mobility, and so on. The medical goal is to cure the diseased organ or limb where possible. Where cure is impossible, we try to make function as normal as possible, through corrective lenses or prosthesis and rehabilitative therapy. But where restoration of function is beyond the ability of medicine per se, we begin to enter another area of services—nonmedical social supports. Such support services provide
the blind person with the closest he can get to the functional equivalent to vision—for example, he is taught how to navigate, provided with a seeing-eye dog, taught braille, and so on. From the point of view of their impact on opportunity, medical services and social support services that meet health care needs have the same rationale and are equally important. Yet for various reasons, probably having to do with the profitability and glamor of personal medical service and careers in them as compared to services for the handicapped, our society has taken only slow and halting steps to meet the health care needs of those with permanent disabilities. These are matters of justice, not charity. We are not facing conditions of scarcity so severe that these steps to provide equality of opportunity must be foregone in favor of more pressing needs. The point also has implications for the problem of long-term care for the frail elderly, but I cannot develop them here (cf. Daniels, 1981d).

A final implication of the account raises a different set of issues, namely how to reconcile the demands of justice with certain traditional views of a physician's obligations to his patients. The traditional view is that the physician's direct responsibility is the well-being of his patients, that (with their consent) he is to do everything in his power to preserve their lives and well-being. One effect of leaving all resource-allocation decisions in this way to the micro-level decisions of physicians and patients—especially where third-party payment schemes mean little or no rationing by price—is that cost-ineffective utilization results. In the current cost-conscious climate, there is pressure to make physicians see themselves as responsible for introducing economic considerations into their utilization decisions.

But the issue raised here goes beyond cost-effectiveness. My account suggests that there are important resource-allocation priorities that derive from considerations of justice. In a context of moderate scarcity, it is not possible for physicians to see as their ideal the maximization of the quality of care they deliver regardless of cost. Pursuing that ideal upsets resource-allocation priorities determined by the opportunity principle. Considerations of justice challenge the traditional, perhaps mythical, view that physicians can act as the unrestrained agents of their patients. The remaining task, which I pursue elsewhere, is to show the level at which constraints should be imposed so as to disturb as little as possible what is valuable in the traditional view of physician responsibility (cf. Donabedian, 1979; Daniels, 1981b).
References


Norman Daniels


1981e. What Is the Obligation of the Medical Profession in the Distribution of Health Care? Social Science and Medicine (December) (In press.)


Ethics, and the Aging Society. Chicago: Committee on Human Development, University of Chicago.


Acknowledgments: The research on which this paper is based was funded by Grant Number HS03097 from the National Center for Health Services Research, OASH. This paper was prepared under special commission by the President’s Commission for the Study of Ethical Issues in Medicine, Biomedical and Behavioral Research. I would like to thank Mary Anne Bailey and Daniel Wikler for helpful suggestions. The theory on which some of the argument here is based is developed more fully in Daniels (1981a) and in my book, “Justice and Health Care Delivery” (Cambridge University Press, 1982 est.).

Address correspondence to: Dr. Norman Daniels, Department of Philosophy, Tufts University, Medford, MA 02155.