The ethical justification for developing and providing the means to reduce the burden of disease in developing countries is self-evident. Nevertheless, those who pursue these laudable ends encounter ethical dilemmas at every turn. The development of new interventions requires testing with human subjects, an activity fraught with controversy since the dawn of scientific medicine and especially problematic with poor and vulnerable participants in developing countries. Ethical dilemmas arising in setting priorities among interventions and among individuals in need of care are most acute when needs are great and resources few.

We address some of these concerns in this chapter, identifying some of the principal ethical issues that arise in the development and allocation of effective interventions for developing countries and discussing some alternative resolutions. We omit discussion of two other aspects of these ethical decisions: ensuring that the process of decision making is fair and involves the subject population (Daniels 2000; Holm 1998), and respecting legal obligations under international human rights treaties (Gruskin and Tarantola 2001).

**HEALTH RESOURCE ALLOCATION**

Resource allocation in health and elsewhere should satisfy two main ethical criteria. First, it should be cost-effective—limited resources for health should be allocated to maximize the health benefits for the population served. A cost-effectiveness analysis (CEA) of alternative health interventions measures their respective costs and benefits to determine their relative efficiency in the production of health. Costs are measured in monetary terms; benefits are measured in health improvements. By dividing costs by benefits, one can obtain a cost-to-effectiveness ratio for each health intervention, and interventions can be ranked by these ratios. Although a CEA is typically an economic analysis performed by health economists, it is also a measure of one ethical criterion for the evaluation of health programs. Cost-effectiveness is not merely an economic concern, because improving people’s health and well-being is a moral concern, and an allocation of resources that is not cost-effective produces fewer benefits than would have been possible with a different allocation. Producing more rather than fewer benefits for people is one important ethical consideration in evaluating actions and social policies.

Second, the allocation should be equitable or just; equity is concerned with the distribution of benefits and costs to distinct individuals or groups. The maximization of benefits, which is associated with the general philosophical moral theory of utilitarianism or consequentialism, however, is routinely criticized for ignoring those considerations (Rawls 1971). Equity in health care distribution is complex and embodies several distinct moral concerns or issues that this chapter delineates (Brock 2003a). There is no generally accepted methodology comparable to CEA for determining how equitable a distribution is; nevertheless, allocations are unsatisfactory if equity considerations are ignored.

Efficiency and equity can sometimes coincide. In some of the world’s poorest countries, for example, health budgets support tertiary care and travel to clinics abroad for the elite and the well connected, even as the poor are denied effective,
low-cost prevention or treatment for life-threatening diseases (Birdsall and Hecht 1995). Moreover, because equity concerns the relative treatment of different individuals, CEA is largely unobjectionable when it is used only for evaluating alternative health interventions that would serve the same patients. However, considerations of equity may conflict with cost-effectiveness and so may provide moral reasons for an allocation that is not cost-effective. The discussion in this chapter accepts that CEA identifies one important ethical criterion in evaluating health care interventions—producing the most benefits possible for individuals served by those interventions—and then focuses on the other ethical criterion of ensuring equitable distribution of those benefits.

This chapter considers two types of equity issues: first, those that arise in the general construction of a CEA—that is, in determining the form of a CEA; second, those that arise in the use of the results of a CEA for resource allocation in the health sector. It is worth noting that, when applied appropriately and broadly to all social conditions and programs that significantly influence health, CEA may often support using resources to affect the so-called social determinants of health—which largely affect the incidence of disease, disability, and premature mortality—rather than using those resources on health care to treat disease. However, we shall focus largely on CEA in the evaluation of health care and public health programs.

**Issues in the Construction of a Cost-Effectiveness Analysis**

Cost-effectiveness analyses require decisions about which costs to include, which if any financial gains should be counted as offsetting costs, whether to include benefits beyond the effects of the intervention on health, and whether all health gains should be valued alike. None of those decisions, in our view, is exclusively a technical issue, and CEA results reflect the analysts’ ethical judgments on those issues.

**Evaluation of Benefits.** Evaluating health benefits within a CEA involves several issues. This chapter assumes that some version of a quality-adjusted life year (QALY) is used to combine the two main benefits of health care—(a) protecting or improving health or health-related quality of life and (b) preserving life. Disability-adjusted life years (DALYs) are a variant of QALYs in that they measure the losses from disability or premature death; a CEA will determine which interventions will maximize QALYs or minimize DALYs. Calculating QALYs requires a metric evaluating the effect of different states of limitations in function on health-related quality of life, such as the Health Utilities Index (Horsman and others 2003). The Disease Control Priorities Project uses the health state valuations or disability weights of the World Health Organization (WHO). The relative value of any particular health state, typically on a scale in which “0” represents death and “1” represents full, undiminished function (or health) is generally determined by soliciting a group of individuals’ preferences for life in that state using standard gambles, time tradeoffs, visual analog scales, or person tradeoffs. In all these methods, a common issue is whose preferences to use for valuing health states. The main debate has been whether to use a randomly selected group of citizens or to use people who have the particular disability or limitation in function being evaluated.

This issue matters because a number of studies have shown that persons without disabilities generally evaluate the quality of life with a particular disability as significantly worse than do persons who have that same disability (Menzel and others 2002). If the preferences of persons without disabilities are used, their lower evaluation of quality of life with various disabilities will mean that fewer QALYs will be produced by life-saving interventions for persons with disabilities than if the preferences of persons with disabilities had been used. However, if we use the preferences of persons with disabilities, then both prevention and rehabilitation will receive less value than if the preferences of persons without disabilities had been used.

This difference in evaluations in part results from ignorance, prejudice, and stereotypes on the part of persons without disabilities about what it is like to live with various disabilities. The difference results as well from the process of adaptation to disability in which disabled persons adjust by learning new skills, cope by adjusting their expectations to their new circumstances, and accommodate by substituting new aims and activities for ones made difficult or impossible by their disabilities (Solomon and Murray 2002). They thus adopt a new valuational perspective for making health and quality-of-life evaluations. Because the adoption of this new perspective resulted from a disability, it will represent a set of values for making choices that reflects a restricted set of abilities. Nevertheless, neither the nondisabled perspective nor the adapted disabled perspective is mistaken; they are only different (Brock 1995). These differences create controversy in the literature over which perspective is correct for cost-effectiveness evaluations in health care.

A second issue is whether, in evaluating interventions that preserve or extend life, we should use life years saved (as QALYs do) or lives saved. Certainly individuals offered two interventions that would preserve their lives for different lengths of time would prefer, all other things being equal, the alternative with the longer period of survival. Moreover, when the differences are extreme—for example, extending group A’s lives by a week or extending an equally numerous group B’s lives by 10 years—virtually everyone would judge this difference to support giving priority to group B. This fact suggests that even the proponent of counting lives saved should require that the lives saved for a shorter period of time must still be saved for a significant period of time; what is significant will depend in part on the
duration of lives saved by the alternative with which it is being compared. Some empirical studies indicate that ordinary people tend not to give much weight to differences in the duration of health benefits to different groups of persons when prioritizing between them, as long as the lesser duration benefits are viewed as significant; this attitude suggests that they favor lives saved over life years saved (Nord and others 1996). The life years saved versus lives saved controversy remains unsettled.

**Should Life Years Be Age Weighted?** The standard assumption in most CEAs using QALYs is that one QALY has the same social value, regardless of the age of the recipient (Gold and others 1996). Thus, equality is adopted as the weighting for QALYs achieved by recipients at different ages, and that is the approach adopted in this volume. The use of any age weighting that gives less value to benefits for the elderly than for younger persons is often charged as unjust age discrimination. Even the use of equally weighted QALYs is often charged as unjust age discrimination because, other things being equal, saving the lives of younger persons will produce more QALYs than saving the lives of older persons. The goal of lives saved, as opposed to life years saved, removes this disadvantage to the elderly from CEAs that use QALYs. However, if the relevant benefit is adding years to life, then standard CEA is neutral or impartial regarding age, in the sense that it gives the same value to a year of life extension whatever the age of its recipient.

WHO, in its burden-of-disease and resource prioritization studies that use DALYs, rejected the equal age weighting that is standard with QALYs. Instead, it gave less value to DALYs prevented for infants, young children, and the elderly, in comparison with persons in their productive adult years. WHO justified this weighting by the fact that the very young and the elderly both tend to be economically, socially, and psychologically dependent on adults during those adults’ productive working and child-rearing years (Murray 1994). This justification is ethically problematic, however, because it assigns different value to meeting people’s health needs on the basis of differences in the instrumental value to others of meeting their needs. This approach differentiates people solely on whether they are a means to benefiting others. The same reasoning would justify giving priority to rich over poor patients with the same medical needs because the rich are more socially productive than the poor, a practice that would be widely regarded as unjust.

Writers in this field have provided different reasons for giving greater value to QALYs for younger patients, however, that are not subject to this moral objection and that are specifically grounded in fairness. For example, Alan Williams has developed an argument to the effect that fairness requires that individuals should each receive “fair innings” of QALYs in their lives (Williams 1997). In this view, the earlier a preventable death could occur and the worse a person’s past health is, the greater is the unfairness the person suffers—so the greater is the moral urgency, grounded in fairness, of preventing the death. The younger a person is, the greater is the moral value of providing a QALY to him or her. This view leaves open to what extent the moral value of QALYs should decline with the age of the recipient. This age weighting to favor the young has been attacked by some as unjust age discrimination, but because an explicit moral justification in terms of fairness is offered for it, critics must show why that justification is unsound.

**What Costs Should Count in Health Cost-Effectiveness Analyses?** No controversy surrounds the inclusion in a CEA of direct costs of a health intervention program or direct health benefits to the intervention’s recipients. Ethical issues do, however, arise in other aspects of the cost calculation (Brock 2003b). A full CEA of alternative health programs should take account of all the economic effects on public or private expenditures of the alternative health interventions or programs under analysis. An example is provided in the consideration of treatment for two alternative health conditions judged to have equally detrimental effects on patients’ health: the first condition permits patients to continue working, and the second interferes with regular work and so has large economic costs to the patients’ employers. Should the costs of treating the second be reduced by the cost savings to the employers from returning the patients to work on a regular basis? If so, the second treatment program will have a more favorable cost-effectiveness ratio than the first, even if it may be no better or worse without consideration of those economic effects. The same issue arises in many other contexts.

From the moral perspectives of both a consequentialist and a standard CEA, these indirect economic effects for others are real benefits or cost reductions and should be part of the CEA. The fundamental moral objection to giving higher priority to treating those who can be treated at lower net cost because of the economic savings to their employers is the same as that with WHO’s instrumental rationale for its age weighting. One condition or group of patients gets higher priority solely because treating it or them is a means to producing economic benefits to others, thereby reducing the net social costs of their treatment. This approach violates the Kantian injunction against treating people solely as means—the first group has lower priority for treatment solely because treating that group is not a means to the economic savings to employers. It fails to give equal moral concern to the health needs of each group of patients because it discriminates against the less socially valuable patients. Conversely, at the macro level of the allocation of resources to health instead of other social goods, the WHO Commission on Macroeconomics and Health has supported increasing health investments in developing countries because such investments often more than pay for themselves in their economic and development benefits (CMH 2001). Using a “separate spheres” view, only the health benefits and health...
costs of alternative health interventions should determine their priority for obtaining resources, but this view remains controversial.

Another aspect of cost calculation concerns whether future health care and other costs, such as old-age payments, that will be incurred as a result of a person's life being saved should be added to the costs of treating that person now. Persons who do not die now because of a life-saving intervention will typically go on to incur future health costs that would not have been incurred had they died now. The U.S. Public Health Service Panel on cost-effectiveness recommended that inclusion of these costs be optional in CEAs (Gold and others 1996). Others have argued that, if CEA is designed to maximize lifetime utility, the future costs should be included (Meltzer 1997). These are costs that would not be incurred if the patient was not saved, but virtually no one would argue that, because of those costs, we should judge a life-saving intervention as not cost-effective and thus deserving of lower priority than interventions that do not have those effects. What does this thinking show? That we are not prepared to allocate health resources on the basis of a full CEA that accounts for all the costs incurred and saved by those interventions—that is, that some should be disregarded on ethical grounds.

**Should Health Benefits and Costs Be Discounted in Cost-Effectiveness Analyses?** As standard practice in CEAs, both health care costs and benefits are discounted at the same rate, for example, 3 percent or 5 percent, and the Disease Control Priorities Project applies a 3 percent discount rate to costs and benefits (Gold and others 1996). Little controversy surrounds the idea that future monetary costs and benefits should be discounted to their present value in a CEA. The same amount of money is worth more if received today than in 10 years because it can be invested at the market rate of interest if received today. For the same reason, costs that can be deferred require fewer present dollars to meet them.

The controversial issue is whether health benefits should be discounted—that is, whether the same magnitude of health benefit has progressively less social value the farther into the future it occurs. This issue is complex and has engendered an extensive literature that cannot be reviewed here, but we can at least try to focus the issue. It is appropriate to discount for the uncertainty about whether potential beneficiaries will survive to receive a future health benefit and to discount for any increased uncertainty about whether a benefit will occur because it is more distant. However, these uncertainties are reflected in the calculation of expected future benefits and do not require that future benefits be discounted. Likewise, if individuals receive a health benefit (such as regaining mobility) sooner rather than later, their total lifetime benefit may be greater, but this fact, too, is reflected in the estimation of the total benefit without discounting.

The ethical issue about discounting is whether, after taking account of such considerations, a health benefit of the same size has progressively less social value the farther into the future that it occurs. To make the issue more concrete, suppose we must decide between two programs: one will save 100 lives now, and the other, say a hepatitis vaccination program, will save 200 lives in 30 years. The vaccination program will save twice as many lives, but if we apply even a 3 percent discount rate to the future lives saved, they are equivalent to only 78 lives saved now, and we should prefer the first program. This example illustrates not only the theoretical issue, but its practical import, too, because discounting future health benefits will systematically tend to disadvantage prevention programs that must be undertaken now but whose benefits occur only at some point in the future. This reasoning applies not only to many vaccination programs, but also to most programs to change unhealthy behaviors in which the benefits generally occur at some later time.

Arguments for discounting health benefits at the same rate as costs have included consistency arguments (Weinstein and Stason 1997), avoidance of paradoxes in allocation concerning research and deferral of spending (Keeler and Cretin 1983), individual or social rates of time preference, and so forth. Those arguments cannot be reviewed here, but whether to discount health benefits is squarely an ethical question about the valuing of health benefits over time and should be explicitly addressed as such in allocating resources.

**Issues in the Use of Cost-Effectiveness Analysis for Resource Allocation**

It is now widely recognized that CEA alone is not a satisfactory guide to resource allocation in all cases. CEA, as customarily formulated, measures the sum of costs and benefits and largely ignores the pattern of their distribution across the affected population. In some cases, the resulting allocation will strike most observers as unfair. Health resource allocators need to take distributional issues into account along with cost-effectiveness.

**Priority to the Worst Off.** Justice requires a special concern for the worst off, as is reflected in aphorisms such as “you can tell the justice of a society by how it treats its least well-off members,” in the well-known Difference Principle in John Rawls’s theory of justice, and by the special concern for the poor within many religious traditions (Brock 2002; Rawls 1971). This concern is often understood to reflect a concern for equality—in particular, equality in outcomes or welfare between people. In the health context, it takes the form of a concern for reducing inequalities in health between persons or groups. A variety of ethical bases underpin a concern for equality in general and for equality in health in particular, and they cannot be explored...
here. It is important, however, to understand that concern for the worst off is different from a concern for equality, because the two can be and often are confused. Raising the position of the worst off will typically reduce inequality, but it need not always do so. Sometimes improving the position of the worst off may unavoidably improve the position of those who are better off even more and thereby increase inequality. Moreover, the concern for equality in outcomes is subject to the “leveling down” objection, in which equality is achieved by making the better-off members worse off, even when doing so in no way benefits those who are worst off. In the face of that objection, many have rejected equality in outcomes in favor of a prioritarian view, according to which benefiting people has greater moral value the worse off those people are (Parfit 1991).

A number of possible lines of reasoning support prioritarianism. For example, the worse off that people are, the greater is the relative improvement that a given size of benefit will provide them, so the more the benefit may matter to them. Alternatively, the greater the undeserved health deprivation or need that an individual suffers, the greater is the moral claim to have it alleviated or met.

However priority to the worst off is justified, an important issue is who the worst off are. In the context of resource allocation in health care, the worst off might be those who are globally worst off, those with the worst overall well-being (such as the poor), or those with the worst health (that is, the sickest). General theories of justice usually focus on people’s overall well-being, often allowing a lower level in one domain of well-being to be compensated for by a higher level in another domain. However, there are both moral and pragmatic reasons for what has been called a separate spheres view, according to which the worst off for the purpose of health resource allocation should be considered to be those with worse health. Morally, for example, Scanlon has argued that “for differences in level to affect the relative strength of people’s claims to help, these differences have to be in an aspect of welfare that the help in question will contribute to” (Scanlon 1997, 227). Pragmatically, it may generally be too difficult, costly, intrusive, and controversial, as well as too subject to mistake and abuse, to have to inquire into all aspects of people’s overall levels of well-being.

Even if health allocation to the worst off should be based on levels of health, other issues remain. For example, are those with worse health those who are sickest now, at the time a health intervention would be provided for them, or those with worse health over time, taking into account past and perhaps expected future health? The latter would give special weight to meeting the health needs of those with long-term chronic diseases and disabilities. Separate spheres would still include past and future health. Should special priority also be given to those whose health is not worse now but is especially vulnerable to becoming worse? Finally, how much priority should the worst off receive? Giving absolute priority to the worst off is implausible because it faces the bottomless pit problem—using very great amounts of resources to produce very limited or marginal gains in the health-related quality of life of the severely ill or disabled. However, there is no apparent principled basis for determining how much priority the worst off should receive.

Aggregation and Cost Differences. The aggregation problem occurs when determining at what point small benefits to a large number of persons should take priority over very large benefits to a few, because the former result in greater aggregate or total benefits (Daniels 1993; Kamm 1993). The issue can be illustrated by the initial effort to prioritize different treatment-condition pairs in the Medicaid program in the U.S. state of Oregon by what was essentially a cost-effectiveness standard. As was widely reported, capping teeth for exposed pulp was ranked just above performing appendectomies for acute appendicitis, even though appendicitis is a life-threatening condition. A variety of methodological problems affected Oregon’s analysis, but this kind of result is to be expected from CEA. The Oregon Health Services Commission estimated that it was possible to provide a tooth capping for more than 100 patients for the cost of one appendectomy, so the aggregate benefits of the many tooth cappings were estimated to exceed the benefit of one appendectomy. As a consequence of results of this sort, the commission fundamentally changed its prioritization methodology to largely ignore cost differences, except in the case of roughly equally beneficial interventions. The commission essentially adopted what might be called a relative effectiveness or benefit standard (Hadorn 1991).

What Oregon’s experience shows is that most people’s sense of priorities is determined by a one-to-one comparison of the benefits of different interventions, in which case appendectomies are clearly a higher priority than tooth capping. That ignores the great differences in costs between different health interventions that a CEA will reflect. Is it then simply a mistake to ignore those cost differences in allocating health resources? At least two moral considerations suggest not. First, empirical studies have shown that many people ignore the cost differences because they believe that patients should not be at a disadvantage in priority for treatment simply because their condition happens to be more expensive to treat than are other patients’ conditions (Nord and others 1995). Second, according to many moral theories, individuals should confront other competitors for scarce resources as individuals, and their priority for treatment should be determined by the urgency of their individual claims to treatment (Scanlon 1997).

Then again, most people and most moral theories do not reject all aggregation of different sizes and costs of health benefits in setting priorities and allocation, although there is no consensus either on when aggregation should be permitted or
for what reasons. However, at a minimum, we suggest that individuals should not be denied very great health benefits—in the extreme case, life-saving interventions—merely to provide small health benefits to a large number of other persons.

**Fair Chances and Best Outcomes.** The thesis that resources should be targeted to interventions in which they will do the most good ascribes a higher priority to those who can be helped more easily or cheaply. This thinking, in turn, implies that some patients will lose out simply because their needs are more difficult or expensive to meet. Consider, for example, a ward with 100 patients, 50 of whom require one pill and 50 of whom require two pills to recover. The patients are otherwise similar. The clinic has 50 pills and must decide how to distribute them. To achieve the best outcome, all 50 pills should be given to the patients who need only one to recover. However, to give each patient an equal chance to recover, entitlement to treatment should be awarded randomly. Seventeen fewer cures would result.

Limited surveys indicate a sharp difference between health professionals and the general public in their responses to this conflict. Most health professionals favor distribution to one-pill patients only, and most members of the general public insist that people should not be penalized for needing two pills (Nord 1999). This division of opinion goes to the heart of CEA, which is precisely a guide to identifying the route to the best outcomes that can be hoped for with existing resources. It also creates a dilemma for those health professionals who maintain that health policy should be based on values most frequently endorsed by the population affected.

The conflict between fair chances and best outcomes arises not only from differences in the costs of treating otherwise similar groups of patients, but also when one group of patients will receive somewhat greater benefits than another at the same cost. The appeal of a fair-chances solution is greater when the difference in cost-effectiveness between the two programs is relatively small compared with the potential gain or loss to individual patients. Suppose that health program A will produce 5,000 QALYs while program B will produce 4,500 QALYs and that the effect on the health or life of each patient served is large—in the extreme, life saving. Patients who would be served by program B could complain that it is not fair that all the resources go to program A and none to B when they have nearly as pressing health needs and would be benefited by treatment nearly as much as the patients served by program A. If all cannot be treated, they might go on to argue, they deserve a fair chance to have their needs met rather than having no chance for treatment only because treating them would produce slightly less benefit than treating the patients served by program A. The small difference in benefits produced for the two groups—for example, a slightly greater life expectancy or more serious disability averted in program A—they argue, is too small to justify the tremendous difference in how the two groups are treated. In the extreme case, some live and others die. The better outcome is produced by funding program A rather than program B, but that additional good is insufficient to justify morally the huge difference in the way the two groups of patients are treated. The conflict between fair chances and best outcomes can arise in a variety of contexts (Kamm 1993).

Preferring the most cost-effective program can also seem unfair because it compounds existing unfair inequalities. For example, screening slum-dwelling black men for hypertension targets the group with the highest incidence and greatest risk of premature death. However, it is more cost-effective to target well-to-do suburban white men, because they have more ordered lives, comply better, have personal doctors and the means to obtain medical services, are more educated, and are more likely to modify their lifestyles wisely. However, if the poor black men are not screened for this reason, it only compounds their existing unjust deprivation and, of course, is also in conflict with giving priority to the worst off.

If those who need a less cost-effective program deserve a fair chance to have their needs met, what would be a fair chance? Some argue that a fair chance is an equal chance, so some random method of selecting which program to fund should be used (Broome 1991). Others suggest proportional chances or a weighted lottery, in which the chance of each program being selected is proportional to the amount of health benefit each would produce, as a way of balancing fair chances against best outcomes (Brock 1988). Alternatively, some resources might go to each program (which is usually possible at the macro level), thereby benefiting some patients in each group—at least if their relative benefits are not strikingly dissimilar—instead of all going to the most cost-effective programs.

Another consideration supports spreading some resources to less cost-effective programs instead ofdevoting them all to the most cost-effective: to give all—or at least more—patients a reason to hope that their health needs will be met. This consideration may be especially important in developing countries where resource scarcity is more severe and adhering strictly to cost-effectiveness criteria could result in large numbers of patients with serious—or even life-threatening—health needs having no hope that their needs will be met.

**Discrimination against Persons with Disabilities.** The use of CEA in resource allocation to maximize the QALYs produced by available resources will often discriminate against persons with disabilities. Many persons with disabilities such as cystic fibrosis, HIV/AIDS, and chronic pulmonary or heart disease have reduced life expectancies or health-related quality of life as a result of their disabilities. Life-extending health care for those people will produce fewer QALYs than for people without them, all else being equal.

When health interventions are aimed at improving quality of life rather than extending life, similar discrimination can
arise. The presence of disabilities can act as comorbidities, making treatment less effective or more expensive (or both) than it would otherwise be, thereby worsening its cost-effectiveness ratio relative to comparable treatment for persons without disabilities. These effects of treatment can result from a disability that exists before treatment and is unrelated to the treatment provided. So it seems that a cost-effectiveness standard for resource allocation discriminates against such persons specifically because of their disabilities. Moreover, this effect will arise not only in the case of preexisting disabilities, but also in the case of patients who become disabled as a result of treatment that is only partially effective.

Several strategies to avoid this discrimination in resource allocation have been suggested. Perhaps the most plausible, at least for the case of life-sustaining treatment, is to ignore differences in patients’ posttreatment quality of life as long as each patient accepts and values that quality of life and to ignore differences in life expectancy after treatment as long as each will receive a significant gain in life extension; obviously, what counts as significant is vague and needs finer definition. Ignoring differences in life expectancy posttreatment fits with empirical evidence that individuals give little weight to duration of benefits in prioritizing between health interventions that serve different individuals.

**Cutoffs for Cost per Quality-Adjusted Life Year.** It is not uncommon in health care allocation to suggest the use of cutoffs tied to cost per QALY, although the cutoffs suggested vary substantially depending on the overall wealth of the country and on the amount that it spends on health care. The cutoffs can be of some value in identifying health interventions that are either good or poor buys, given the society’s overall wealth and overall level of health spending. However, it is important to be clear that such cutoffs should never function as anything more than a rough initial guide in health resource allocation. The various equity considerations discussed briefly above can serve as justification for departing from or violating any cutoffs related to cost per QALY.

**Responsibility for Health Needs.** Some have suggested that health needs for which individuals are morally responsible should have lower priority than health needs for which individuals are not responsible (Moss and Siegler 1991). If individuals are responsible for their health needs and could have taken steps to avoid them, they have weaker claims on social resources to meet those needs than do individuals whose health needs are no fault of their own and could not have been prevented. Smoking and substance abuse are two of the most prominent examples of behaviors often cited. However, differentiating patients by whether they deserve care on the basis of whether they are responsible for their health needs does not fit the practice or norms of medicine, which have the goal of meeting patients’ medical needs.

There are strong moral reasons for considerable caution in letting health resource allocation depend on individuals’ responsibility for their health needs (Wilder 2002). For that practice to be fair to those whose needs receive lower priority because of behavior, (a) the needs must have been caused by the behavior, (b) the behavior must have been voluntary, and (c) the persons must have known that the behavior would cause the health needs and that if they engaged in it their health needs resulting from it would receive lower priority. Smoking shows that these conditions are not easily satisfied. Smoking is one causal factor in much cancer and heart disease, but many smokers do not get those diseases, indicating that other factors, no doubt in part genetic differences for which individuals are not responsible, also play an important causal role. Smoking is typically begun when individuals are young adolescents, and as discussed in chapter 46, it is highly addictive, which undermines the voluntariness of continuing to smoke. Individuals in industrial countries are now generally familiar with the health risks of smoking, but this is less true among less educated populations in developing countries, where smoking is an increasing problem. No one anywhere has been informed before they smoke that, if they do, their health needs from smoking will receive lower priority for treatment than will other health needs.

Thus, it would generally be unfair to give smokers lower priority for treatment of smoking-related diseases on the grounds that they were morally responsible for those health needs, although there may be other behaviors for which individuals could more justifiably be held responsible. Moreover, attempting to make those judgments in individual cases would be extremely difficult and controversial. Given the difficulty of instituting a fair practice that allocates health resources according to people’s moral responsibility for their health needs, we generally have good moral reason to preserve the egalitarian feature of the practice of medicine that looks to patients’ needs for care rather than to whether they deserve care.

**ETHICS IN RESEARCH AND NEW PRODUCT DEVELOPMENT**

All new drugs and other medical products must be tested on human subjects before they are sold. Although participation in health research is often a valuable opportunity for participants, what happens to them is determined not only by their clinicians’ therapeutic intent (if any) but also by the need to ensure that the research yields useful information. Managing the potential conflict between those motivations is often an ethically challenging task, and the issues become particularly contentious when research is conducted in developing countries.
Developing Consensus on Ethics and Human Subjects Research

The central ethical question in health research that involves human subjects is what may be asked of some individuals so that others may benefit. The question arises in any research in which human subjects are asked to participate, but is most pressing if the care that is offered to subjects provides no therapeutic benefit or if that care is compromised by the requirements of the study design. Informed consent, while in most cases a requirement for ethical justification of research involving risk, does not relieve the scientist of responsibility. The ethical question is what potential subjects may be recruited for, even if they do consent.

A rough consensus exists worldwide on the elements of research ethics and, increasingly, on the central role of the ethical review committee, or institutional review board (IRB). This consensus can be traced back to the post–World War II international determination to ensure that the kind of barbaric research practiced by Nazi scientists would not again stain the good name of medical science.1 Three advisory documents have been particularly influential. The Nuremberg Tribunal that conducted the postwar Doctors’ Trial promulgated a code of conduct for medical research that stressed the requirement of informed consent. The World Medical Association issued the first version of its Declaration of Helsinki in 1964 and has revised it several times. A further set of guidelines, issued in 1993 and revised a decade later, was published by the Geneva-based Council for International Organizations of Medical Sciences. Although they lack the force of law, these documents are widely acknowledged as international standards. Indeed, the World Medical Association’s periodic revisions of the Declaration of Helsinki have become focal points for international debates over outstanding issues in research ethics.

The most elaborate codification of research ethics is the so-called Common Rule of Conduct of the U.S. Code of Federal Regulations (title 45, section 46), which derived from the work of the National Commission for the Protection of Human Subjects of Biomedical and Behavior Research of the mid 1970s. In addition to proposing rules governing many aspects of research with human subjects, the commission proposed that the IRB be given the central role in research ethics and be responsible for prior review of research proposals.2 The IRB was a compromise, granting a measure of self-regulation to scientists and an assurance of ethical conduct to the government and the public for publicly funded investigations.

The basic elements of research ethics engender little disagreement. The research must never be brutal or inhumane, and all unnecessary risks should be eliminated. Any net risks to subjects must be justified by the prospect of potential benefits to others. Prospective participants must be told that they are in a study and must be informed of its nature and its risks and benefits. In the case of research that offers therapeutic benefit, scientists must explore the range of reasonable therapeutic alternatives with the patient. Potential subjects must understand that their participation is completely voluntary and that they may withdraw at any time and for any reason. Because they cannot voluntarily shoulder risks, further protection must be provided to those who cannot give consent. Such people include, among others, mentally incompetent or immature participants and those involved in research (chiefly in social psychology) that requires initial deception. Consent, however, is not sufficient to ensure fairness; there should be additional safeguards against unfair distribution of the burdens and benefits of research. Finally, all research that involves potential risks should be reviewed by an IRB acting on the basis of internationally recognized ethical principles.

The global acceptance of these principles and the rapid development of capacity for ethical review attest to the perceived validity of this system of rules and procedures of ethical review. However, there has been relatively little research on how IRBs actually perform. Many IRBs in smaller institutions lack the necessary expertise to review novel or complex proposals, and their institutional setting creates a potential conflict of interest. Government investigations of the adequacy of IRBs for the tasks that are now assigned to them have often been critical (for example, Office of the Inspector General 1998, 2000). IRBs are often overworked and understaffed, resulting in ever-lengthening delays between initial submission of protocols and final approval. Regardless of the value of IRBs, predicting what will pass through them and what might provoke delay or rejection has become a significant concern for medical researchers. The system thus has costs as well as benefits, a fact that lends additional gravity to the controversies that it must resolve.

Goals of Ethical Review of Research

Although the overall purpose of ethical review is to ensure that research with human subjects is ethically defensible, the international consensus specifies several distinct goals that are sometimes in tension with each other:

- **Protection.** Ethical review committees can protect subjects by alerting investigators to unforeseen hazards and by suggesting research designs that can avoid unnecessary risk or reduce the number of subjects exposed to risk. By insisting that a clear explanation of risks and benefits be provided to potential participants, ethical review committees also help potential participants to protect themselves. Ethical review committees often take the name “Committee for the Protection of Human Subjects,” reflecting a central preoccupation of research ethics today.

- **Assurance that participation is voluntary.** Some research cannot be conducted without asking some participants to
endure discomfort or pain, to delay relief from symptoms of their disease, or to risk other harm so that future patients may benefit. Permitting investigators to approach potential subjects in these cases requires an ethical judgment. In approving such a proposal, the function of the ethical review committee is not, strictly speaking, only to protect the subjects (the goal of protection would often be served more effectively by declining to do the research), but also to permit them to be enlisted in the effort to improve health care for others. Thus, a second function of ethical review is to ensure that those who agree to participate do so voluntarily and freely and that they understand what is being asked of them.

- **Equality and fairness.** Although research ethics committees have little authority to address persistent social injustices, a third concern of research ethics is that the benefits and burdens of health research be distributed fairly. This function receives relatively little attention in the literature of research ethics, despite its prominence in such well-known documents as the Belmont Report of the National Commission for the Protection of Human Subjects (1979). Many of the most notorious abuses of research subjects, including the Nazi investigations in the concentration camps, the Japanese biowarfare experiments on Chinese and other civilians, and the Tuskegee research on African Americans suffering from syphilis, were committed on subjects chosen exclusively from disadvantaged groups.

Those three goals of ethical review—protecting subjects; ensuring voluntary, informed participation; and reviewing the fairness of recruitment—are promulgated in the international guidelines and in the Common Rule (and in the regulations of other countries), but they do not always point in the same direction. For example, a research project that asks participants to endure a burden or risk—thus failing to offer full protection—can still meet the requirement of equality if the burden is equally shared.

Ethical review, thus, is not a matter of applying a checklist, but it imposes an obligation of substantial ethical judgment. A key challenge for IRBs is to earn and retain the trust of participants and of the public, a task made more difficult by the unavoidable absence of explicit criteria for approval. This problem is exacerbated by the institutional conflict of interest inherent in the placement of the IRB within the research institution, which prompts concern that the committees will downplay risks to subjects for projects that profit or benefit the institution or its influential staff members. Conversely, IRBs that are fearful of institutional embarrassment or legal sanction in the event of any harm befalling research participants might lean too far in the direction of overprotection of subjects, at the expense of important scientific research initiatives. Both concerns have been raised about the IRB system.

### Current Controversies in Research Ethics

Some of the most sharply disputed issues have arisen in international collaborative research involving scientists and sponsors from wealthy countries conducting experiments in developing countries. Some of the problems are procedural. For example, U.S. agencies have insisted on the same kind of recordkeeping for IRBs in developing countries that is required of IRBs in U.S. research institutions. IRBs in developing countries may accept the same principles of accountability, but they do not have the elaborate staffs and budgets that leading IRBs rely on.

The most difficult disputes involving the ethics of research in developing countries are, however, substantive rather than procedural.

#### Standard of Care

The international guidelines used in navigating the ethical dilemmas of research in developing countries were created for the very different purpose of ensuring that what happened at Dachau and Auschwitz would not recur. It is not clear whether those rules usefully resolve the kinds of dilemmas that arise in, say, Uganda or Peru.

The Declaration of Helsinki, following the Nuremberg Tribunal, requires informed consent of all competent research subjects, and in section 29 states that “the benefits, risks, burdens, and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and therapeutic methods.”

To its supporters, any departure from the letter of the Declaration of Helsinki that would permit an experiment in a poor country that would be forbidden in a rich one would constitute a double ethical standard. In their view, this clause of the Declaration of Helsinki affirms the equal importance of human lives, regardless of wealth or nationality, and stands as a safeguard against exploitation of those made vulnerable by poverty, sickness, and absence of governmental protection.

Opponents, however, argue that this position seems to rule out the possibility of testing cheap new products that may be effective, although perhaps not as effective as other products that the population could not afford. If so, it would be difficult to understand whom the single-standard-of-care position would be protecting, for surely it is better for a seriously ill person to receive a good drug, even if it is not the best, than to receive no drug at all.3

Both points of view deserve respect. The single-standard approach is consistent with the postwar consensus on principles of research ethics, and it offers a bright line between research that amply respects human subjects and that which might result if sponsors and scientists were tempted to roam the globe in search of human subjects who could be used as experimental material with a minimum of expense or trouble.

Opponents of the universal-standard view, however, challenge its premise. It made sense to insist on a single, universal
standard when the problem was Nazi barbarity, because the prevailing standard was high and the medical criminals in the death camps denied it to the imprisoned minority—people unjustly stripped of their entitlements. In Uganda or Nepal, however, care at the highest world standard is available, if at all, to only a small elite.

A full reconciliation of those points of view may not be possible. The authors suggest that a relativized standard should be considered only when the beneficiaries will include the impoverished, sick population. Even in those cases, however, the local standard of care could be adopted in the experiment only if it met or exceeded the standard provided by other countries at similar levels of development.

Placebo Controls and Other Issues Involving Research Design. For certain purposes, scientists use a placebo control even though a proven treatment exists. Patients in these control groups thus receive care that is inferior to what they would experience in good clinical care. Until very recently, the Declaration of Helsinki flatly condemned this practice (its current language is somewhat less restrictive), but the U.S. Food and Drug Administration (FDA) accepted results of these trials in applications for approval of new drugs. The FDA’s justification for this acceptance rests on two claims, one scientific and one ethical. The first is that in certain contexts (for example, for conditions such as depression, in which eligibility criteria and outcomes are subjective, to an appreciable extent, and in which symptoms fluctuate in both treated and untreated patients), active controls may produce misleading indications of equivalency, yielding seemingly positive results that may be spurious. The second is that when only placebo controls can be informative, it is sometimes justifiable to ask participants to be randomized with placebo and thereby to risk discomfort and distress (but not any appreciable risk of death or long-term impairment).

Debates over placebo controls are often joined in the context of disputes over the appropriate standard of care that arise in the case of research in developing countries, but placebo controls are controversial in trials in high-income countries, too. Placebo controls are one instance of a large category of ethical issues in research that require weighing the importance of a scientifically ideal research design against the well-being of participants. For example, a study of long-term chemotherapy to prevent the recurrence of breast cancer was halted before the designated endpoints had been reached after the study’s Data Safety and Monitoring Board decided that continuing the study after a strong trend had been established favoring the chemotherapy would be unfair to the control group. It is notable that in this instance severe criticism of this decision was voiced by an organization representing women at risk for breast cancer, as well as by the editorial board of the New York Times. Critics of the early termination of the trial were, in effect, aligning themselves with the interests of future beneficiaries of the research and possibly against the immediate interests of the women in the control group.

Rights of Host Communities. Ideally, research involving human subjects would be a cooperative endeavor for mutual advantage among free citizens who understand and endorse the need for research and who expect to share both in the burden of serving as research subjects and in the eventual benefit of improved health care. Societies that recruit subjects primarily from lower socioeconomic strata fall short of this ideal; those that do not offer new advances in care to all of their citizens fall even further short, raising serious questions about fairness. Furthest of all from this ideal are some instances of the increasingly common practice of recruiting research subjects among the poorest people in the poorest countries. The means for protecting human subjects in these countries are often nonexistent. Most of their citizens will be unable to afford new drugs developed by firms in industrial countries. It is not clear that these subjects participate voluntarily. Their lack of scientific education or even literacy limits their ability to understand the terms of the proposed agreement with the scientists and sponsors (particularly when consent forms, on legal advice, run to 20 dense pages), and poverty often deprives them of any alternative means of recovering their health.

Despite these potential ethical shortcomings, international collaborative research is assured of continued growth. Some of this research targets diseases affecting mainly poor people, who as a group suffer more from too little research on their populations than from too much. Even research intended to develop therapies that will be affordable only to much wealthier patients can be defended. Individual participants may receive better care than they would otherwise, and visiting scientists offer employment and technical training.

To right the perceived imbalance in what is asked of research subjects in poor countries and the value that is obtained by scientists who experiment on them, some have proposed that sponsors of research in the poorest countries compensate their hosts by offering a supplementary benefit (Glantz and others 1998). One much discussed option is access following the end of the study to any drugs or other therapies whose effectiveness is confirmed in the research. The most limited proposals would restrict this entitlement to individuals who were enrolled in the study (those who received placebo as members of a control group, for instance), and time limits (such as three years) have been proposed in the case of chronic diseases such as HIV/AIDS. More expansive community benefit proposals have called for lifetime access to the treatments by all participants, their families, other members of the local community, or even all citizens of the country. Other proposed benefits include a specified amount of technology transfer, including scientific training and the construction of clinics and laboratories, and cash payments earmarked for health care. A moderate proposal
is to encourage these benefits but to require only that they be discussed and agreed on before investigations are initiated (National Bioethics Advisory Commission 2001).

These proposals are intended to restore fairness to the relationship between participants and those who benefit from research, including scientists and their sponsors and also future beneficiaries of advances in medical science. Among the potential drawbacks are the inability to specify, even roughly, how much is owed to host communities; the inability to determine whether community benefit should be required even of research funded by governments or philanthropists for the benefit of people living in the host communities; and the risk that placing these demands on proposed research projects will drive them away from these very needy sites. Some of these uncertainties may be resolved over time as a variety of approaches are attempted, particularly if they are studied and reported to officials in potential research sites.

These international collaborations would draw less scrutiny if it were clear that all subjects knew what they were getting into and participated of their own free will. Although evidence on this point is mixed, special circumstances in some countries introduce problems that will have to be addressed over the long run. Cultural differences between host populations and scientists may lead to conflicts over who has the authority to speak for the individuals invited to participate in a given study. Regulatory authorities in high-income countries have been reluctant to accept permission by a woman’s husband or by a village chief on behalf of his people in lieu of individual consent. It is often unclear—particularly from the vantage point of an IRB in Europe or the United States—whether the cultural norms of the host population designate the husband or village chief as decision makers in these transactions and whether insistence on concurrent individual consent would be viewed as intrusive or insulting.

Another recurring issue is whether people enrolled in a trial of a promising therapy who are ill and very poor can rightly be viewed as volunteers. The prospect of a cure for a person who would otherwise die would seem to be irresistible, even if the treatment is not up to the standards that even less well-to-do citizens of richer countries would expect. Financial incentives, too, would predictably have a powerful effect on an individual who may always be looking for a day’s wage to feed hungry children. Some IRBs limit payments to compensation for lost wages and travel expenses, but even at this level researchers are asked to change the amounts offered to avoid forcing a choice on the potential participant. As with alleged cultural differences regarding individual informed consent, IRBs operate with scant evidence on this point. It is difficult from a long distance to decide what amount of compensation undermines freedom of choice. It is also unclear whether the moral categories used in these disputes have been adequately thought through. The fact that a poor person finds an attractive offer irresistible will be viewed as evidence of coercion by some observers but nothing more than common sense by others.

Most of these controversies can be traced back to underdevelopment and the inequalities of wealth and education that prevail among and within nations today, but progress in resolving the ethical controversies that have become obstacles to badly needed health research must be made even as these disparities persist. Viewing health research in the context of development and emphasizing research that is targeted to the needs of the poor minorities in poor countries can provide a context in which trust rather than fear or suspicion is the default response in host countries. Efforts to build capacity for ethical review within the host countries, such as financial support for ethical review committees, can place the locus of decision making closer to the people who serve as subjects. Research on the effectiveness of current ethical and regulatory requirements and mechanisms might enhance the process of ethical review while reducing its bureaucratic burden. Meanwhile, the quality and appropriateness of ethical review of this research that takes place in the sponsors’ countries would be enhanced by eliciting the views of officials in developing countries, clinicians, scientists, and community leaders.

NOTES

1. Because our current system of ethical review and regulation of research with human subjects derives from our resolve to prevent the recurrence of earlier abuses, it deserves mention that the standard historical account of research ethics has been seriously incomplete. While the Allies sat in judgment of the Nazi scientists at Nuremberg, abuses of similar scope and savagery practiced by Japanese biowarfare researchers on Chinese and other civilians and prisoners of war were kept secret (and their perpetrators were unpunished) following a pact with the criminal scientists to exchange data for war crimes immunity. Moreover, the Allied governments did not always honor the Nuremberg principles. In the Soviet Union, scientists attempting to develop for clandestine operations poisons that would not be identified on autopsy practiced their craft, with predictably lethal results, on hapless prisoners (Birstein 2001). Abuses in the United States, such as the Tuskegee syphilis study (Brant 2000), have been more widely publicized, but ethical lapses in large-scale Cold War–related studies, ranging from radiation studies on urban populations (Advisory Committee on Human Radiation Experiments 1995) to surreptitious administration of mind-altering drugs such as LSD (Rockefeller Commission 1975), were state secrets.

2. In the United States, the Office of Human Research Protections, an agency of the Department of Health and Human Services, has overall responsibility for oversight of IRBs administering research using U.S. government funds. Its Web site is http://ohrp.osophs.dhhs.gov/.

3. Supporters of the single-standard view might point out that, in its current version, the Declaration of Helsinki does not require that everyone in an experiment receive the best available care, but rather that new treatments be tested against the best available care. But this defense faces further objections. In some cases, testing against the best available care (rather than against the care currently provided to the population or against placebo) will fail to provide the evidence needed to convince the ministry of health or potential donors that funds should be provided. There is a potential contradiction in any view that claims both that all patients in experiments deserve the best care and that it is ethically acceptable to test a new treatment that is not expected to be quite as good as the best currently available.