Health-care reform has remained a controversial sociopolitical issue for the last 2 decades. Part of the controversy at the policy level arises from the question of whether health-care reform will involve rationing medical care. This topic raises fears about unfair treatment of individuals,1 which have been inflamed by assertions that rationing devalues human life.2 Physicians have struggled with the controversy surrounding rationing.3,4 Some deny that rationing occurs and contend that their professional obligations require them not to participate in rationing.5-7 Others admit to rationing8,9 and see just allocation of medical care as part of physicians’ ethical duties.10 Intensivists share this ambivalence. In a recent survey, only 60% vouched that they provide “every patient all beneficial therapies without regard to costs.”11

To be thoughtful participants in the social debate about rationing in medicine, physicians must be well informed. The purpose of this article is to address the following topics: (1) the inevitability of rationing of social goods, including medical care; (2) types of rationing; (3) ethical principles and procedures for fair allocation; and (4) whether rationing ICU care to those near the end of life would result in substantial cost savings.
treatments is appropriate because virtually no treatment in medicine offers certain benefit for an individual patient and because a central point of controversy is whether the potential benefit is large enough or likely enough to occur in order to justify the expense. In this document, we use the terms “rationing” and “resource allocation” synonymously, although we acknowledge that the emotional valence of the two terms is clearly different.

It is also important to note that not all efforts to control health-care costs involve rationing. For example, choosing a less expensive treatment over a more-expensive one does not entail rationing if both are equally effective, because selecting the less costly of the two does not result in the patient being denied a potentially beneficial treatment. In addition, strategies focused on reducing administrative costs and waste in health care (eg, reducing duplicative testing and administrative inefficiencies) are generally not rationing because they do not entail denying patients potentially beneficial care.

Rationing Is Unavoidable

In many industrialized countries, social goods—including health care, education, infrastructure, environmental protection, and public health—draw funding from a common pool. Although need for such social goods is limitless, the resources available to supply them are limited. Inevitably, difficult choices must be made to allocate finite resources in a way that achieves a reasonable balance across the range of important social goods. Attempting to meet all health-care needs would likely overwhelm our capacity to supply basic elements of other social goods, such as public safety, education, and defense. Therefore, some degree of rationing of health care is necessary for the overall well-being of society.

Rationing decisions pervade daily practice in ICUs. For example, it is common to transfer a patient out of an ICU when she might still derive some small degree of benefit from ongoing monitoring; such transfers accommodate the needs of sicker patients in the face of a finite number of ICU beds. Physicians in ICUs also routinely ration their time. They must decide which patients to see first and how much time to spend with each. Physicians also must balance the needs of patients against their nonprofessional obligations, such as responsibilities to their families. It is undoubtedly true that physicians cannot provide every potential benefit to every critically ill patient. Therefore, the reality of practice in ICUs is that patients are routinely denied some potential benefit—however small—through implicit rationing decisions made by physicians at the bedside.

The Appropriateness of Rationing Is Context Specific

The necessity of some rationing in medicine does not mean that all such rationing is ethically justifiable, and a justifiable rationing decision in one health-care system may not be similarly justifiable in another. One example is the rules in many health systems requiring less expensive, less beneficial drugs to be first-line choices over more expensive, more beneficial drugs. This type of rationing is relatively easy to justify in single-payer systems (eg, the government-sponsored health-care plans in Canada and many European countries), in which savings are reinvested in programs to improve the health of the population. Such rationing decisions are harder to justify in a for-profit health system with wasteful administrative mechanisms and in which most profits are passed on to employees and shareholders rather than invested in improving the quality of care for patients.

Levels and Transparency of Rationing

Rationing can occur at multiple levels. The clearest conceptual distinction exists between “macroallocation” and “microallocation” decisions. Macroallocation occurs at the societal level and includes decisions about how to allocate funds across a range of public goods. For example, macroallocation decisions determine how a particular society’s public funds are allocated across social goods, such as defense, education, infrastructure, public health, and health care. Microallocation decisions involve bedside decisions about whether an individual patient will or will not receive a scarce medical resource. Although conceptually distinct, macroallocation decisions and microallocation decisions are related. For example, restrictive macroallocation decisions regarding health-care funding will create more situations in which individual patients must be denied potentially beneficial treatments.

Perhaps the most straightforward examples of the rationing in medicine occur when there is an absolute scarcity of a medical resource, such as organs for transplantation. The United Network for Organ Sharing (UNOS) has developed policies to ration according to weighted organ-specific criteria, such as time on the waiting list, severity of illness, human leukocyte antigen matching, prognostic information, and other considerations. These policies are examples of rationing at the micro level. UNOS explicitly acknowledges that many will die without receiving an organ because of the need to ration. Conceivably, more funding of initiatives to encourage organ donation at the macro level would decrease deaths of patients on transplant waiting lists but would likely come at the cost of funding other important social programs. Scarcity is unavoidable.
in the realm of social goods, and the need to ration is one consequence.

Rationing also occurs because of general fiscal scarcity rather than an absolute scarcity of a particular medical resource. For example, in the early 1990s, Oregon had to cope with escalating medical expenditures for Medicaid recipients in the face of budget deficits. The resulting Oregon Health Plan concurrently set a firm annual health-care budget and expanded the Medicaid eligibility criteria to include all below the federal poverty level. The initial macroallocation decision balanced state health-care spending against competing social goods, such as education, infrastructure, and prisons. The second macroallocation traded providing a larger range of health-care services to less than one-half the state’s poor for providing a basic level of health care to all Oregonians living in poverty.

Oregon covered services according to a published priority list until projected expenditures exhausted the budget; there was not publicly funded coverage for the remaining services. This entailed denying beneficial therapies to some patients (microallocation).

Both the UNOS strategy for organ allocation and the Oregon Health Plan are examples of explicit rationing; these rationing decisions arise from stated principles and rules. In contrast, implicit rationing occurs without formally stated rules or principles. The 46 million uninsured in the United States are an example of implicit rationing at the macro level. Intensivists’ decisions about how much time to spend with each patient are also examples of implicit rationing because they are generally not based on publicly disclosed reasons. In general, implicit rationing raises more concerns about fairness than explicit rationing because the basis of the decisions is not disclosed and because unspoken and illegitimate biases may exert undue influence on the decisions.

Empiric Data on Rationing in ICUs

Empiric data from multiple countries document the rationing of medical services in ICUs. In >10,000 ICU bed triage decisions across North America, Europe, Israel, and Hong Kong, at least 15% of patients were refused ICU admission, of which approximately 15% were attributed to lack of beds. Additionally, during times of ICU bed shortages, admitted patients were more ill at both ICU admission and discharge, average lengths of stay were shorter, and fewer patients were admitted for monitoring, which suggests that some patients are denied potentially beneficial treatment in times of ICU bed shortages. Some centers have attempted to reduce ICU use by making mechanical ventilation available on the wards. This also constitutes rationing because ICU care is associated with lower rates of adverse events and mortality compared with providing mechanical ventilation outside ICUs.

A survey of US intensivists suggests that many believe that they do not ration. These results may reflect a lack of understanding of what rationing is or may reflect a symbolic belief about what physicians should do. In either case, the lack of insight about the inevitability of rationing in ICUs is problematic, because it suggests that many intensivists are not well positioned to be informed participants in the social conversation about how best to make the difficult decisions regarding competing social goods.

What Principles Could Guide Rationing?

A substantial barrier to moving from implicit to explicit approaches to rationing health care is the failure to specify what principle(s) should guide allocation. Many principles could form the basis of rationing decisions in health care, each of which represents a different interpretation of distributive justice. For example, the following have been proposed as valid material principles of distributive justice: (1) to each person an equal share, (2) to each according to need, (3) to each according to effort, (4) to each according to free market conditions, (5) to each so as to maximize overall usefulness. A more comprehensive description of the principles—and how they might be combined into multiprinciple allocation strategies—can be found elsewhere.

A foundational debate about distributive justice is how to navigate the conflicting impulses to maximize efficiency (making decisions so as to produce the most good with the least expenditure), equity (treating individuals equally), and prioritarian conceptions of justice (favoring the worst off). Therefore, we briefly discuss three approaches to allocating scarce resources grounded in these radically different philosophical notions of justice: utilitarianism, egalitarianism, and prioritarianism. We also introduce the “rule of rescue.”

To Each to Maximize Overall Quality-Adjusted Life Years: Utilitarianism

In general terms, utilitarianism seeks to maximize overall benefits at the societal level. There are numerous approaches to quantifying benefits related to health care. Many health economists advocate use of the quality-adjusted life years (QALYs) as the best metric. Rationing by QALYs involves two steps: selecting outcome measures that adjust life-years for quality, and then allocating so as to maximize QALYs. Use of QALYs allows comparisons regarding effectiveness across diseases and services that would otherwise be difficult to compare.
approximately $620 per QALY, ICU treatment of acute renal failure costs approximately $30,625 per QALY, and drotrecogin alfa treatment of patients with systemic inflammatory response syndrome and APACHE (Acute Physiology and Chronic Health Evaluation) II scores <25 costs > $400,000 per QALY.36-38

Rationing by maximizing QALYs has limitations. First, there are important unanswered questions regarding the best methods to quantify quality of life.23,33 For example, a person who has over time adapted to using a wheelchair may rate her quality of life the same as someone who is ambulatory, whereas someone recently confined to a wheelchair might rate her quality of life lower. These differences would lead to substantially different cost per QALY calculations depending on the time point at which quality-of-life assessments were obtained.

Additionally, simple strategies to maximize QALYs fail to consider how the benefits are distributed. For example, saving 95 QALYs distributed among two people in a population of 10 with the disease is not necessarily superior to saving 94 QALYs that are equally distributed across all 10 patients (9.4 QALYs per patient).33,39 because of egalitarian concerns about equal distribution of benefits among similarly situated patients. Discounting lower quality of life may also systematically disadvantage those with chronic illness compared with those with good health; such practice opposes a commonly held moral intuition that it is important to help the worst off, or at least not to enable their poor health to be a self-fulfilling prophesy.

Despite these limitations, the National Institute for Health and Clinical Excellence in the United Kingdom uses QALYs to guide coverage decisions.40 For example, drug treatments costing >£20,000-30,000 per QALY are not considered cost-effective and often are not approved for funding. In the United States, public mistrust of policies incorporating cost considerations has made the use of QALYs and cost-effectiveness analysis a political quagmire.35,41

To Each an Equal Opportunity: Egalitarianism

Egalitarianism emphasizes the equal moral status of individuals by trying to provide equal opportunity to have the basic goods in life.22 A straightforward example of an egalitarian approach to rationing is a lottery to determine priority for receiving a scarce resource.18,33 Many citizens have strong moral intuitions toward egalitarian allocation strategies, even when they come at the expense of utility maximization.42 For example, if there were an insufficient supply of ICU beds for the number of patients in need, an egalitarian might advocate for a lottery to randomly select which patients would be admitted. Lotteries require little knowledge about recipients, can occur rapidly, and resist corruption.33 On the other hand, lotteries—and egalitarian principles of justice in general—are insensitive to factors that are also intuitively important to many, such as patients’ need and likelihood of deriving benefit from treatment.33

First-come, first-served strategies to allocate scarce resources appear to be egalitarian, but often are not.18,33 Existing guidelines support allocating ICU beds in this way,43,44 and prior to 2005 waiting time was the primary criterion for allocating lungs for transplantation.10 However, time on the wait list for organ transplantation is not “random” in two ways. First, it favors those with diseases who are well enough to wait the longest.19 Second, those with power, knowledge, and connections often have the social resources to more quickly secure a position in the queue compared with those who have poor health-care access.18,33

To Each to Favor the Worst Off: Prioritarianism

In general terms, prioritarianism attempts to help those who are considered the worst off by giving them priority in situations in which all cannot receive a particular resource.33 For example, a prioritarian might preferentially allocate medical resources to the young over the old because the young have had the least chance to live through life’s stages.33,34 This “life cycle principle”—which is one example of a prioritarian allocation strategy—has been advocated as a way to allocate scarce organs for transplantation and mechanical ventilators during an influenza pandemic.34,45 The justification for this principle does not rely on considerations of one’s intrinsic worth or social usefulness. Rather, the goal is to give all individuals equal opportunity to live a normal life span. When used alone to guide allocation decisions, the life cycle principle ignores prognostic differences among individuals. This type of objection points to the possibility that multiprinciple allocation strategies may better account for the complex moral considerations at play in such decisions compared with single-principle allocation strategies.34

The Rule of Rescue

The rule of rescue describes a powerful psychologic impulse to attempt to save those facing death, no matter how expensive or how small the chance of benefit. The philosopher Albert Jonsen coined the term and describes it as “the moral response to the immi-
ence of death [which] demands that we rescue the doomed.”46 In many ways, the impulse underlying the rule of rescue is an admirable human response to suffering. However, it also can lead to decisions that confound priority setting meant to maximize population-level outcomes. When Oregon refused to
cover a potentially lifesaving bone marrow transplantation for 7-year-old Coby Howard, there was tremendous public outrage and negative media coverage, which likely arose as a consequence of not satisfying the psychologic impulse to rescue identifiable persons facing death. The emotional costs of rationing ICU care would likely be similarly high because it would lead to the loss of identifiable lives.

Conflicts Between Efficiency, Equity, and the Rule of Rescue

The deep moral tensions between efficiency, equity, and responding to those facing death should not be underestimated. In surveys of physicians, citizens, and economists about how to balance such trade-offs, people generally prioritize treatment that can be made available to everyone, but this view is tempered by impulses to maximize usefulness and to rescue those in need. Finding an acceptable balance between these competing ethical goals remains a serious challenge for the development of explicit rationing policies.

Fair Processes of Rationing

In morally pluralistic societies, reasonable people may be unable to agree about which principles should guide rationing. When such conflicts arise concerning high-stakes outcomes, using fair processes to make decisions acquires special ethical importance. Daniels and Sabin and Daniels have proposed four characteristics of fair processes related to allocation: oversight by a legitimate institution, transparent decision making, reasoning according to information and principles that all can accept as relevant, and procedures for appealing and revising individual decisions. A fifth aspect of procedural fairness is meaningful public engagement. This step is important to identify unanticipated needs and values and to obtain public support. The approach used to develop the Oregon Health Plan priority lists had many elements of procedural fairness: The process was under the authority of the state government, which is a legitimate authority for such policies; there was extensive public engagement; priority setting was explicit and incorporated expert opinion; and mechanisms were created for review and refinement of the priority list.

Recent work by Baum and colleagues and Danis and colleagues has demonstrated the feasibility of public engagement related to priority setting in health care. For example, focus groups with citizens about priority setting during a severe influenza outbreak revealed a strong sense of support for interventions focused on the well-being of the community at large. Citizens also raised other ethical, economic, religious, and social concerns that policy makers must consider to develop just policies that will garner compliance. Other research has focused on engaging community participants in setting priorities for health insurance plans. Similar exercises have been used for research and policy settings in nine states. Their use improved understanding of the need to limit benefits in order to limit health-care spending, increased community mindedness of group decisions, and allowed groups to set priorities that at least 85% of participants were willing to abide by.

Will Fair Processes Fail for Tragic Choices?

Although public engagement and transparency seem indispensable for ethical priority setting in medicine, critics have argued that the emotionally and morally difficult choices raised by the rationing of life-saving medical therapies may prove resistant to rational debate. In their book Tragic Choices, Calabresi and Bobbitt argue that society is unlikely to be able to produce a durable, acceptable solution to the issue of scarcity in medicine because the consequences of denying these treatments to individual patients are intolerable.

They argue that individuals collectively attempt to deny moral responsibility for their role in choices—no matter how ethical or necessary—that consign individuals to death. This denial involves creating the illusion that the suffering arises out of nature rather than from conscious choices. For example, the safety standards in the mining industry do not create the safest possible environment for coal miners; doing so would be prohibitively expensive and threaten the market competitiveness of mining companies. However, when there is a mine accident and identifiable miners are trapped, nothing is spared to save them. This response supports the illusion that the mining accident was not preventable and that all was done to safeguard the lives at stake, while ignoring the initial decision that allowed people to work in conditions with a certain level of risk.

Two repeating processes characterize tragic choices. First, society iteratively remakes macro- and micro-allocation decisions to make human suffering appear as infrequent and random as possible. Second, society chooses ostensibly noncontroversial values to justify rationing decisions until the inherent conflict with basic values is exposed. For example, when hemodialysis was first developed as a life-saving therapy for patients with renal failure, demand outstripped supply, and the Seattle Dialysis Committee was formed to determine who would receive dialysis. This panel made decisions that entailed refusing treatment to patients who died as a result. An exposé of the committee’s
decisions was published in *LIFE* magazine, which generated a national public firestorm. The public’s distaste for allowing identifiable patients to die partly led Congress to authorize universal coverage for hemodialysis. In doing so, society was able to better tolerate the (still unresolved) societal question of how to allocate scarce medical resources because the proposed solution minimized the number of identifiable lives lost. In the last decade, the debate has reemerged in a predictable way, now focused on controlling spending while ensuring a minimum acceptable level of basic care for all. It is not yet clear whether the next iteration of health care reform will produce substantive changes rather than changes that appease our consciences but leave unaddressed the inevitability of tragic choices.

**Would Rationing ICU Care Near the End of Life Save Money?**

ICU care is expensive and not always successful. In the United States, upward of 0.66% of the gross domestic product is spent on critical care services, and care for those who die in ICUs totals tens of billions of dollars a year. It would seem then that the ICU might be an ideal location for rationing. In our experience, some physicians believe that healthcare costs should be substantially reduced by strategies that allow unilateral withdrawal of life support in ICUs when patients do not respond fully to a trial of intensive care. However, Luce and Rubenfeld’s analysis of ICU cost structures reveals that the truth is less straightforward. Because > 80% of hospitals’ budgets are independent of the volume of patients treated (ie, fixed—mortgage, maintenance, utilities, and essential personnel salaries), only 20% of costs are modifiable on a per-patient basis (ie, variable—medications, diagnostic and therapeutic equipment, or patient care supplies). The analysis suggests that authorizing unilateral withdrawal of life support when ICU care appears to be failing is unlikely to meaningfully reduce costs. Several empiric studies support this claim. Limiting the number of ICU beds built—and closing existing ICU beds—presents much greater opportunities for cost savings because both fixed and variable costs would be reduced.

Nonetheless, it is certainly true that some cost savings could result from rationing ICU care for patients with relatively poor chances of benefit, especially if rules were developed that delineated situations in which palliative care rather than ICU care would be provided. However, these types of policies would likely be socially divisive and politically challenging, because they would violate the rule of rescue and result in the deaths of identifiable patients. Because the modest savings achieved may be outweighed by the psychologic costs and social outrage, efforts to explicitly ration health care should likely begin with less controversial medical decisions.

**Conclusions**

Rationing of health care is necessary, unavoidable, and ethically complex. The levels at which health care is rationed, and the transparency of rationing, are important structural considerations in creating a sustainable and just health-care system. Ethical rationing requires deliberate choices guided by reasonably applied principles and fair procedures. How rationing occurs is important because it not only affects individual lives but also expresses what values are most important to society. We live in a world in which need is boundless but resources are not—and medicine is not immune to the consequences of this reality.

**Acknowledgments**

Financial/nonfinancial disclosures: The authors have reported to CHEST that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Role of sponsors: The sponsors had no role in the conception, drafting, or decision to submit this manuscript for publication.

Other contributions: All work on this manuscript was performed at the University of North Carolina at Chapel Hill and the University of Pittsburgh.

**References**


