Principles for allocation of scarce medical interventions
Govind Persad, Alan Wertheimer, Ezekiel J Emanuel

Allocation of very scarce medical interventions such as organs and vaccines is a persistent ethical challenge. We evaluate eight simple allocation principles that can be classified into four categories: treating people equally, favouring the worst-off, maximising total benefits, and promoting and rewarding social usefulness. No single principle is sufficient to incorporate all morally relevant considerations and therefore individual principles must be combined into multiprinciple allocation systems. We evaluate three systems: the United Network for Organ Sharing points system, quality-adjusted life-years, and disability-adjusted life-years. We recommend an alternative system—the complete lives system—which prioritises younger people who have not yet lived a complete life, and also incorporates prognosis, save the most lives, lottery, and instrumental value principles.

In health care, as elsewhere, scarcity is the mother of allocation.1 Although the extent is debated,1,2 the scarcity of many specific interventions—including beds in intensive care units,3 organs, and vaccines during pandemic influenza4—is widely acknowledged. For some interventions, demand exceeds supply. For others, an increased supply would necessitate redirection of important resources, and allocation decisions would still be necessary.5

Allocation of scarce medical interventions is a perennial challenge. During the 1940s, an expert committee allocated—without public input—then-novel penicillin to American soldiers before civilians, using expected efficacy and speed of return to duty as criteria.6 During the 1960s, committees in Seattle allocated scarce dialysis machines using prognosis, current health, social worth, and dependants as criteria.7 How can scarce medical interventions be allocated justly? This paper identifies and evaluates eight simple principles that have been suggested.8–12 Although some are better than others, no single principle allocates interventions justly. Rather, morally relevant simple principles must be combined into multiprinciple allocation systems. We evaluate three existing systems and then recommend a new one: the complete lives system.

Simple allocation principles
Eight simple ethical principles for allocation can be classified into four categories, according to their core ethical values: treating people equally, favouring the worst-off, maximising total benefits, and promoting and rewarding social usefulness (table 1). We do not regard ability to pay as a plausible option for the scarce life-saving interventions we discuss.

Some people wrongly suggest that allocation can be based purely on scientific or clinical facts, often using the term “medical need”.13,14 There are no value-free medical criteria for allocation.13,14 Although biomedical facts determine a person’s post-transplant prognosis or the dose of vaccine that would confer immunity, responding to these facts requires ethical, value-based judgments.

When evaluating principles, we need to distinguish between those that are insufficient and those that are flawed. Insufficient principles ignore some morally relevant considerations. Conversely, flawed principles recognise morally irrelevant considerations: inherently flawed principles necessarily recognise irrelevant considerations, whereas practically flawed principles allow irrelevant considerations to affect allocation. Principles that are individually insufficient could form part of an acceptable multiprinciple system, whereas systems that include flawed principles are untenable because they will always recognise irrelevant considerations.

Treating people equally
Many scarce medical interventions, such as organ transplants, are indivisible. For indivisible goods, benefiting people equally entails providing equal chances at the scarce intervention—equality of opportunity, rather than equal amounts of it.17 Two principles attempt to embody this value.

Lottery
Allocation by lottery has been used, sometimes with explicit judicial and legislative endorsement, in military conscription, immigration, education, and distribution of vaccines.18–20

Lotteries have several attractions. Equal moral status supports an equal claim to scarce resources.20 Even among only roughly equal candidates, lotteries prevent small differences from drastically affecting outcome.20 Some people also support lottery allocation because “each person’s desire to stay alive should be regarded as of the same importance and deserving the same respect as that of anyone else”.21 Practically, lottery allocation is quick and requires little knowledge about recipients.18 Finally, lotteries resist corruption.21

The major disadvantage of lotteries is their blindness to many seemingly relevant factors.21,22 Random decisions between someone who can gain 40 years and someone who can gain only 4 months, or someone who has already lived for 80 years and someone who has lived only 20 years, are inappropriate. Treating people equally often fails to treat them as equals.21 Ultimately, although allocation solely by lottery is insufficient, the lottery’s
simplicity and resistance to corruption suggests that it could be incorporated into a multiprinciple system. 28

First-come, first-served
Within health care, many people endorse a first-come, first-served distribution of beds in intensive care units or organs for transplant. The American Thoracic Society defends this principle as “a natural lottery—an egalitarian approach for fair [intensive care unit] resource allocation.” Others believe it promotes fair equality of opportunity, and allows physicians to avoid discontinuing interventions, such as respirators, even when other criteria support moving those interventions to new arrivals. Some people simply equate it to lottery allocation. 31

As with lottery allocation, first-come, first-served ignores relevant differences between people, but in practice fails even to treat people equally. It favours people who are well-off, who become informed, and travel more quickly, and can queue for interventions without competing for employment or child-care concerns. Queues are also vulnerable to additional corruption. As New York State’s pandemic influenza planners stated, “Those who could figuratively (and sometimes literally) push to the front of the line would be vaccinated and stand the best chance for survival”. First-come, first-served allows morally irrelevant qualities—such as wealth, power, and connections—to decide who receives scarce interventions, and is therefore practically flawed.

Favouring the worst-off: prioritarianism
Franklin Roosevelt argued that “the test of our progress is not whether we add more to the abundance of those who have much; it is whether we provide enough for those who have too little”. Philosophers call this preference for the worst-off prioritarianism. Some define being worst-off as currently lacking valuable goods, whereas others define it as lacking valuable goods throughout one’s entire life. Two principles embody these two interpretations.

Sickest first
Treating the sickest people first prioritises those with the worst future prospects if left untreated. The so-called rule of rescue, which claims that “our moral response to the imminence of death demands that we rescue the doomed”, exemplifies this principle. Transplantable livers and hearts, as well as emergency-room care, are allocated to the sickest individuals first. Some people might argue that treating the sickest individuals first is intuitively obvious. Others claim that the sickest people are also probably worst off overall, because healthier people might recover unaided or be saved later by new interventions. Finally, sickest-first allocation appeals to prognosis if untreated—a criterion clinicians frequently consider.

On its own, sickest-first allocation ignores post-treatment prognosis: it applies even when only minor gains at high cost can be achieved. To circumvent this result, some
misleadingly claim that sick people with a small but clear chance of benefit do not have a medical need. Sick recipients’ prognoses are wrongly assumed to be normal, even though many interventions—such as liver transplants—are less effective for the sickest people. If the failure to take account of prognosis were its only problem, sickest-first allocation would merely be insufficient. However, it myopically bases allocation on how sick someone is at the current time—a morally arbitrary factor in genuine scarcity. Preferential allocation of a scarce liver to an acutely ill person unjustly ignores a currently healthier person with progressive liver disease, who might be worse off when he or she later suffers liver failure. Favouring those who are currently sickest seems to assume that resource scarcity is temporary: that we can save the person who is now sickest and then save the progressively ill person later. However, even temporary scarcity does not guarantee another chance to save the progressively ill person. Furthermore, when interventions are persistently scarce, saving the progressively ill person later will always involve depriving others. When we cannot save everyone, saving the sickest first is inherently flawed and inconsistent with the core idea of priority to the worst-off.

Youngest first
Although not always recognised as such, youngest-first allocation directs resources to those who have had less of something supremely valuable—life-years. Dialysis machines and scarce organs have been allocated to younger recipients first, and proposals for allocation in pandemic influenza prioritise infants and children. Daniel Callahan has suggested strict age cut-offs for scarce life-saving interventions, whereas Alan Williams has suggested a system that allocates interventions based on individuals’ distance from a normal life-span if left unaided. Prioritising the youngest gives priority to the worst-off—those who would otherwise die having had the fewest life-years—and is thus fundamentally different from favouritism towards adults or people who are well-off. Also, allocating preferentially to the young has an appeal that favouring other worst-off individuals such as women, poor people, or minorities lacks: “Because [all people] age, treating people of different ages differently does not mean that we are treating persons unequally.” Prudent planners would allocate life-saving interventions to themselves earlier in life to improve their chances of living to old age. These justifications explain much of the public preference for allocating scarce life-saving interventions to younger people.

Strict youngest-first allocation directs scarce resources predominantly to infants. This approach seems incorrect. The death of a 20-year-old young woman is intuitively worse than that of a 2-month-old girl, even though the baby has had less life. The 20-year-old has a much more developed personality than the infant, and has drawn upon the investment of others to begin as-yet-unfulfilled projects. Youngest-first allocation also ignores prognosis, and categorically excludes older people. Thus, youngest-first allocation seems insufficient on its own, but it could be combined with prognosis and lottery principles in a multiprinciple allocation system.

Maximising total benefits: utilitarianism
Maximising benefits is a utilitarian value, although principles differ about which benefits to maximise.

Save the most lives
One maximising strategy involves saving the most individual lives, and it has motivated policies on allocation of influenza vaccine and responses to bioterrorism. Since each life is valuable, this principle seems to need no special justification. It also avoids comparing individual lives. Other things being equal, we should always save five lives rather than one.

However, other things are rarely equal. Some lives have been shorter than others; 20-year-olds have lived less than 70-year-olds. Similarly, some lives can be extended longer than others. How to weigh these other relevant considerations against saving more lives—whether to save one 20-year-old, who might live another 60 years if saved, or three 70-year-olds who could only live for 10 years each—is unclear. Although insufficient on its own, saving more lives should be part of a multiprinciple allocation system.

Prognosis or life-years
Rather than saving the most lives, prognosis allocation aims to save the most life-years. This strategy has been used in disaster triage and penicillin allocation, and motivates the exclusion of people with poor prognoses from organ transplantation waiting lists. Maximising life-years has intuitive appeal. Living more years is valuable, so saving more years also seems valuable.

However, even supporters of prognosis-based allocation acknowledge its inability to consider distribution as well as quantity. Making a well-off person slightly better off rather than slightly improving a worse-off person’s life would be unjust; likewise, why give an extra year to a person who has lived for many when it could be given to someone who would otherwise die having had few? Similarly, giving a few life-years to many differs from giving many life-years to a few. As with the principle of saving the most lives, prognosis is undeniably relevant but insufficient alone.

Promoting and rewarding social usefulness
Unlike the previous values, social value cannot direct allocation on its own. Rather, social value allocation prioritises specific individuals to enable them to promote other important values, or rewards them for having promoted these values. In view of the multiplicity of reasonable values in society and in view of what is at stake, social value allocation must
Reciprocity allocation, like instrumental value allocation, might potentially require time-consuming, intrusive, and demeaning inquiries, such as investigating whether a person adhered to a healthy lifestyle.\textsuperscript{12,22} Furthermore, unlike instrumental value, reciprocity does not have the future-directed appeal of promoting important health values. Ultimately, the appropriateness of allocation based on reciprocity seems to depend in a complex way on several factors, such as seriousness of sacrifice and irreplaceability. For instance, former organ donors seem to deserve reciprocity since they make a serious sacrifice and since there is no surplus of organ donors. By contrast, laboratory staff who serve as vaccine production workers do not incur serious risk nor are they irreplaceable, so reciprocity seems less appropriate for them.

**Assessing principles: allocation systems**

Which principles best embody morally relevant values? First-come, first-served is flawed in practice because it unwittingly allows irrelevant considerations, such as wealth, to affect allocation decisions, whereas a lottery is insufficient but not flawed. Similarly, sickest-first allocation is inherently flawed, whereas the youngest-first principle, though insufficient, recognises the important value of priority to the worst-off. Both utilitarian principles—maximising lives saved and prognosis—are relevant but insufficient, and usefulness and reciprocity are relevant where irreplaceable individuals make serious sacrifices, such as those during public health emergencies.

Ultimately, no principle is sufficient on its own to recognise all morally relevant considerations. Combining principles into systems increases complexity and controversy, but is inevitable if allocations are to incorporate the complexity of our moral values (table 2). People disagree about which principles to include and how to balance them. Many allocation systems do not make their content explicit, nor do they justify their choices about inclusion, balancing, and specification.\textsuperscript{1} Elucidating, comparing, and evaluating allocation systems should be a research priority.\textsuperscript{7}

**United Network for Organ Sharing (UNOS) points systems**

The UNOS points systems are used for organ allocation (table 2). They combine three principles: sickest-first (current medical condition); first-come, first-served (waiting time); and prognosis (antigen, antibody, and blood type matching between recipient and donor). UNOS weights principles differently depending on the organ distributed. Kidney and pancreas allocation is mainly by waiting time, with some weight given to sickest-first and prognosis.\textsuperscript{23} Conversely, heart allocation weights sickest-first principles heavily and waiting time less so.\textsuperscript{23} Lung and liver allocation takes into account waiting time, sickest-first, and prognosis.\textsuperscript{23} Historically, no UNOS system has emphasised prognosis, although
UNOS’s most recent policy discussions on lung allocation suggest such a change.66

The UNOS point systems are flexible: conceivably, they could include any simple principle by translating it into a points framework. The systems are easily revisable to weight one principle more heavily than others.

Current UNOS systems incorporate two flawed simple principles: first-come, first-served and sickest first. They are also vulnerable to additional exploitation. Taking advantage of the first-come, first-served principle, well-off patients place themselves on multiple waiting lists.57

Exploiting the sickest-first element, some transplant centres have temporarily altered or misrepresented their patients’ health status to get them scarce organs, making sickest-first both practically and inherently flawed.56,59

Furthermore, UNOS points systems do not appropriately consider the benefit-maximising principles, prognosis, and saving the most lives, nor do they include youngest-first allocation. Most dramatically, multiple-organ transplants to one individual are permitted, even when a heart-lung-liver combination could save three lives if transplanted separately.66 Similarly, policy revisions during the 1990s de-emphasised organ-recipient matching even though poorer matching leads to fewer lives saved.52

Attempts to remedy these deficiencies have been covert and haphazard. In an effort to implement prognosis allocation tacitly, ill or old people have been excluded from supposedly first-come, first-served waiting lists.62

Physicians can misdiagnose comorbidities as contraindications, wrongly implying that transplants will harm recipients, rather than explicitly practising prognosis-based allocation.63 Some have proposed so-called old-for-old policies that match donor organ age to recipient age—misrepresenting both youngest-first and prognosis-based allocation as biological fact.64 Others have advocated local rather than national waiting lists to circumvent sickest-first allocation.68,69 Explicit and public acknowledgment of allocation strategies would be preferable to this surreptitious and piecemeal approach.

**Quality-adjusted life-years**

Allocation systems based on quality-adjusted life-years (QALY) have two parts (table 2). One is an outcome measure that considers the quality of life-years. As an example, the quality-of-life measure used by the UK National Institute for Health and Clinical Excellence (NICE).65,66 Both the ethics and efficacy of QALY allocation have been substantially discussed.66

The QALY outcome measure has problems. Even if a life-year in which a person has impaired mobility is worse than a healthy life-year, someone adapted to wheelchair use might reasonably value an additional life-year in a wheelchair as much as a non-disabled person would value an additional life-year without disability.70 Allocators have struggled with this issue.62

More importantly, maximising the number of QALYs is an insufficient basis for allocation. Although QALY advocates appeal to the idea that all QALYs are equal,
Department of Ethics

Figure: Age-based priority for receiving scarce medical interventions under the complete lives system

People, not QALYs, deserve equal treatment. Treatment of a serious disease such as appendicitis gives a few people many more QALYs, whereas treatment of a minor problem like uncapped teeth gives many people a few more QALYs. Even though the two strategies produce equal numbers of QALYs, they treat individuals very differently. Likewise, giving QALYs to someone who has had few life-years differs morally from giving them to someone who has already had many. Ultimately, QALY allocation systems do not recognise many morally relevant values—such as treating people equally, giving priority to the worst-off, and saving the most lives—and are therefore insufficient for just allocation.

Disability-adjusted life-years

WHO endorses the system of disability-adjusted life-year (DALY) allocation (table 2). As with QALY allocation, DALY allocation does not consider interpersonal distribution. DALY systems also incorporate quality-of-life factors—for instance, they equate a life-year with blindness to roughly 0.6 healthy life-years. Additionally, DALY allocation ranks each life-year with the age of the person as a modifier: “The well-being of some age groups, we argue, is instrumental in making society flourish; therefore collectively we may be more concerned with improving health status for individuals in these age groups.” This argument, although used to justify age-weighting, would equally justify counting the life-years of economically productive people and those caring for others for more.

DALY allocation wrongly incorporates age into the outcome measure, claiming that a year for a younger person is in itself more valuable. Priority for young people is better justified on grounds of distributive justice. Also, the use of instrumental value to justify DALY allocation resembles that used in Seattle’s dialysis allocation, which inappropriately favoured wage earners and carers of dependants.

The complete lives system

Because none of the currently used systems satisfy all ethical requirements for just allocation, we propose an alternative: the complete lives system. This system incorporates five principles (table 2): youngest-first, prognosis, save the most lives, lottery, and instrumental value. As such, it prioritises younger people who have not yet lived a complete life and will be unlikely to do so without aid. Many thinkers have accepted complete lives as the appropriate focus of distributive justice: “individual human lives, rather than individual experiences, are the units over which any distributive principle should operate.” Although there are important differences between these thinkers, they share a core commitment to consider entire lives rather than events or episodes, which is also the defining feature of the complete lives system.

Consideration of the importance of complete lives also supports modifying the youngest-first principle by prioritising adolescents and young adults over infants (figure). Adolescents have received substantial education and parental care, investments that will be wasted without a complete life. Infants, by contrast, have not yet received these investments. Similarly, adolescence brings with it a developed personality capable of forming and valuing long-term plans whose fulfilment requires a complete life. As the legal philosopher Ronald Dworkin argues, “It is terrible when an infant dies, but worse, most people think, when a three-year-old child dies and worse still when an adolescent does.” This argument is supported by empirical surveys. Importantly, the prioritisation of adolescents and young adults considers the social and personal investment that people are morally entitled to have received at a particular age, rather than accepting the results of an unjust status quo. Consequently, poor adolescents should be treated the same as wealthy ones, even though they may have received less investment owing to social injustice.

The complete lives system also considers prognosis, since its aim is to achieve complete lives. A young person with a poor prognosis has had few life-years but lacks the potential to live a complete life. Considering prognosis forestalls the concern that disproportionately large amounts of resources will be directed to young people with poor prognoses. When the worst-off can benefit only slightly while better-off people could benefit greatly, allocating to the better-off is often justifiable. Some small benefits, such as a few weeks of life, might also be intrinsically insignificant when compared with large benefits.

Saving the most lives is also included in this system because enabling more people to live complete lives is better than enabling fewer. In a public health emergency, instrumental value could also be included to enable more people to live complete lives. Lotteries could be used when making choices between roughly equal recipients, and also potentially to ensure that no individual—irrespective of age or prognosis—is seen as beyond saving. Thus, the complete lives system is complete in another way: it incorporates each morally relevant simple principle.

When implemented, the complete lives system produces a priority curve on which individuals aged between roughly 15 and 40 years get the most substantial chance, whereas the youngest and oldest people get chances that are attenuated (figure). It therefore superficially resembles...
the proposal made by DALY advocates; however, the complete lives system justifies preference to younger people because of priority to the worst-off rather than instrumental value. Additionally, the complete lives system assumes that, although life-years are equally valuable to all, justice requires the fair distribution of them. Conversely, DALY allocation treats life-years given to elderly or disabled people as objectively less valuable.

Finally, the complete lives system is least vulnerable to corruption. Age can be established quickly and accurately from identity documents. Prognosis allocation encourages physicians to improve patients’ health, unlike the perverse incentives to sicken patients or misrepresent health that the sickest-first allocation creates.8,10

Objections

We consider several important objections to the complete lives system.

The complete lives system discriminates against older people.8,10 Age-based allocation is ageism.10 Unlike allocation by sex or race, allocation by age is not invidious discrimination; every person lives through different life stages rather than being a single age.1,9 Even if 25-year-olds receive priority over 65-year-olds, everyone who is 65 years now was previously 25 years.8 Treating 65-year-olds differently because of stereotypes or falsehoods would be ageist; treating them differently because they have already had more life-years is not.

Age, like income, is a “non-medical criterion” inappropriate for allocation of medical resources.14,15 In contrast to income, a complete life is a health outcome. Long-term survival and life expectancy at birth are key health-care outcome variables.14 Delaying the age at onset of a disease is desirable.8,15,16

The complete lives system is insensitive to international differences in typical lifespan. Although broad consensus favours adolescents over very young infants, and young adults over the very elderly people, implementation can reasonably differ between, even within, nation-states.8,15 Some people believe that a complete life is a universal limit founded in natural human capacities, which everyone should accept even without scarcity.17 By contrast, the complete lives system requires only that citizens see a complete life, however defined, as an important good, and accept that fairness gives those short of a complete life stronger claims to scarce life-saving resources.

Principles must be ordered lexically: less important principles should come into play only when more important ones are fulfilled.8 Rawls himself agreed that lexical priority was inappropriate when distributing specific resources in society, though appropriate for ordering the principles of basic social justice that shape the distribution of basic rights, opportunities, and income.1 As an alternative, balancing priority to the worst-off against maximising benefits has won wide support in discussions of allocative local justice.1,8,9 As Amartya Sen argues, justice “does not specify how much more is to be given to the deprived person, but merely that he should receive more”.8

Accepting the complete lives system for health care as a whole would be premature. We must first reduce waste and increase spending.11,10 The complete lives system explicitly rejects waste and corruption, such as multiple listing for transplantation. Although it may be applicable more generally, the complete lives system has been developed to justly allocate persistently scarce life-saving interventions.8,16 Hearts for transplant and influenza vaccines, unlike money, cannot be replaced or diverted to non-health goals; denying a heart to one person makes it available to another. Ultimately, the complete lives system does not create “classes of Untermenschen whose lives and well being are deemed not worth spending money on”,10 but rather empowers us to decide fairly whom to save when genuine scarcity makes saving everyone impossible.

Legitimacy

As well as recognising morally relevant values, an allocation system must be legitimate. Legitimacy requires that people see the allocation system as just and accept actual allocations as fair. Consequently, allocation systems must be publicly understandable, accessible, and subject to public discussion and revision.10 They must also resist corruption, since easy corruptibility undermines the public trust on which legitimacy depends. Some systems, like the UNOS points systems or QALY systems, may fail this test, because they are difficult to understand, easily corrupted, or closed to public revision. Systems that intentionally conceal their allocative principles to avoid public complaints might also fail the test.11

Although procedural fairness is necessary for legitimacy, it is unable to ensure the justice of allocation decisions on its own.9,12 Although fair procedures are important, substantive, morally relevant values and principles are indispensable for just allocation.9,12

Conclusion

Ultimately, none of the eight simple principles recognise all morally relevant values, and some recognise irrelevant values. QALY and DALY multiprinciple systems neglect the importance of fair distribution. UNOS points systems attempt to address distributive justice, but recognise morally irrelevant values and are vulnerable to corruption. By contrast, the complete lives system combines four morally relevant principles: youngest-first, prognosis, lottery, and saving the most lives. In pandemic situations, it also allocates scarce interventions to people instrumental in realising these four principles. Importantly, it is not an algorithm, but a framework that expresses widely affirmed values: priority to the worst-off, maximising benefits, and treating people equally. To achieve a just allocation of scarce medical interventions, society must embrace the challenge of implementing a coherent multiprinciple framework rather than relying on simple principles or retreating to the status quo.
Conflict of interest statement
We declare that we have no conflict of interest.

Acknowledgments
We thank Dan Brock, Daniel Callahan, David Heyd, Frances Kamm, Dennis Thompson, members of the Department of Bioethics, and two peer reviewers for helpful suggestions. The views expressed are the authors’ own. They do not reflect any position or policy of the National Institutes of Health, US Public Health Service, or Department of Health and Human Services.

References


79 Richardson J. Age weighting and discounting: what are the ethical issues? Melbourne: Centre for Health Program Evaluation; 1999.


