Value Judgment in the Oregon Medicaid Experiment

ROBERT M. KAPLAN, PhD

Oregon proposed a unique social experiment in which combinations of medical conditions and treatments were prioritized. Under the proposed program, providers would not be reimbursed for services relevant to 17% of the condition-treatment pairs. The program was designed to expand access and allow significantly more residents to qualify for Medicaid. The original Oregon proposal used four levels of human judgment: community values assessed in town meetings; ratings of the desirability of health states; medical judgment of treatment efficacy; and subjective reordering of the list by Oregon Health Services Commissioners. In August 1992, the Department of Health and Human Services rejected Oregon's application to proceed with the experiment, objecting to the use of one of the four types of subjective data: ratings of the desirability of health states. A revised application that eliminated this one subjective component was approved in March 1993. This paper demonstrates that among the four levels of judgment, the ratings of health states were supported by the most evidence of reliability and validity. Implications for future prioritization experiments are discussed. Key words: Oregon Medicaid Experiment; prioritization; cost-effectiveness; QALY; Medicaid. (Med Care 1994;32:975-988)

The American Health Care system is in a serious crisis. Estimated 1994 health care costs exceeded $900 billion, whereas access to the system is severely restricted for more that 37 million uninsured Americans. Medicaid programs in nearly all states are failing because of increasing costs and little potential to keep up with growing demand. The problems with the health care system are threefold and might be described by the 3 A's: Affordability, Access, and Accountability. Beyond the problems of excessive expense (affordability) and limited access, there is an accountability problem because the system is poorly equipped to document what it produces. In other words, we have no evidence that greater expenditures on health care translate into better health for American citizens.

In response to these problems, the State of Oregon proposed an innovative experiment that has been the subject of considerable commentary. The Oregon proposal attempted to address all three issues simultaneously. The strategy was to rank all medical services (excluding long-term care, mental health, and services for the disabled) by cost-effectiveness. A substantial proportion of medical procedures have little or no effect on patient outcomes. Nevertheless, physicians and hospitals are reimbursed for offering these services to Medicaid recipients. At
the same time, more citizens are uninsured or significantly underinsured than are covered by the Medicaid programs. Because of the continuing financial difficulties, Medicaid budgets have been unable to keep pace with the rising costs of health care. As a result, most Medicaid programs have continually readjusted eligibility thresholds downward. Clearly, this has resulted in the rationing of health care. Rationing occurs by allowing some individuals into Medicaid while excluding others who are equally in need. The peculiarities of Medicaid eligibility force the system to ration by default. \(^9\)

Oregon's plan used principles of cost-effectiveness that have been endorsed in the literature of various academic fields. \(^4\)-\(^5\),\(^10\)-\(^11\) However, the proposal to actually use these methods for resource allocation was bold and different, and it generated considerable controversy. \(^12\)-\(^13\) After detailed study by several groups, the plan was regarded as a reasonable candidate for social experimentation. \(^6\)

**Department of Health and Human Services Rejection of Oregon Application**

Because Oregon intended to revise traditional Medicaid policies, they needed a waiver to perform their demonstration experiment. Such experimentation had been widely promoted by the Bush administration. However, in August 1992, the Department of Health and Human Services rejected Oregon's application for a waiver that would have allowed them to proceed with the demonstration. The rationale for this rejection was that the Oregon proposal violated the Americans with Disability Act which became law in July 1992. Specifically, the department stated that the Oregon preference survey on quality of life "quantified stereotypic assumptions about persons with disability." According to the statement, scholars have found that people without disabilities systematically undervalue the quality of life of those with disabilities. An unpublished discussion paper by David Haddock and an analysis by the US Congress Office of Technology Assessment OTA \(^7\) were cited to support this statement. A revised application that eliminated this quality of life component was approved by the DHHS in March 1993.

The DHHS decision to exclude quality of life data fails to acknowledge that resource allocation decisions necessarily require human judgment. Ultimately, decisions are made by patients, physicians, administrators, or their surrogates. Oregon clearly recognized this and attempted to separate aspects of human judgment. For example, when decisions required medical knowledge, they depended on clinicians. When the decisions required in-depth understanding of human values, they used discussions held in open forums in Oregon towns. When the judgments involved an assessment of quality of life for those with either symptoms or disabilities, they depended on the ratings by Oregonians. This exercise was unusual because all of these judgments were made publicly using methods that could be replicated by others.

The focus of the DHHS review of the Oregon Experiment was on the use of subjective judgments about quality of life. However, there were four types of subjective data in the proposal, not just one. Ultimately, it will be argued here, the approved plan is more subjective than the one rejected. The plan that was accepted may have more potential for allocation errors and discrimination against persons with disabilities.

**Four Subjective Judgments**

Four different types of subjective judgment were used to create the original Oregon Medicaid list. These were: 1) values assessed in town meetings; 2) ratings for health states; 3) subjective judgments about treatment effectiveness; and 4) reprioritization based on commissioner judgment. The following sections review each of these
types of judgments. In the denial of the Oregon Medicaid application in August 1992, only preference weights for health states were specifically cited. Among the four types of judgments, it could be argued that these were the most scientifically justifiable.

Oregon Town Meetings

As part of the process the commission hired a grass roots citizens group that conducted 47 community meetings to determine which values were commonly held by members of the Oregon population. The participants completed a questionnaire concerning eight theoretical health situations. The participants were asked to place nine types of care into three categories: essential, very important, and important. The nine categories ranged from treatment of conditions where the health care is likely to extend life by more than 2 years or to improve quality of life, and “treatment not likely to extend life or make any big improvement in quality of life.” The participants then took part in small group decisions that focused on these issues. On the basis of these discussions group, consensus was estimated and results were recorded.7

The community meetings were well attended. Groups ranged in size from small (7 participants) to fairly large (132 participants) with an average of about 20 participants. Overall, the 47 meetings were attended by more than 1,000 people. Nearly 64% of the participants were women and two-thirds were health care workers. On the basis of the 47 town meetings, 13 community values emerged. These 13 values were grouped by the commission into three attributes: value to society; value to an individual at risk of needing service; and essential to basic health care.

The commission also held a series of 12 public hearings in various parts of the state. During these meetings, testimony was solicited from seniors, handicapped persons, mental health consumers, low income Oregonians, and health care providers. A variety of grass roots organizations participated, providing publicity for the meeting and door-to-door shuttle service for those who had disabilities. On the basis of the testimony, the commissioners concluded that the general public wanted coverage for services that may not be part of standard basic benefit packages. These services included dental care, prevention, mental health care, and chemical dependency services.

Critics were quick to point out that those who attended the community meetings were not representative of all Oregon citizens.12 For example, the experiment was designed to impact on those with no medical insurance and 91% of those who attended the meetings were, themselves insured. More than one-third of those who participated in the meetings had annual incomes in excess of $50,000 and two-thirds had graduated from college. Only 1% of the meeting participants were black. As a result of this imbalance, there was concern that the process was biased in favor of procedures that would benefit the white, educated, and wealthy subpopulations of the society.

In response to these concerns, the Health Services Commission pointed out that they made extensive efforts to recruit all members of society and engaged in special outreach efforts for the poor. Clearly, the process would have been better if the town meeting participants were more representative of the general population. However, public hearings are common components of public policy making. In this respect the meetings in Oregon were not different than the policy forums in nearly all levels of government. The fact that those who testify at public hearings are not representative of the general population has not stopped local, state, or federal governments from enacting legislation that they perceive to be the will of the people. If those participating in the town meetings were reckless in advocating policies to support their own self-interest, we would expect their opinions to be dis-
crepant from those of the general public. The results of the town meetings indicated that the public wanted programs that benefit many, those that emphasize prevention, and those that improve quality of life. Is the suggestion that poor people are opposed to prevention and to equity in health care? Public opinion polls that include random samples from the general population show that the values expressed in the town meetings are consistent with those of the general population.14

Perhaps the strongest rebuttal is that nearly all states are currently rationing health care by changing the Medicaid eligibility criteria and excluding categories of people. These rationing decisions are made with little or no public input. The attempt in Oregon to gain public input represents a significant, although imperfect improvement over the current system.15

Weights for Health States

The Oregon experiment used a model for valuing health states developed at the University of California, San Diego.10,16-20 This general health policy model estimates the impact of any illness or medical treatment in a unit that is equivalent to a year of life. These units are necessary to make direct comparisons between medical interventions that have different specific objectives. Traditional outcome measures that are disease specific cannot be used for these comparisons. For example, a treatment for diabetes mellitus might be evaluated in terms of blood glucose while a treatment for hypertension might be evaluated in terms of blood pressure. Blood sugar and blood pressure outcomes cannot be compared directly to one another.

The common denominator for many illnesses has been life expectancy which is typically analyzed through survival analysis. In survival analysis those that are alive are coded as 1.0, and those who are dead are coded as 0.0. This coding scheme assigns the same score to all those alive. Thus, a person in a coma is scored as alive (1.0) as is a person who is completely functional with no symptoms. The Oregon approach assigned value to these states in recognition that near death and wellness cannot be considered equivalent. This is accomplished by obtaining descriptions of health states and rating the desirability for each. In Oregon these ratings were obtained form random samples of Oregon citizens. These ratings represent utilities for health states and are described as utility weights, preferences, or ratings. We use these terms interchangeably. Once ratings are obtained, they can be used to weight or "quality adjust" years of life. Two years in a state rated as 0.5 (mid way between optimum health and death) are equivalent to 1 year of wellness. Theoretically, treatments that improve quality of life by 10%, (for example from 0.5 to 0.6) produce the equivalent of 1 year of life for each 10 people they affect over the course of 1 year (10 people X 0.1 X 1.0 year = 1.0 QALY). The Oregon commission did not calculate QALYs because it did not consider how many people would be affected.

Criticisms on the use of desirability ratings abound. The most common criticism stems from the assumption that mean ratings vary across patient or demographic groups. For example, in most areas of preference assessment, it is easy to identify differences between different groups of different individuals. Judgments about net health benefits for white Anglo-Saxon men possibly should not be applied to Hispanic men who may give different weight to some symptoms. Social groups may have different utilities for movies, clothing, or political candidates, and these same differences are assumed to extend to health states. Thus, the entire analysis may be highly dependent on the particular group that provided the rating data. In Oregon, for example, critics declared the whole process meaningless because the program was aimed at Medicaid recipients when the ratings came form both Medicaid
recipients and nonrecipients. Other analysts have suggested that ratings from the general population cannot be applied to any particular patient group. Rather, patient utilities from every individual group must be obtained. The August 1992 rejection of the waiver application was based on the assumption that people with disabilities have different utilities than those without disabilities.

The difference between instrumental and terminal preferences is important to understanding this debate. The difference between instrumental and terminal preference is analogous to the difference between a means and an end. Instrumental preferences describe the means by which various assets are attained. For instance, socialists and capitalists hold different instrumental values with regard to the methods for achieving a fully functioning society. Different individuals may have different preferences for how they would like to achieve happiness and evidence suggests that social and demographic groups vary considerably on instrumental values.

Terminal values are the ends, or general states of being that individuals seek to achieve. The classic Rokeach study of values demonstrated that there is very little variability among social groups for terminal preferences. Within health states, there is less reason to believe that different social or ethnic groups will have different utilities for health outcomes. All groups agree that it is better to live free of pain than to experience pain. Freedom from disability is universally preferred over disability states. Although it often is suggested that individuals adapt to disabilities, studies have consistently shown that those with disabilities rate being disabled as less desirable than being disability free. If disability states were preferred to nondisability states there would be no motivation to develop interventions to help those with problems causing the disabilities.

The common assumption that utilities vary across social groups is challenged by the evidence. A study of California preference weights demonstrated some significant, but very small differences, between social and ethnic groups on preferences. Studies have found little evidence for preference difference between patients and the general population. For example, Balaban and colleagues compared preference weights obtained from arthritis patients with those obtained from the general population in San Diego. They found remarkable correspondence for ratings of cases involving arthritis patients (Figure 1). Nerenz and colleagues performed a similar study with cancer patients. Again, they found that preference weights for these patients and the cognitive strategies used to evaluate these descriptions were remarkably similar to those from the general population.

There are few differences by location. Patrick and his colleagues found essentially no differences between utilities for another health status measure among those who live in the UK and those who live in Seattle. We have compared residents of the Navaho nation living in rural Arizona with the general population in San Diego and found few differences. One of the crucial comparisons was between Oregon citizens with disabilities and those without these problems. Among respondents to the Oregon survey, 76 had used a wheelchair or walker at some point in their life. In comparison to people who had not experienced these conditions, those who had used a wheelchair or walker rated 7 of 31 cases significantly higher. Differences for the other 24 cases were not significant. However, the seven statistically significant differences were typically small (less than 0.05 on a scale from 0.0 to 1.0) and rarely were the rank orders different. Figure 2 shows that the ratings from those with and without the experience of these problems were strongly linear and highly correlated.

The utility judgments obtained by the Oregon Health Services Commission used a different scaling methodology and different wording in the case descriptions than in
other studies. Nevertheless, differences between San Diego citizens evaluated in the mid 1970s and Oregon citizens evaluated in the 1990s were small. With the exception of three outliers, the relationship was linear and strong ($r = 0.92$). Figure 3 compares Oregon women and Oregon men, whereas Figure 4 compares Oregon citizens with health insurance with those without insurance. Neither gender nor insurance status has substantial impact. A similar scaling methodology was used by the EuroQol Group in a series of European communities. The data from those studies suggest that differences in preference among citizens in the three different European communities are small and typically not significant. We have used EuroQol scenarios and estimated approximate San Diego preferences for these cases. The results suggest that preferences are similar. We do recognize that there is considerable variability in estimating preferences for a particular case. However, averaged across individuals, the mean preference for different cases in different groups is remarkably similar.

We should not leave the impression that there are never any mean differences in preference. For example, our original study identified some significant differences between social groups. Further, the Oregon Health Services Commission identified
small, but significant, preference differences among those who had previously experienced a condition and those who had not. However, these differences were typically small and it is unlikely that these small differences could affect which services would be funded under the program. Although the OTA analysis showed that it was hypothetically possible for condition-treatment pairs to move from the fundable to the unfundable range on the list, such shifts were very improbable and could only happen for services that already had marginal placements.

**Physician Judgment of Effectiveness**

The health outcomes for people in different treatments was estimated using clinical judgment. For each pair of medical condition and treatment (referred to as condition-treatment or CT pair), clinicians estimated the Quality of Well-Being (QWB) change for the average patient. A clinician panel placed patients into defined categories of mobility, physical activity, and social activity. In addition, clinicians were asked to identify which of 23 symptoms or problems would be most likely at the beginning of treatment. Then, the clinicians were asked to repeat the exercise with their expectations for their average patient following treatment. The time frame for these judgments was 5 years.

The initial task for providers was enormous. In their first attempt at prioritization the group examined more than 1,600 condition treatment pairs and several groups of

![Graph](image-url)

**Fig. 2** Comparison between ever and never in wheelchair or walker for 31 items: Data from Oregon Health Services Commission, Oregon State University.
providers were needed. Ultimately, 54 provider groups and over 200 individual providers from many specialty groups participated. The groups included essentially all licensed practitioner associations in the state of Oregon and represented most medical subspecialties. In addition, the practitioner panels included chiropractors, acupuncturists, and massage therapists.

The major difficulty with the initial Oregon exercise was the attempt to prioritize a large number of services in a relatively short period of time. The first "trial run" release of the list used both cost and effectiveness information to form a cost-utility ratio for each of nearly 1,600 condition-treatment pairs. Although the list was never published, it was widely circulated. Early inspection of the list revealed many inconsistencies. The replacement of some condition-treatment pairs were, in the opinion of the commissioners, counter-intuitive. Almost certainly, these peculiar placements were based on faulty analysis. For example, treatment for thumb-sucking and acute headaches received higher rankings than treatment for AIDS or cystic fibrosis. The problem was not the method but rather the way data were generated by the medical committees. We should not fault the committees for doing a poor job. Indeed, tremendous personal effort went into estimation of treatment effectiveness. The difficulty was that the committee attempted to do several decades of work within the confinement of a few months. Health policy analysts sometimes take 2 to 3 years to thoroughly analyze the expected benefits of a single condition-treatment
pair. For example, Weinstein and Stason engaged in a several year exercise to estimate the cost/utility of the screening and treatment for high blood pressure. A similar analysis might have been performed in a single meeting of the Oregon medical committee. Under this time pressure, one could expect that some analyses would be problematic. Yet in response to criticism the commission gave up the more systematic aspects of the analysis by dropping the cost component.

As part of their evaluation of the Oregon Experiment, the Office of Technology Assessment (OTA) hired clinicians, (two internists and two pediatricians) to evaluate the list. On careful scrutiny, the clinicians identified several problems. For example, they found 21 condition-treatment pairs where the Oregon Group had significantly underestimated the treatment effect. These included medical therapy and thymectomy for myasthenia gravis and medical therapy for chronic bronchitis. The clinicians also found several condition-treatment pairs where the Oregon Group had overestimated the treatment effect. For example, the Oregon Group had rated excision ganglion of tendon or joint as 495. The OTA reviewers felt that the condition was essentially trivial and needed no treatment. The Oregon Group rated as 606 medical therapy for hepatorenal syndrome. Even with this low and unfundable ranking, the OTA reviewers commented that the ranking was too high since the treatment for this condition is regarded as futile. The clinicians were also concerned about some cases in which surgi-
medical therapy was higher on the list than medical therapy. The reviewers argued that medical therapy should always be tried before surgery is employed. Thus, medical therapy should rank higher than surgery.

One of the most serious concerns raised by the clinical reviewers was the inexplicable grouping of some condition-treatment pairs. For example, line 264 was for diseases of white blood cells. However, this category groups together some conditions which are quite trivial with others that are life threatening. Line 640 (testicular hyperfunction) combines a condition which may require no treatment, with Schmidt's syndrome, which is likely to be fatal without treatment.

Another medical concern was the problem of comorbidity. It is often difficult to assess the importance of a condition without knowing the other diseases or disabilities that go along with it. For example, medical therapy for problems in blood clotting is difficult to evaluate because it depends on whether the problem is caused by a transient infection or by a serious disease such as cancer. Sometimes, it was difficult for the consultants to evaluate the condition-treatment pairs since identification of the problem requires treatment. An example includes surgery for peritoneal adhesions (line 508). The difficulty is that the diagnosis of this problem requires a surgical procedure or laparotomy. In addition to these problems, the consultants identified several problems of apparent miscoding or of mismatches between the international classification of disease (ICD-9) and the CPT code matches.7

Advocates from Oregon argued that their medical committee included clinicians with at least equal experience and that their own judgments arose from detailed group dis-
cussion. There were differences of opinion, but the Oregon group saw few justifications for trusting the OTA reviewers over their own committees. Any identifiable differences of opinion could be resolved through further discussion. Several lines of debate were advanced. For example, Oregon providers suggested that it was only OTA’s opinion that chronic bronchitis was underestimated. This condition-treatment pair included smokers cough without other evidence of pulmonary function loss. Alternatively, they did not feel that they overestimated the treatment effect of Schmidt’s syndrome. Although treatment of the syndrome is not effective, individual conditions that make up the syndrome are listed and covered elsewhere on the list. Our point is that determination of treatment effectiveness had a strong subjective component.

The greatest problem facing implementation of the model is lack of scientific data. We need considerably more research on the efficacy of clinical interventions. In the absence of systematic data, subjective judgments are used. The high degree of subjectivity in evaluating medical effectiveness has gone largely unchallenged by philosophical critics of the Oregon proposal.

Commission Reshuffling

One of the most important concerns regarding Oregon’s model is that the commissioners took it upon themselves to reorganize the list. They argued that this was necessary because many of the rank orders were illogical or in other ways inconsistent with the commissioners’ expectations.

In their final report, the Oregon Health Services Commission commented that change in rankings were done rarely. However, a review of the February 1, 1991 versus the May 1, 1991 list suggests that changes in rankings were common. These changes were reviewed in detail by the Congress Office of Technology Assessment. Further, the introduction of the 1992 and 1993 lists saw considerably more reshuffling by the commissioners. According to the Oregon Health Services Commission report, CT pairs were to be ranked by net benefit within category. For example, within the maternal and child health category, it would be expected that the condition with the highest net benefit would be ranked first while the condition with the next highest net benefit would be ranked second, and so on. However, after the rerankings, there was essentially no correlation between ranking and net benefit within some categories. Figure 5 shows a scatter plot comparing net benefit versus ranking. Net benefit was not an important determinant of ranking for services high on the list. However, few CT pairs at the bottom of the list had high net benefit. The OTA analysis examined which services went up and which services went down as a function of the commission review. Services that had high benefit rarely moved far down the list (more than 100 lines). However, services that moved up the list more than 100 ranks were likely to be those that had few benefits. Few of the services that ended up on the lowest end of the list (right portion of Figure 5) had much expected benefit. This subjective readjustment of the list has also received considerably less attention than the use of preference weights.

Resubmission and Acceptance

After the disappointing August 1992 rejection by DHHS, Oregon revised and resubmitted its application. In November 1992 they submitted a modified program that changed the prioritization method to eliminate any consideration of quality of life. The rationale for eliminating this aspect of the methodology was that is was “subjective.” The new method ranked 688 pairs of conditions and treatments in two stages. First, the treatments were subjectively ranked by effectiveness. The criteria for establishing effectiveness included preventing death, returning patients saved from death to an
asymptomatic state, returning symptomatic patients to an asymptomatic state, and average cost of the procedure. Instead of using public preferences, conditions were ranked by medical experts and the commissioners reordered the lists to be consistent with the values they perceived to be representative of Oregon citizens. By eliminating the Quality of Life portion of the model, Oregon answered some of the criticisms in the initial DHHS review. However, they created new problems because they eliminated the possibility of estimating cost per quality adjusted year of life in an objective manner.

Afraid of further criticism, the Oregon commissioners moved items or reassigned the content of many condition-treatment pairs. For example, two of the placements that drew most criticism in the 1991 exercise were life support and medical support for very low birth weight babies and heroic care for end-stage AIDS patients. In both of these cases, evidence suggested that medical care contributed nothing to either quality of life or duration of life. However, many critics made emotional pleas to the media and suggested that deserving citizens would be condemned to death by the plan.

Life support for babies born less than 500 grams and less than 23 weeks gestation was ranked 708 in the 1991 list because children in this condition have an estimated probability of survival of 0.00. To avoid criticism, the commissioners eliminated this line and created a new one for “Low birth weight (under 2,500 grams).” This line also allowed the elimination of another controversial treatment (medical therapy for intraventricular and subarachnoid hemorrhage of fetus or neonate) that was ranked 687 in 1991. The commissioners then placed medical and life support services for low birth weight babies as number 40 on the list. Medical therapy for end stage HIV disease was line 702 on the 1991 list. This drew heavy criticism from groups in advocacy for people with AIDS. However, the placement was based on the evidence that end stage treatment has no effect on quality of life and may even shorten life expectancy because the toxic effects of treatment may exceed their benefits. Although AIDS patients would not get essentially useless treatment, comfort care would be given to these patients because it was near the top of the list. In the revised application, the Oregon commissioners dealt with the problems by simply eliminating the line. An analysis of the revised 1992 list by the Office of Technology Assessment showed that only 25% of the CT pairs on the 1992 list were within 25 lines of where they were in 1991 and 30% of the pairs had moved 100 or more ranks. Change in methodology and commissioner subjective judgment had a substantial effect on ultimate rankings.

Ironically, the new system has greater potential to discriminate against people with disabilities. Using the preference weighted system, it is possible to quantify benefits of treatments that make small quality of life improvements for people with chronic medical problems. However, many treatments for people with disabilities neither improve life expectancy or move people to asymptomatic states. It is precisely these small changes that are most likely to be disregarded.

The Clinton administration has encouraged states to experiment with different approaches to Medicaid. In March 1993, they approved the revised plan and it was implemented in early 1994.

**Discussion**

At least four levels of human judgment were used in the construction of Oregon prioritized list. Among these four, there was substantial justification for the use of the utility rating system.

The August 1992 rejection of the application emphasized the subjective ratings for quality of life. Yet, among the four levels of subjective judgment, only the preference weight component was obtained using a systematic methodology that has replicated other findings. The list that was approved in March of 1993 left considerably more room for subjective biases.
Cost-effectiveness analysis is receiving increasing attention as a approach to resolving health care problems. For example, both the province of Ontario, Canada and the Australian government have endorsed prioritization schemes to make formulary decisions. Application of these methods should rely on the strongest methodologies available. Consumer utilities for health states are an important part of health status and can be measured reliably.

Elimination of preferences from the Oregon resource allocation model was not only misinformed, it was incorrect. It assumed that there would be discrimination against persons with disabilities because treatment could not improve their chronic problems. However, this analysis makes a serious conceptual error. Effectiveness of treatment is based on estimated course of the illness with and without treatment. A treatment that sustains life, even without improvements in quality of life, produces very substantial benefits. For example, suppose a person is an accident that leaves him or her in a state rated 0.5 on a 0 to 1.0 scale. Further suppose that a treatment will maintain them at this level while absence of treatment will result in death. According to the Oregon model, the treatment will produce 0.50 (calculated as 0.50 - 0) for each year the person remains in that state. That is a powerful treatment effect in comparison to most alternatives. The crucial element is that the treatment works. The system does attempt to exclude treatments that neither extend life nor make patients better in contrast to those not receiving treatment. In other words, the targets for elimination are only treatments that use resources and do not make a difference.

The August 1992 DHHS decision also misrepresented the meaning of quality of life scores. They assumed that having a low quality of life score was discriminatory because people with disabilities and those without disabilities would not be rated the same. Alternatively, the statement is contradictory with the notion that people with disabilities need medical services. People who are at optimum health (1.0 on the QWB scale) may need fewer services than those who occupy lower levels. Quantifying these differences allows us to set priorities for future resource allocation. If, for the sake of argument, we decide to score people with disabilities 1.0 and they with stay there without treatment, it would follow that we should not provide any services for these individuals because they have already achieved the optimum level of wellness. Scores lower that 1.0 suggest that resources should be used to improve these conditions. This in no way implies that the lives of people with disabilities are valued less that those of people without disabilities. The most important point is that judgment was not made on the basis of rating of disability, but rather on the basis of expected improvement as a result of treatment. The only reference cited when former Secretary of Health and Human Services Louis Sullivan rejected the Oregon application in 1992, was an unpublished paper by Hadorn.28 However, Hadorn’s conclusion was ignored. He noted that “It is the change in quality of life, or net benefit realized from treatment that matters, not the point-in-time quality of life of the patient”.

Because of the DHHS challenge, Oregon eliminated the quantitative utility data and developed a prioritized list that excluded quality of life decisions. In doing so they gave up the most replicable part of their methodology. Subjective, judgments are an unavoidable part of priority setting. Policymakers must acknowledge what aspects of their decisions depend on human judgment and they must use the most reliable and valid methods to capture them.

Acknowledgments

The author thanks Maria Hewitt, Paige Spies-Meltzer, Darren Coffman, Lucy Lord-Lippincott and Bob DiPrate for their comments on an earlier draft of this paper.
References


4. Eddy DM. What care is "essential"? What services are "basic"? JAMA 1991; 265:782.


