Chapter 7

Measuring Health Outcomes for Resource Allocation

Robert M. Kaplan

The Health Care debate has been naively conceptualized as a simple struggle between "good guys and bad guys." When President Clinton addressed Congress in September 1993, he suggested that the good guys were physicians, nurses, and hospitals. The bad people were lawyers, insurance companies, and holders of MBA degrees. I argue that the distinctions are not so clear. To place this in context, we must examine the problems in health care.

THE THREE *A*s

Some people argue that there is no health care crisis. For example, some elected representatives have suggested that we need only minor changes in our current system. Yet their constituents frequently testify with challenging stories. For example, a family with a sick child may not be able to get health insurance because the illness is too expensive. Testimony on health care reform produced countless stories of families seriously distressed because they could not obtain health care. When Hillary Clinton headed the task force on health care reform, she received literally thousands of letters describing personal complications in relation to the health care system. One case was that of a family in Cleveland who had three daughters. The first was healthy, but the other two were born with serious chronic illnesses that required constant care. The parents were employed but were unable to purchase private health insurance. In a communication to Mrs. Clinton, the mother noted, "I finally realized how futile it was when I was talking to an insurance agent explaining our medical problems with our daughters and he looked at me and he said, you don't understand; we don't insure burning houses'' (from remarks by the president and the first lady, Health Care Forum, Century Village East, Deerfield Beach, Florida, distributed over Internet, March 24, 1994). There are countless cases in which people in need of help are unable to get basic services. Some of the issues in health care might be described by the ''Three *A*s'' (Kaplan, 1993a, 1993b). The first is affordability—health care costs too much. A second issue is *access*. We have too many people who do not have a regular source of health care. They are either uninsured or underinsured. In addition, there are other barriers to health care. For example, some people do not have transportation. The third is *accountability*. Despite the fact that we spend more on health care than any other country, we have failed to document that the care that we provide makes a difference (Eddy, 1994; Kaplan, 1993b; Wennberg, 1994).

The three *A*s are connected. For example, providing access to everyone will resolve the problem of access but may make the problem of affordability severer because costs will go up. To solve the health care crisis, the three *A*s must be addressed simultaneously. We have to find ways to reimburse for services that make people better and not to use resources for services that do not work. Refusing to spend money on nonefficacious services may save enough money to expand access to basic care for people who were currently underinsured and uninsured. To explore these issues in more detail, each of the three *A*s will be addressed briefly.

Affordability

Some people argue that we do not have to worry about expanding health care costs. Reinhardt (1993) has questioned whether we need to limit health care spending. Spending more money on health care in relation to other public services, such as education, defense, and so forth, may be our way of making resource allocation decisions. However, it is important to recognize that if we care about other services, we must also address the health care problem. As more resources are used for health care, fewer are available to improve schools, support a national defense, and so on.

The exponential growth in health care costs over the last 50 years is shown graphically in Figure 7.1. However, it is important to focus attention on the two right-hand bars of the figure. Health care costs have risen exponentially since 1940. Between 1990 and today, there has been substantial de facto health care reform. In the last few years, there have been significant reductions in hospital admissions and lengths of hospital stay.

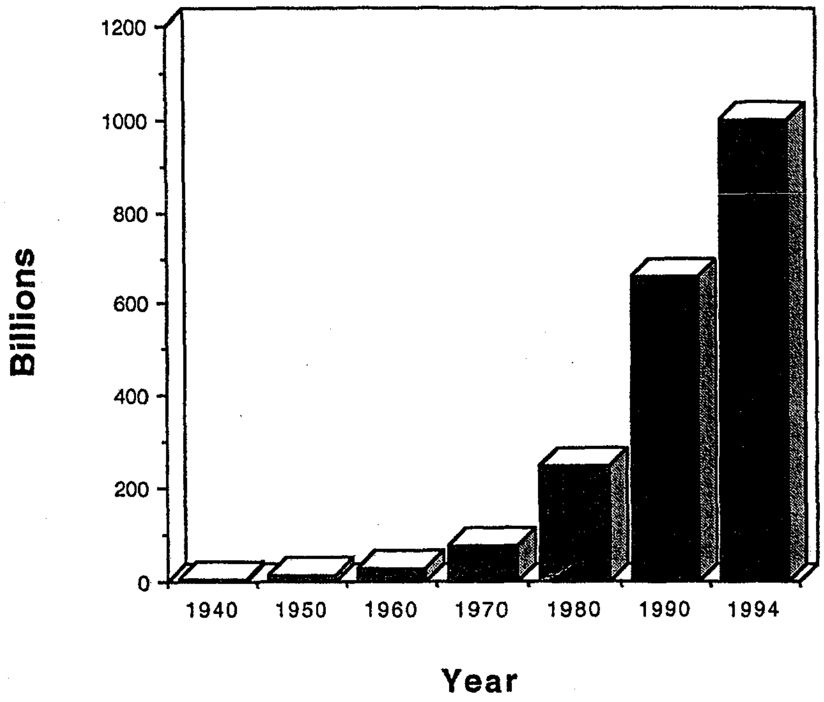

FIGURE 7.1 Costs of health care in the United States, 1940–1994.

Nevertheless, health care costs in the United States have risen an estimated $350 billion since 1990. In 1990, experts estimated that health care costs would reach $1 trillion by the turn of the century. Current estimates indicate that we may hit $1 trillion by 1997.

To put these expenditures in other terms, consider an agency like the National Heart Lung and Blood Institute, one of the best funded institutes in the National Institutes of Health. If we had capped our health care expenditures in 1990, the savings each day would pay for the entire National Heart Lung and Blood Institute for an entire year. As we spend more and more of our gross domestic product on health care, we have less for other sectors of the economy. In particular, the purchase of consumer goods may be slowed (Warshawsky, 1994). Yet the purchase of these goods drives the rest of the economy (Fuchs, 1990). So, for example, as our expenditures in health care have increased relative to other countries, our balance of trade has gone in the negative direction (Levit et al., 1994). As an increasing percentage of the economy is devoted to health care, the rest of the economy will be starved for capital, and there may be economic decline and a reduction in the standard of living (Warshawsky, 1994).

To summarize my first argument, we cannot continue to spend this way. We have to set some sort of expenditure ceiling. This will most likely be accomplished by government intervention or through a regulated competition.

In some ways our problems are not that different from other countries. We are spending about 14.5% of our gross domestic product on health care while most other countries are below 10%. However, no country can do all the things that they want to do in health care. All of them face difficult decisions. Rich countries may be able to set their allocation thresholds higher, but all countries face choices (Williams, 1988). For example, many European countries have formal policies to control costs. The United Kingdom accomplishes cost control through a government-run system with explicit policies on what services will be covered. The Netherlands restricts growth in health care to a fixed portion of the gross domestic product (Kirkman-Liff, 1991).

Access

Despite these high expenditures, the United States is unique among developed countries because we have large numbers of people who are uninsured or underinsured. An estimated 38 million people in the United States have no health insurance, and 58 million are uninsured for at least part of each year (Health Care Financing Administration, 1994). Most of the uninsured are employed, and the poor and minority group members are disproportionately represented among those without coverage.

Opportunity Cost.

Opportunity costs are the foregone opportunities that are surrendered as a result of using resources to support a particular decision. To put it in other terms, if we spend a lot of money in one sector of health care, we necessarily spend less money elsewhere. This problem is heightened by the tendency to follow a "rule of rescue." The rule of rescue, in the words of philosophers, is a moral obligation to provide rescue services whenever saving a life is a possibility (Hadorn, 1991). However, the decision to invest in rescue may necessarily mean two decisions have been made. With limited resources, a decision to perform an expensive liver transplantation surgery for one person often means giving up the opportunity to perform less expensive services for many people. Sometimes a liver transplant has extremely limited potential for producing a health benefit, while the preventive services that are neglected have substantial potential to help other people. Indeed, the U.S. health care system is rich with applications of the rule of rescue. Large investments in dramatic and often futile care have resulted in the unfunding or underfunding of substantial opportunities in primary care, mental health, and prevention. For example, the state of Illinois passed a 1985 bill that guaranteed reimbursement of up to $200,000 for any citizen who needed an organ transplant. At the same time, more than 60% of black children in Chicago's inner cities did not receive routine medical care and were not even immunized against common diseases, such as polio. In 1990, Florida's Governor Martinez committed $100,000 to a heroic attempt to save the life of a single child who had nearly drowned in a swimming pool accident. All experts agreed that the case was futile. While the governor received great acclaim for his compassion, thousands of Florida children were denied basic services through Florida's underfunded Medicaid program (Kitzhaber, 1990). When funds are directed toward rescue, prevention programs are typically the first victim of the revenue shortfall.

Cost Shifting.

Some members of our society feel that we should not care about the uninsured. They believe that the uninsured are typically poor people who do not take care of themselves and their failure to plan should not be our problem. However, it is not necessarily true that people without insurance do not get care. In fact, they do get health care by going to emergency rooms, but they are often unable to pay for their services. When a patient is unable to pay, the hospital still has to reimburse its nurses, it still has to support its pathology laboratories, and so forth. Someone also has to pay for the surgeons and other physicians.

It is not necessarily true that patients who are uninsured get free care. In fact, the costs are just shifted. When an uninsured patient comes to the hospital and cannot pay, his or her charges are shifted to insured patients. When the insured patients get charged, insurance rates go up. There are a whole series of shifts. As a result, charges in fancy suburban clinics, where most patients are insured, may be lower than they are in inner-city hospitals, where a high percentage of patients are uninsured.

The reason is that the people who are able to pay in inner-city hospitals are subsidizing a larger number of patients who are uninsured. A corollary is that insurance rates should be higher in areas of the country where there are high rates of people who are medically uninsured. In fact, this seems to be so. For example, in cities, such as Los Angeles and Miami, where the rates of medical insurance are low, charges to businesses for health insurance are higher than in cities, such as Minneapolis and Seattle, where the rates of uninsurance are lower (Kaplan, 1993b).

In summary, cost shifting suggests that costs are not avoided; they are just charged to someone else. However, the uninsured obtain services in a costly way, because they are often more seriously ill at entry to the system and are cared for inefficiently through emergency rooms. In Hawaii, providing universal coverage has actually decreased health care costs (Lewin & Sybinsky, 1993). The following sections explore the accountability problem in more detail. Specifically, it will be suggested that better accountability can contribute to the solution to the health policy crisis. Accountability is focused on using resources to make people healthier. By using resources more efficiently, it may be possible to achieve better health outcomes, to save money, and to use the savings to expand access.

Accountability

Health care must achieve greater accountability. Part of the problem in producing greater accountability has been that the health care system has not quantified its major product. The health status benefit of most of the services delivered by the system is rarely measured or reported.

To understand health outcomes, it is necessary to build a comprehensive theoretical model of health status. This model includes several components. The major aspects of the model include mortality (death) and morbidity (health-related quality of life). We have suggested elsewhere that diseases and disabilities are important for two reasons. First, illness may cause the life expectancy to be shortened. Second, illness may make life less desirable at times before death (health-related quality of life) (Kaplan & Anderson, 1988a, 1988b, 1990).

A GENERAL HEALTH POLICY MODEL (GHPM)

Over the last two decades, a group of investigators at the University of California, San Diego, has developed the GHPM. Central to the General

Health Policy Model is a general conceptualization of quality of life. The model separates aspects of health status and life quality into distinct components. These are life expectancy (mortality), functioning and symptoms (morbidity), preference for observed functional states (utility), and duration of stay in health states (prognosis).

Components of GHPM

Mortality.

A model of health outcomes necessarily includes a component for mortality. Indeed, many public health statistics focus exclusively on mortality through estimations of crude mortality rates, age-adjusted mortality rates, and infant mortality rates. Death is an important outcome that must be included in any comprehensive conceptualization of health.

Morbidity.

In addition to death, quality of life is also an important outcome. The GHPM considers functioning in three areas: mobility, physical activity, and social activity. Descriptions of the measures of these aspects of function are given in many different publications (Kaplan & Anderson, 1988a, 1988b). Most public health indicators are relatively insensitive to variations toward the well end of the continuum. Measures of infant mortality, to give an extreme example, ignore all individuals capable of reading this chapter because they have lived beyond 1 year following their births. Disability measures often ignore those in relatively well states. For example, the RAND Health Insurance Study reported that about 80% of the general populations have no dysfunction. Thus, they would estimate that 80% of the population is well. Our method asks about symptoms or problems in addition to behavioral dysfunction (Kaplan et al., 1976). In these studies, only about 12% of the general population report no symptoms on a particular day. In other words, health symptoms or problems are a common aspect of the human experience. Some might argue that symptoms are unimportant because they are subjective and unobservable. However, symptoms are highly correlated with the demand for medical services, expenditures on health care, and motivations to alter lifestyles. Thus, we believe that the quantification of symptoms is important.

Utility (Relative Importance).

Given that various components of morbidity and mortality can be tabulated, it is important to consider their relative importance. For example, it is possible to develop measures that detect minor symptoms. Yet, because these symptoms are measurable, it does not necessarily mean they are important. A patient may experience side effects of a medication but be willing to tolerate them because the side effects are less important than the probable benefit that would be obtained if the medication is consumed. Not *all* outcomes are equally important. A treatment in which 20 of 100 patients die is not equivalent to one in which 20 of 100 patients develop nausea. An important component of the GHPM attempts to scale the various health outcomes according to their relative importance. This exercise adds the "quality" dimensions to health status. In the preceding example, the relative importance of dying would be weighted more than developing nausea. The weighting is accomplished by rating all states on a quality continuum ranging from 0 (for dead) to 1.0 (for optimum functioning). These ratings are typically provided by independent judges who are representative of the general population (Kaplan et al., 1978). Using this system, it is possible to express the relative importance of states in relation to the life-death continuum. A point halfway on the scale (0.5) is regarded as halfway between optimum function and death. The quality-of-life weighting system has been described in several different publications (Kaplan et al., 1976, 1978, 1979). Although there are differences between cultures and religious groups on the definition of wellness, preferences for health states are remarkably constant across demographic and cultural groups (Kaplan, 1994)

Prognosis.

Another dimension of health status is the duration of a condition. A headache that lasts 1 hour is not equivalent to a headache that lasts 1 month. A cough that lasts 3 days is not equivalent to a cough that lasts 3 years. In considering the severity of illness, duration of the problem is central. As basic as this concept is, most contemporary models of health outcome measurement completely disregard the duration component. In the GHPM, the term *prognosis* refers to the probability of transition among health states over the course of time. In addition to consideration of duration of problems, the model considers the point at which the problem begins. A person may have no symptoms or dysfunction currently, but may have a high probability of health problems in the future. The prognosis component of the model takes these transitions into consideration and applies a discount rate for events that occur in the future. Discount rates are used to value resources and health outcomes differently if the onset is delayed as opposed to immediate. A headache that will begin a year from now may be less of a concern than a headache that will start immediately.

Quality of Well-Being Scale (QWB)

The QWB is one of several different approaches for computing quality-adjusted life years (QALY) (Kaplan & Anderson, 1988b). Using this method, patients are classified according to objective levels of functioning. These levels are represented by scales of mobility, physical activity, and social activity (see Table 7.1). In addition to classification into these observable levels of function, individuals are also classified by the symptom or problem that they found to be most undesirable (see Table 7.2). On any particular day, nearly 80% of the general population is optimally functional. However, fewer than half of the population experience no symptoms. Symptoms or problems may be severe, such as serious chest pain, or minor, such as having to take medication or adhering to a prescribed diet for health reasons.

Human value studies have been conducted to place the observable states of health and functioning onto a preference continuum for the desirability of various conditions, giving a ''quality'' rating between 0 for death and 1.0 for completely well. These weights are shown in Tables 7.1 and 7.2. A QALY is defined as the equivalent of a completely well year of life, or a year of life free of any symptoms, problems, or health-related disabilities. The well-life expectancy is the current life expectancy adjusted for diminished quality of life associated with dysfunctional states and the durations of stay in each state. It is possible to consider mortality, morbidity, and the preference weights for the various observable states of function. Table 7.3 gives formulas and an example of the calculation of the QWB for a patient with chronic lung disease. The model quantifies the health activity or treatment program in terms of the years that it produces or saves.

A mathematical model integrates components of the model to express outcomes in a common measurement unit. Using information on current functioning and duration, it is possible to express the health outcomes in terms of equivalents of well years of life. The model for point in time QWB is

$$\begin{aligned} \text{QWB} = 1 \quad & (\text{observed morbidity} \times \text{morbidity weight}) \\ & (\text{observed physical activity} \times \text{physical activity weight}) \\ & (\text{observed social activity and social activity weight}) \\ & (\text{observed symptom/problem} \times \text{symptom/problem weight}) \end{aligned}$$

The net cost/utility ratio is defined as

TABLE 7.1 Quality of Well-Being/General Health Policy Model: Elements and Calculating Formulas (Function Scales, With Step Definitions and Calculating Weights)

Step no.	Step definition	Weight
	Mobility Scale	
5	No limitations for health reasons	–.000
4	Did not drive a car, health related; did not ride in a car as usual for age (younger than 15 years), health related, or did not use public transportation, health related; or had or would have usedmore help than usual for age to use public transportation, health related	–.062
2	In hospital, health related	–.090
	Physical Activity Scale	
4	No limitations for health reasons	–.000
3	In wheelchair, moved or controlled movement of wheelchair without help from someone else; or had trouble or did not try to lift, stoop, bend over, or use stairs or inclines, health related; or limped, used a cane, crutches, or walker, health related; or had any other physical limitation in walking, or did not try to walk as far as or as fast as other the same age are able, health related	–.060
1	In wheelchair, did not move or control the movement of wheelchair without help from someone else, or in bed, chair, or couch for most or all of the day, health related	–.077
	Social Activity Scale	
5	No limitations for health reasons	–.000
4	Limited in other (e.g., recreational) role activity, health related	–.061
3	Limited in major (primary) role activity, health related	–.061
2	Performed no major role activity, health related, but did perform self-care activities	–.061
1	Performed no major role activity, health related, and did not perform or had more help than usual in performance of one or more self-care activities, health related	–.106

$$\frac{\text{Net cost}}{\text{Net QWB} \times \text{duration in years}} = \frac{\text{Cost of treatment} - \text{cost of alternative}}{[\text{QWB}_2 - \text{QWB}_1] \times \text{duration in years}}$$

Where QWB_2 and QWB_1 are measures of quality of well-being taken before and after treatment.

TABLE 7.2 Quality of Well-Being/General Health Policy Model: Symptom/Problem Complexes (CPX) With Calculating Weights

CPX No.	CPX Description	Weights
1	Death (not on respondent's card)	–.727
2	Loss of consciousness such as seizure (fits), fainting, or coma (out cold or knocked out)	–.407
3	Burn over large areas of face, body, arms, or legs	–.387
4	Pain, bleeding, itching, or discharge (drainage) from sexual organs—does not include normal menstrual (monthly) bleeding	–.349
5	Trouble learning, remembering, or thinking clearly	–.340
6	Any combination of one or more hands, feet, arms, or legs either missing, deformed (crooked), paralyzed (unable to move), or broken—includes wearing artificial limbs or braces	–.333
7	Pain, stiffness, weakness, numbness, or other discomfort in chest, stomach (including hernia or rupture), side, neck, back, hips, or any joints or hands, feet, arms, or legs	–.299
8	Pain, burning, bleeding, itching, or other difficulty with rectum, bowel movements, or urination (passing water)	–.292
9	Sick or upset stomach, vomiting or loose bowel movement, with or without chills, or aching all over	–.290
10	General tiredness, weakness, or weight loss	–.259
11	Cough, wheezing, or shortness of breath, *with* or *without* fever, chills, or aching all over	–.257
12	Spells of feeling, upset, being depressed, or of crying	–.257
13	Headache, or dizziness, or ringing in ears, or spells of feeling hot, nervous or shaky	–.244
14	Burning or itching rash on large areas of face, body, arms, or legs	–.240
15	Trouble talking, such as lisp, stuttering, hoarseness, or being unable to speak	–.237
16	Pain or discomfort in one or both eyes (such as burning or itching or any trouble seeing after correction	–.230
17	Overweight for age and height or skin defect of face, body, arms, or legs, such as scars, pimples, warts, bruises, or changes in color	–.188
18	Pain in ear, tooth, jaw throat, lips, tongue; several missing or crooked permanent teeth—includes wearing bridges or false teeth; stuffy, runny nose; or any trouble hearing—includes wearing a hearing aid	–.170
19	Taking medication or staying on a prescribed diet for health reasons	–.144
20	Wore eyeglasses or contact lenses	–.101
21	Breathing smog or unpleasant air	–.101
22	No symptoms or problem (not on respondent's card)	–.000
23	Standard symptom/problem	–.257
X24	Trouble sleeping	–.257
X25	Intoxication	–.257
X26	Problems with sexual interest or performance	–.257
X27	Excessive worry or anxiety	–.257

TABLE 7.3 Calculating Formulas and Example of QWB for Patient With Lung Disease

Calculating formulas

Formula 1. Point-in-time well-being score for an individual (W):

$$W = 1 + (\mathrm{CPXwt}) + (\mathrm{MOBwt}) + (\mathrm{PACwt}) + (\mathrm{SACwt})$$

where "wt" is the preference-weighted measure for each factor, and CPX is symptom/problem complex. For example, the W score for a person with the following description profile may be calculated for one day as follows:

CPX-11	Cough, wheezing or shortness of breath, with or without fever, chills, or aching all over	–.257
MOB-5	No limitations	–.000
PAC-1	In bed, chair, or couch for most or all of the day, health related	–.077
SAC-2	Performed no major role activity, health related but did perform self-care	–.061

$$W = 1 + (-.257) + (-.000) + (-.077) + (-.061) = .605$$

Formula 2. Well years (WY) as an output measure:

$$\mathrm{WY} = [\text{No. of persons} \times (\mathrm{CPXwt} + \mathrm{MOBwt} + \mathrm{PACwt} + \mathrm{SACwt}) \times \text{Time}]$$

Consider, for example, a person who is in an objective state of functioning that is rated by community peers as 0.5 on a 0 to 1.0 scale. If the person remains in that state for 1 year, he or she would have lost the equivalent of one half of 1 year of life. Thus, a person limited in activities who requires a cane or walker to get around the community would be hypothetically rated at 0.50. If they remained in that state for an entire year, such an individual would lose the equivalent of one-half year of life. However, a person who has the flu may also be rated as 0.50. In this case, the illness might only last 3 days, and the total loss in well years might be 3/365 × 0.50, which is equal to 0.004 well years. This may not appear as significant an outcome as a permanent disability. But suppose that 5,000 people in a community get the flu. The well years lost would then be 5,000 × 0.004, which is equal to 20 years. Now suppose that a vaccination has become available and that the threat of the flu can be eliminated by vaccinating the 25,000 people in the community. The cost of the vaccine is $5 per person, or $125,000. The cost/utility of the program would be

$$\frac{\$125{,}000 \text{ (cost)}}{20 \text{ years (utility)}} = \$6{,}250/\text{well year}$$

Ideally, the outcomes are assessed in systematic clinical studies. For example, patients might be randomly assigned to a treatment or to a control group and followed at regular intervals. Well-designed studies take both outcome and duration into consideration and the benefit is shown in QALYs. Although the model does not depend on any particular experimental design, the weight given to a particular finding might be lower for nonsystematic experiments.

The GHPM has been used in a wide variety of population studies (Anderson et al., 1989; Erickson et al., 1989). In addition, the methods have been used in clinical trials and studies to evaluate therapeutic interventions in a wide range of medical and surgical conditions. These include chronic obstructive pulmonary disease (Kaplan et al., 1984), acquired immunodeficiency disorder (Kaplan et al., 1995), cystic fibrosis (Orenstein et al., 1989), diabetes mellitus (Kaplan et al., 1987), atrial fibrillation (Ganiats et al., 1993), lung transplantation (Squier et al., 1994), arthritis (Bombardier et al., 1986; Kaplan et al., 1988), cancer (Kaplan, 1993c), Alzheimer's disease (Kerner et al., 1996), Sinus disease (Hodgkin, 1994), and a wide variety of other conditions (Kaplan, 1993b). Further, the method has been used for health resource allocation modeling and has served as the basis for an innovative experiment on rationing of health care by the state of Oregon.

Is the Model Applicable to Mental Health?

Despite widespread interest in the model among practitioners in many different specialties, the concept of a QALY has received little attention in the mental health fields. We believe that this reflects the widespread belief that mental health and physical health outcomes are conceptually distinct. Ware and Sherbourne (1992) emphasized that mental and physical health are different constructs, and that attempts to measure them using a common measurement strategy is like comparing apples to oranges. We recognize the distinction between mental health and physical health outcomes, and acknowledge the need to measure the effects of treatment using different units. However, we also suggest that a common measurement strategy is required so that the productivity of mental and physical health providers can be compared directly.

Several years ago, Kaplan and Anderson (1988a) argued that there are many similarities in mental and physical health outcomes. The QWB system includes the basic dimensions of observable functioning, symptoms, and duration. Mental health problems, like physical health problems, can be represented by symptoms and disrupted role functioning. Consider some examples. Suppose that a patient has the primary symptom of a cough. If the cough does not disrupt role function, the QWB score might show a small deviation from 1.0. If the cough is more serious and keeps the person at home, the QWB score will be lower. If the cough is severe, it might limit the person to a hospital and may have serious disruptive effects on role functioning. This would necessitate an even lower QWB score. Coughs can be of different duration. A cough associated with an acute respiratory infection may have a serious impact on functioning that may last only a short period. This would be indicated by a minor deviation in well years. A chronic cough associated with obstructive lung disease would be associated with significant loss of QALY because duration is a major component of the calculation.

Now consider the case of a person with depression. Depression may be a symptom reported by a patient just as a cough is reported by other patients. Depression without disruption of role function would cause a minor variation of wellness. If the depression caused the person to stay at home, the QWB score would be lower. Severe depression might require the person to be in a hospital or special facility, and would result in a lower QWB score. Depressions, like coughs, are of different durations. Depression of long duration would cause the loss of more QALYs than would depression of short duration.

Some evidence supports the validity of the QWB in studies of mental health. One recent study evaluated the validity of the QWB as an outcome measure for older psychotic patients. Seventy-two psychotic patients and 28 matched controls from the San Diego Veterans Affairs Medical Center completed the QWB, the Structured Clinical Interview for the DSM-III-R patient version (SCID-P), Scales for the Assessment of Positive and Negative Symptoms (SAPS and SANS), and the Global Severity Index (GSI) from the Brief Symptom Inventory (BSI) were administered to all subjects. The QWB correlated with the SANS –.52 (*p*<.001) with the SAPS –.57 (*p*<.001) and the GSI –.62 (*p*<.001). Patients and controls were significantly different on the QWB. We also identified a linear relationship between QWB and severity of illness (as classified by the SANS and the SAPS). In addition, component scores of the QWB (i.e., mobility, physical activity, social activity, and worst symptom) were significantly lower among patients compared with controls and declined systematically as psychiatric symptoms increased (Patterson et al., 1996).

Using the GHPM, it is possible to estimate the benefit of any health care intervention in terms of the QALY the treatment produces. Suppose, for example, that a treatment for anxiety elevates patients from a level of .65 to a level of .75. Suppose, further, that this treatment benefit lasted for 1 year. Each patient would gain .1 QALY (.75 − .65 = .10 × 1 year = .1 QALY) for each year the benefit was observed. The treatment benefit would be expressed in terms of general QALY units. The productivity of the providers could be compared with providers in other areas of health care. All providers in health care use resources. Dividing the cost of a treatment by the QALY productivity provides the cost-utility ratio. Measuring mental health productivity in QALY units would allow the assessment of investments in mental health services to be compared directly with those in other aspects of health care.

In summary, the general QWB measure has evidence for validity in a variety of different specific diseases. It has been shown to be responsive to change, and its application has been found to be feasible in a wide variety of different populations. We will now turn to applications of the measure in clinical studies and public policy making.

Applications of the GHPM: Potential and Problems

Resource Allocation Decisions.

One of the advantages of the GHPM is that it allows for comparison between different types of services. QALY are a common metric, and all providers in the health care system have the common objectives of increasing length of life and improving quality of life. The general model allows evaluations of the production function for each of these specialties in comparison with the resources that they use. Allocating resources based on systematic data has been proposed by several different governments (Neumann & Johannesson, 1994). For example, the Australian government now requires evidence on effectiveness, as do a variety of European governments (Freund, Evans, Henry & Dittus, 1992). Canada has officially proposed the QALY as a basis for making decisions about which drugs will be purchased by the different provinces (Detsky, 1991). This approach has also been considered in the United Kingdom (Williams, 1988). Perhaps the most interesting experience in the United States has been in the state of Oregon. Oregon attempted to prioritize the cost/utility of different health services in an experiment with their Medicaid Program. One of the innovative features of the Oregon experiment was the attempt to put mental health services and other health services into the same prioritization. In an earlier version of the list, the cutoff for funding was roughly 600. At the top of the list were services, such as treatment for rumination disorder of infancy, treatment for delirium resulting from the use of psychoactive substance, or treatment for a single episode of major depression. In the middle of the list were services, such as psychotherapy for anxiety disorders and panic disorders, and treatment for conversion disorders in childhood. These services would clearly be funded by the program. However, at the bottom of the list were services, such as psychotherapy for antisocial personality disorder, psychotherapy for transsexualism, and psychotherapy for pica (see Table 7.4).

Black Hats and White Hats.

It is popular to conceptualize the health care debate as a fight between the good and bad people. For example, the American Medical Association funded a series of ads in *Time*, *Newsweek*, and elsewhere saying, "Who should make medical decisions, M.D.s or M.B.A.s?" All readers are supposed to know the answer: The former wear white hats and the latter wear black hats. However, this situation is somewhat more complicated. In the early days of the Medicare Program, physicians were reimbursed for essentially any service they wanted to deliver. Although this has changed somewhat in the last decade, most of the history of the Medicare Program provided physicians opportunities to bill for expensive services with little external review. There is overwhelming evidence that providing rich reimbursement for high expense medical treatments led to the overuse of some services (Hillman, 1994). For example, Figure 7.2 compares the rate of use for several services in the United States, Germany, and Canada. For a wide variety of services including magnetic resonance imaging, lithotrypses, radiation therapies, organ transplantations, cardiac catheterizations, and open heart surgeries, the United States performs more of these services than the other two countries. Yet despite this, we have no evidence that Americans live longer than people in other industrialized countries. In fact, our infant mortality rate is not lower, but it is somewhat higher than other Westernized countries.

Once again, the real challenge is linking the three *A*s together. The U.S. system is expensive (affordability), and it is assumed that we are buying more services to obtain better outcomes. In fact, we have been unable to show that there has been greater value for the money we have spent (accountablility). Further, so much is spent on services, resources are not available to care for those who are less advantaged (access). For

TABLE 7.4 Examples of Mental Health Items From Oregon Integrated List

Rank	Item
76	Rumination Disorder of Infancy
99	Delirium from Psychoactive Substance Use
108	Major Depression, Single Episode
152	Youth Abuse of Hallucinogen, Amphetamine, or Alcohol
347	Anxiety Disorder
348	Panic Disorder
353	Conversion Disorder, Child
760	Antisocial Personality Disorder
761	Transsexualism
763	Pica

example, the costs of Medicaid in most states have escalated in recent years. Financially strapped states have had no choice but to restrict services for Medicaid recipients.

As mentioned earlier, one of the common arguments is that doctors (white hats) are offering too many unnecessary services, but it is really the fault of the lawyers (black hats). Doctors know that many of the services and tests are unnecessary, but, if they did not perform them, they would be at risk of being sued by lawyers. This practice is often called "defensive medicine" because the doctors are defending themselves against lawyers. Certainly, defensive medicine is practiced in the United States. However, there are also reasons to believe that financial interest in offering unnecessary services has also contributed to overuse. One example comes from a study by Hillman and colleagues (1990). This study considered the cost of an episode of care provided by family practitioners. This study was an observational analysis of patients cared for by family doctors who were carefully matched on a variety of different variables. The independent variable in the study was whether or not the doctor owned his or her own X-ray equipment. The results of the study are summarized in Figure 7.3. As the figure shows, the cost per episode of care for upper respiratory infections, pregnancy, back pain, or difficulty urinating, were all significantly higher for patients whose attending doctors owned their own X-ray machines (self-referred) in comparison with family doctors who referred their patients to radiologists to receive these exact same services. In these cases, the risk of being sued should have been the same. What differs is the pecuniary interest in using X-ray tests.

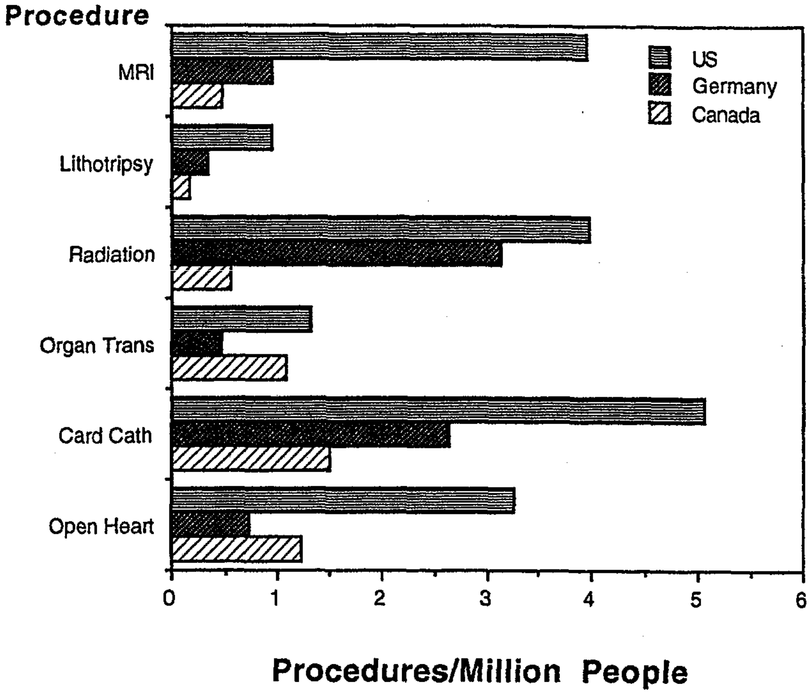

FIGURE 7.2 Rates of various medical procedures in the United States, Germany, and Canada.

Source: Adapted from *Health Care in the Nineties*, Blue Cross, 1990, p. 18—Canada data, 1989; German and U.S. data, 1987

It is widely believed that malpractice lawsuits are a major component in total health care costs. Analyses suggest that malpractice costs are not a big piece of total health care expenditures. For example, malpractice premiums are about 1% of total health care costs. Thus, paying all practice premiums from some other source would reduce health care costs by only 1%. Providers often counterargue that, in fact, it is really not malpractice premiums that cause costs to be high. Instead, it is all the tests, some of which are unnecessary, that are required to protect oneself from litigation. Although experts disagree on the exact costs (Hudson, 1990), one estimate

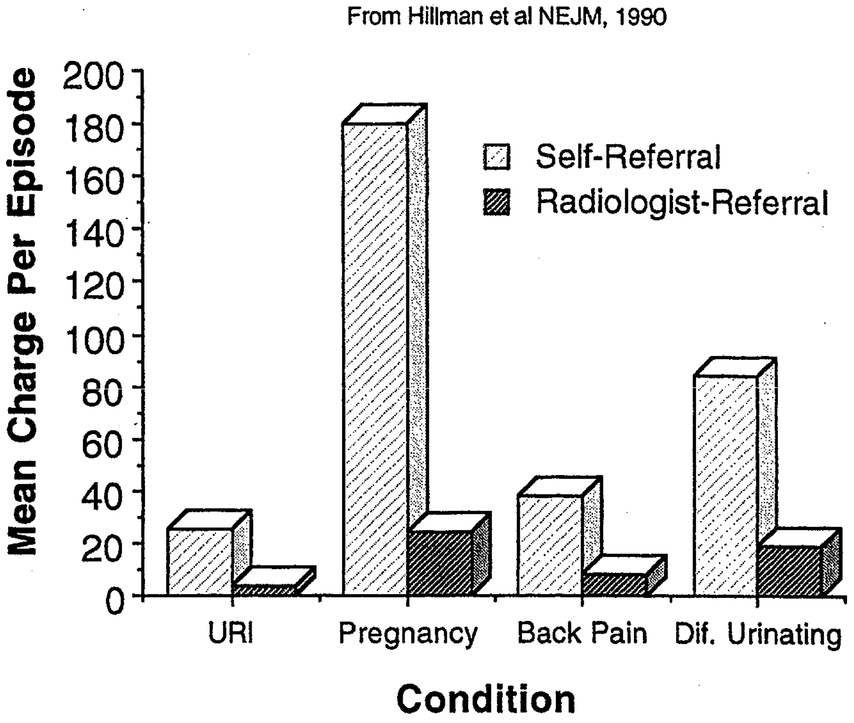

FIGURE 7.3 **Comparison of cost per episode of care for patients cared for by family physicians who owned their own X-ray equipment (self-referred) or those who referred to radiologists, for four diagnoses. (Adapted from Hillman et al., 1989)**

Source: From Hillman et al., *NEJM*, 1990.

is that for each $1 spent on malpractice premiums, $5 are spent on defensive medicine. Multiplying the estimated $9 billion spent on malpractice premiums by 5 yields $45 billion. Although, this is a substantial amount of money, it is still less than 5% of total health care costs. In contrast, consider expenditures on services for which there is no efficacy data. This has been analyzed by investigators at the RAND Corporation. They suggest that we are spending between $280 billion and $450 billion per year (adjusting for inflation) for services for which there is no evidence in the literature that the treatment works. In other words, as many as 30% to 50% of all health care services might be eliminated with no consequent effect on health status (Brook & Lohr, 1987). These procedures may benefit patients in theory, but there is no current documentation that they make people better.

The Gray Hats.

Many professional groups have an interest in health care reform. Literally, billions of dollars are at stake. Various professional organizations and public charities have prepared communications and testified on the value of their contribution. Psychology is not unlike other groups. APA, like other organizations, met with the first lady, who chaired the president's task force on health, to offer evidence of the value of psychology. Virtually every professional health care organization argued that their activities should not be neglected, but few organizations have considered the needs of their competitors or the needs of the nation as a whole.

To place this in prospective, consider the lobbying efforts of private nonprofit charities. These are particularly important because, by virtually every standard, these organizations would be considered the good guys (white hats). The Arthritis Foundation has publicly stated that health care reform must improve outcomes research, effectiveness research, and related activities. However, the Arthritis Foundation also argues that health care reform must make rheumatologic and orthopedic surgical care available regardless of economic considerations (Arthritis Foundation, 1992). In other words, they are arguing that rheumatologists and orthopedic surgeons should be reimbursed for everything that they do. The Arthritis Foundation provides a template for people to write to Congress and offers instructions for people to testify on their behalf. Here the foundation's statement becomes ambiguous. In their public statements, the foundation urges more outcomes research. However, in the template for testimony they suggest that the Clinton Plan overemphasizes outcomes, effectiveness, and clinical trials research. Instead, they argue that the plan does not recognize the cost-saving potential of basic science research. In particular, they worry about the support of molecular biology with the hope it will someday produce a total cure for arthritis.

The American Cancer Society (ACS) argued their guidelines for mammography every 1 to 2 years for women age 40 to 49 must be taken more seriously by providers and insurers. The organization takes issue with early statements by the Clinton administration health task force suggesting that screening begin at age 50. According to the ACS, the Clinton administration health task force proposals were based on "economic considerations rather than good science" (ACS, 1994). The ACS does acknowledge that there should be practice guidelines and emphasizes that these guidelines should be created by the ACS, not by any other group. The difficulty is that the ACS guidelines most clearly favor services offered by oncologists and other ACS-affiliated providers.

A few examples may clarify why the suggestions by these charitable organizations may be problematic. First, consider the Arthritis Foundation suggestion that we should reimburse orthopedic surgeons without challenge. Table 7.5 summarizes findings by Deyo and colleagues based on an observational trial of patients receiving spinal fusion for back pain. Spinal fusion is a complex surgical procedure that requires the fusion of vertebrae in the back. There has been a significant increase in the use of this procedure in recent years. The Deyo Study shows that, in comparison with patients not receiving spinal fusion, those undergoing the procedure are 4 times as likely to have a reoperation, twice as likely to die, 2.2 times more likely to end up in a nursing home, and 5.8 times as likely to have a transfusion. In comparison with controls, those undergoing surgery are twice as likely to have complications. Because this is an observational study rather than a randomized trial the ordinary precautions in interpreting the data are necessary. For example, there is no assurance that those who received surgery had the same risk factors for bad outcomes as those who did not get surgery. Nevertheless, these results significantly challenge the idea that orthopedic surgeons should be reimbursed without question (Deyo et al., 1994).

The ACS argument about mammography is a particularly interesting one. It is emotionally arousing, and it does highlight some important problems in public decision making. Two viewpoints must be considered. First, a variety of advocacy groups have become almost exclusively focused on mammography as a center point for women's health policy. They are infuriated by suggestions that there be any limitation whatsoever in the use of this cancer screening procedure. Another constituency includes providers who have made significant profit providing these tests. As we will see shortly, there is controversy over whether a screening mammography should be offered to women less than 49 years of age who do not have other risk factors for breast cancer. The medical establishment's position is best exemplified in statements by Paul Goldfarb, who is the past president of the ACS of California. According to the *San Diego Union*, Dr. Goldfarb has stated, "I don't know if mammograms are effective under the age 50, but I don't see any reason not to have them. Nobody is going to get hurt by them" (*San Diego Union*, April 24, 1994). I will argue that, in fact, women are hurt by these policies. This is not because mammograms are dangerous. Rather, the problem is one of opportunity cost. Devoting resources to mammography is harmful when it detracts from the opportunity to use the resources for other services that may be necessary to enhance the health of other women.

Within these last few years, there has been an extensive campaign designed to increase the use of screening mammography. Virtually all magazines targeted for female readerships have produced articles on the need for greater use of mammography. *McCalls Magazine* entitled their

TABLE 7.5 Ratios of Complications for Patients Undergoing Spinal Fusion in Comparision With Controls

Complication	Ratio
Blood transfusion	5.8
Nursing home placement	2.2
Reoperation	4.0
Mortality (6 weeks)	2.0

Based on cohort study of 27,111 Medicare recipients of whom 1518 had fusions. (Adapted from Deyo et al., 1994.)

article: "Breast Cancer Alert" and the cover of *Self* displayed "Saving Your Breasts." These magazines are marketed to a younger readership. In most cases, it is suggested that failure to provide mammography for all women is a political scandal. The article went on to say that breast cancer is the leading cause of death for younger women, and that breast cancer is the most common cancer diagnosis for women between the ages of 35 and 50. The article suggested that younger women should insist on mammograms and demand further examinations when the mammograms are negative because denser breast tissue (characteristic of young women) may obscure the visualization of a tumor.

The difficulty is that scientific evidence tends not to support the use of screening mammography for women younger than age 50. Various countries around the world have examined the evidence. Virtually all countries, except Sweden, have recommended that screening mammography begin at age 50. The United States is somewhat unique because we have recommended screening mammography begin at age 35. Last year the ACS and the National Cancer Institute (NCI) were split in their opinions. The NCI, after reviewing the evidence, suggested that screening begin at age 50, and the ACS still insists that screening begin at age 35 to 40 (Fletcher, 1993).

Part of the controversy is in the way outcomes are examined. If we consider a narrow outcome, such as the number of tumors detected, more mammography will find more cases. However, if we consider an outcome, such as deaths resulting from breast cancer, screening women younger than age 50 appears to have little or no benefit. The reasons for this are complex and poorly understood, and it may be valuable to review them briefly.

Figure 7.4 shows the rates of breast cancer detected between 1970 and 1990. Figure 7.4 shows that there was an increase in cases of breast cancer

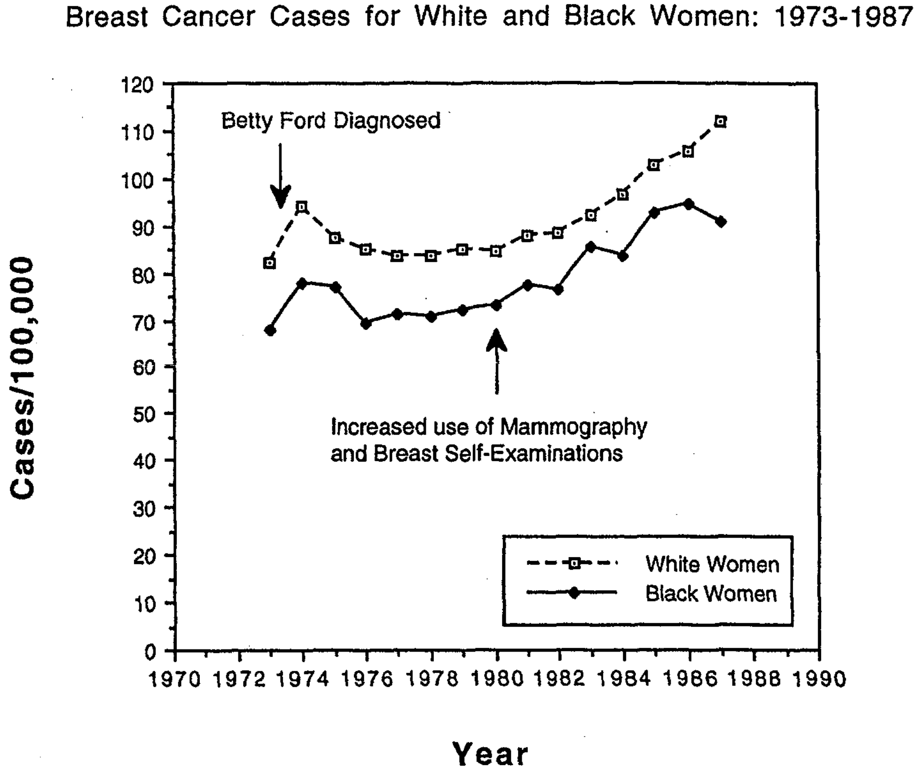

FIGURE 7.4 Breast cancer cases for White and Black women: 1973–1987.

for both White and Black women in the early 1970s. This coincides with the diagnosis of both Happy Rockefeller and Betty Ford. The diagnoses of breast cancer in these prominent women led to significantly more breast cancer screening. However, within a few years the rates of new cases declined. The 1980s were associated with a greater awareness of breast cancer and mass-scale mammography began in the mid-1980s. At that point, cases of breast cancer rose significantly.

If there is a public health benefit of early detection, then we should see a decrease in the rate of death due to breast cancer. Figure 7.5 summarizes the mortality experience during this same period. Interestingly, there has been no change at all in the rate of death because of breast cancer. In fact, the rate of death from breast cancer has been approximately the same over the last 50 years. Figure 7.5 marks the beginning of President Nixon's war on cancer, breakthroughs in surgical techniques, and the introduction of adjuvant chemotherapy for breast

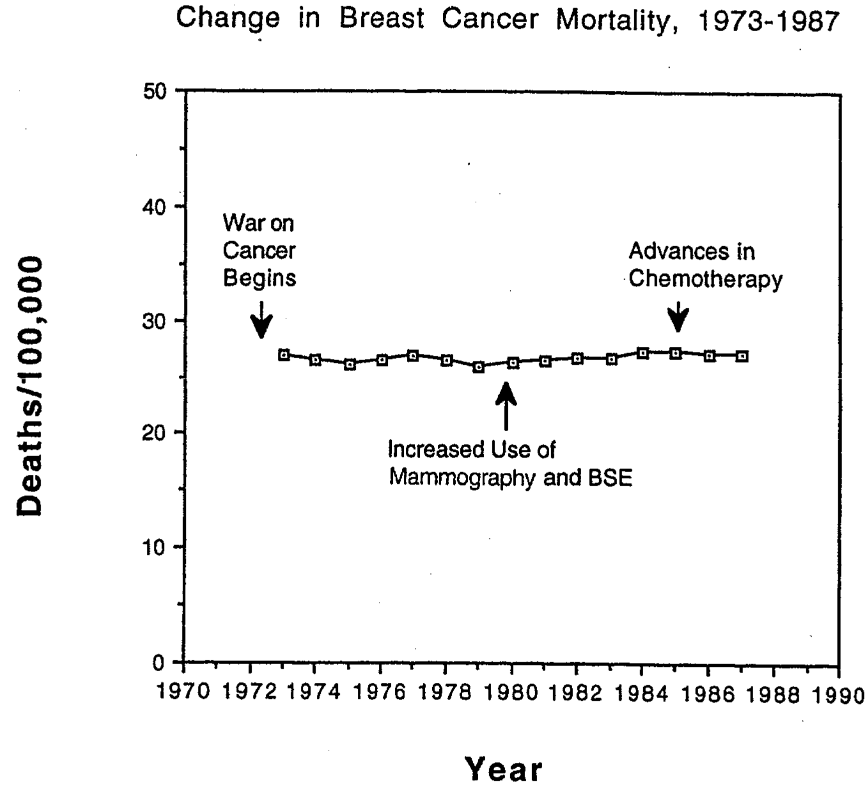

FIGURE 7.5 Change in breast cancer mortality, 1973–1987.

cancer. Each of these was announced as a major advance in breast cancer care. Nevertheless, the mortality rate associated with breast cancer has remained unchanged over this era.

The studies on breast cancer screening have also produced some confusion. There is no inconsistency across studies about the advantages of screening women past age 50. For post-menopausal women, the studies consistently show that breast cancer screening detects cases and results in significant reductions in mortality (Miller, 1994). The confusing aspect is for women 40 to 49 years of age. The most controversial study was the large-scale trial done in the Canadian National Breast Cancer Screening Study (Miller, 1993). In this randomized trial, women in the ages 40 to 49 years were more likely to die of breast cancer if they were regularly screened. A later reanalysis suggested that breast cancer screening may not have been harmful, but there was certainly no evidence that it was helpful.

We have performed a metaanalysis of these studies and shown that the risk ratio for women in the 40- to 49-year age bracket is higher than 1.0. In other words, averaged across studies, women screened regularly have a slightly (although not statistically significantly) *higher chance of dying* (Navarro & Kaplan, 1997, in press). Figure 7.6 summarizes an example of one of these studies. This is the Swedish Two-Countries Study, which compared the effect of screening for breast cancer by age. Risk ratios less than 1.0 imply a protective effect, whereas those greater than 1.0 suggest a damaging effect. Figure 7.6 shows that for women 50 to 59 and 60 to 69 years, there is a significant survival advantage of screening mammography. However, for women 40 to 49 years and those older than 70 years, there appears to be no advantage. In fact, eight out of eight studies in the literature fail to show any advantage of screening mammography for women 40 to 49 years of age. It is important to emphasize that these studies deal with asymptomatic women. Women with a family history of breast cancer or those who experience lumps are well advised to use mammography. Further, we do not want to suggest that these findings are conclusive. For example, it has been suggested that the failure to detect benefit for 40- to 49-year-old women might result from low power because of insufficient sample size or biases in older studies. However, most reviewers fail to find justification for population screening of women younger than the age of 50 (Kerikowski, Grady, Rubin, Sandrock, & Ernster, 1995).

The ACS is disturbed by the suggestions that women younger than age 50 do not need mammography. They argue that screening mammography is good for women. Clearly, it is also good for health care providers. But what about the argument that screening mammography will not hurt anybody? According to the opportunity cost argument, screening mammography may cause harm because it uses resources that could have been better spent elsewhere. Table 7.6 summarizes the cost per tumor detected for women of different ages. The most extreme case would involve screening women 20 years old. For 20-year-olds, the probability of having a breast tumor is about 1 in 100,000. If we assign mammography a cost of $100, approximately $10 million would be spent to detect a single case. Some may consider this expenditure would be worthwhile if it resulted in saving the life of that 20-year-old. However, the question of cost is mute. According to the available evidence, the woman would have no greater chance of survival than women who are unscreened. Considering the evidence, there is no reason to believe that screening before age 50 produces any significant public health benefit.

Analyses by Eddy (1989) also raise significant questions about the regular use of mammography. For example, a woman between ages 35

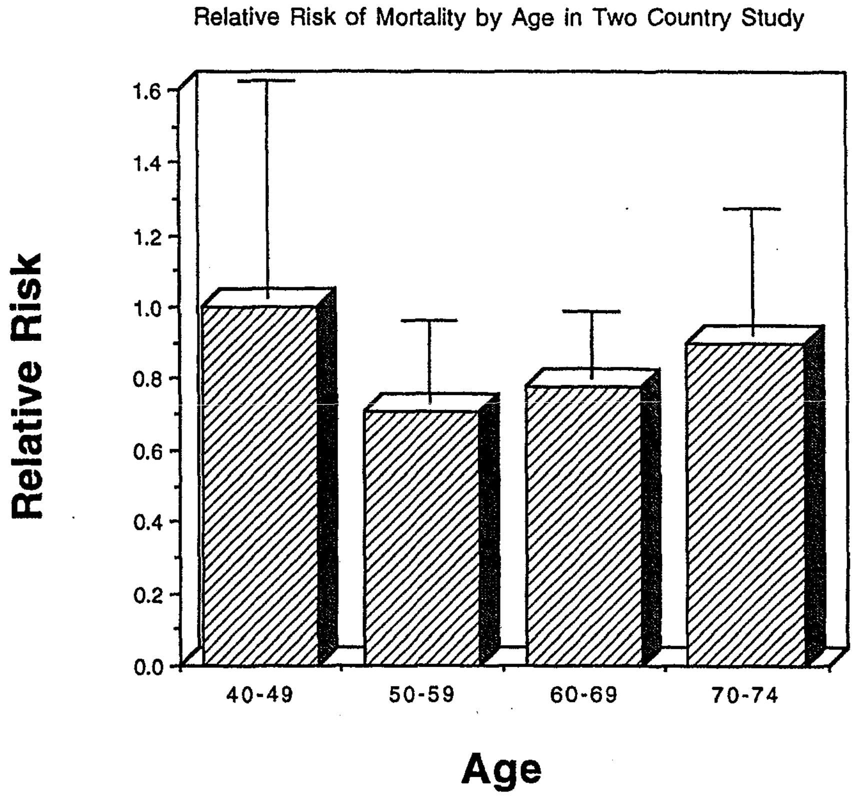

FIGURE 7.6 Relative risk of mortality by age in two-country study.

and 50 who obtains yearly screening mammography has either little or no probability of benefiting from the screening. However, in about one third of these women, findings will emerge that will require additional workup, including biopsy. These workups are not without consequence because they cause significant anxiety and can be costly.

Some of the implications of screening policies have recently been evaluated. Eddy (1994) used data from the Kaiser-Permanente Medical Group of Southern California. Currently, this HMO performs about 300,000 mammograms each year. About one half of these mammograms are completed on women between the ages 50 and 75 years, and about 45% are done for women younger than 50 years. The remaining 5% are done for women who are older than 75 years. Among the population of women that Kaiser serves, mammograms are given to about 22% of the women between the ages of 30 and 40 years, 60% of women between

TABLE 7.6 Cost per Tumor Detected by Age

Age (years)	Cost/Tumor
20	10,000,000
30	416,666
40	112,000
50	54,000
60	38,167
70	31,645

(Adapted from Kaplan, 1993a.)

40 and 50 years, and 69% of women between 50 to 75 years. In addition, Kaiser screens about 57% of the women between 75 and 85 years.

Using computer simulation, Eddy estimated that the current policy will prevent approximately 909 women from dying of breast cancer by the year 2010 at a cost of $707 million. There are alternative uses of the mammogram budget. One policy might be to discourage strongly the use of mammography for women younger than 50 years and older than 75 years. Instead, the policy might aggressively recruit women for mammography between ages 50 and 75 years, and those with risk factors for breast cancer (Eddy, 1994). In the 1990 National Health Interview Survey, less than 40% of women older than 50 years report screening mammography in the last year. An aggressive education program might significantly increase use of mammography in this group. Eddy (1994) estimated that if this program were successful in attracting 95% of the women in the 50- to 75-year age group, the number of breast cancer deaths prevented would increase to 1206 from 909 (a net increase in 297 lives). Further, the program would cost $210 million less than the current program. In other words, a cost-saving maneuver might result in about a 33% reduction in breast cancer deaths.

There could be biases in the studies. For example, let us assume that we are incorrect about the lack of benefit of mammography for younger women. Considering the most optimistic studies in the literature, screening mammography may reduce breast cancer mortality by 13% for women younger than 50 years of age (Nystrom et al., 1993). In Eddy's analysis, this would mean that the screening program would prevent 1045 premature deaths instead of 909. However, this is still significantly short of the 1206 deaths prevented by an age-targeted screening program. On the basis of a detailed review, of the literature, a group of policy analysts from the

RAND Corporation suggested that basic benefit packages in health insurance include screening mammography only for women 50 to 69 years of age (Kattlove, Liberate, Keeler, & Brook, 1995).

Opportunities Costs in the Mammography Example.

The real consequences of screening all women may accrue to the pool of women who cannot afford services. An estimated 17% of the U.S. population do not have health insurance. Today, public programs, such as Medicaid, cannot afford to support basic services for large numbers of people. In part, this results because public funds have been used to support some unnecessary services. Restricting the use of unnecessary services will free resources that could be used by others who are seriously in need of basic health care.

There are many potential alternative uses of the funds. For example, in a program like Kaiser of Southern California, restriction of mammograpy to women between the ages of 50 to 74 would save about $300 million each year (Eddy, 1994). What could be done with the savings? It is important to emphasize that many programs are not currently available within systems, such as Kaiser. For example, Eddy's analysis estimates that antismoking education programs for pregnant women may add 3,700 years of life that would have been lost to tobacco-related diseases. Other areas in which prevention programs could improve health status and prevent premature death include immunizations for children, prenatal care, and programs to reduce risk factors for cardiovascular diseases that kill more than 300 per 100,000 American women each year and remains the most common cause of death for both men and women in the United States (U.S. Department of Health and Human Services, 1991). These programs could be funded from the savings that would accrue from more effectively targeted mammography screening. The issue is not only to save money but also to use it more wisely (Navarro & Kaplan, 1997, in press).

CONCLUSION

Problems in health care are interconnected. Some of the problems might be characterized by the three words beginning with *A*. *Affordability* is linked to *access*, because creating greater access will use more resources. *Accountability* may help resolve some of these problems, because much of what makes health care unaffordable is the use of resources on services that provide little or no benefit. By changing reimbursement patterns to emphasize value for money, we may save enough to significantly expand access without raising cost.

Because of the growing expense of health care, there are significant consequences of not doing anything. Health care reform advocates must attend to all parts of the problem. The opportunity cost problem emphasizes that decisions necessarily involve choices between competing alternatives. Psychology is part of a larger health care network. A decision to reimburse psychologists for their efforts may displace the opportunity to spend resources on other services. If we allow orthopedic surgeons to be reimbursed for whatever they do, there may be little money left to provide mental health services. At some level, we are all part of the same system.

Competition for health care resources will require mental health providers to compete with other health care professionals. A general conceptualization of health outcome will allow direct comparisons between the productivity of mental health providers and the productivity of other health care providers.

REFERENCES

American Cancer Society. (1994). American Cancer Society Position: Clinton Health Reform Plan. New York: American Cancer Society.

Anderson, J. P., Kaplan, R. M., Berry, C. C., Bush, J. W., & Rumbaut, R. G. (1989). Interday reliability of function assessment for a health status measure: The Quality of Well-Being scale. *Medical Care, 27,* 1076–1083.

Arthritis Foundation. (1992). *Briefing paper for government affairs: Reforming access to health care.* Atlanta: Arthritis Foundation.

Bombardier, C., Ware, J., Russell, I. J., et al. (1986). Auranofin therapy and quality of life for patients with rheumatoid arthritis: Results of a multicenter trial. *American Journal of Medicine, 81,* 565–578.

Brook, R. H., & Lohr, K. (1986). Will we need to ration effective health care? *Issues in Science and Technology, 3,* 68–77.

Detsky, A. (1991). Guidelines for economic analysis of pharmaceutical products. Ontario Ministry of Health. Toronto: Drug Programme Branch, Ministry of Health.

Deyo, R. A., Ciol, M. A., Cherkin, D. C., Loeser, J. D., & Bigos, S. J. (1993). Lumbar spinal fusion: A cohort study of complications, reoperations, and resource use in the Medicare population. *Spine, 18,* 1463–1470.

Eddy, D. M. (1994). Clinical decision making: From theory to practice. Rationing resources while improving quality: How to get more for less. *Journal of the American Medical Association, 272,* 817–824.

Eddy, D. M. (1989). Screening for Breast Cancer, *Annals of Internal Medicine*, *111*, 389–399.

Erickson, P., Kendall, E. A., Anderson, J. P., & Kaplan, R. M. (1989). Using composite health status measures to assess the nation's health. *Medical Care*, *27* (Suppl. 3), S66–S76.

Fletcher, S. W., Black, W., Harris, R., Rimer, B. K., & Shapiro, S. (1993). Report of the International Workshop on Screening for Breast Cancer. *Journal of the National Cancer Institute*, *85*, 1644–1656.

Freund, D. A., Evans, D., Henry, D., & Dittus, R. (1992). Implications of the Australian guidelines for the United States. *Health Affairs*, *11*, 202–206.

Fuchs, V. R. (1990). The health sector's share of the gross national product. *Science*, *247*, 534–538.

Ganiats, T. G., Palinkas, L. A., & Kaplan, R. M. (1992). Comparison of Quality of Well-being scale and functional status index in patients with atrial fibrillation. *Medical Care*, *30*, 958–964.

Hadorn, D. C. (1991). Setting health care priorities in Oregon: Cost-effectiveness meets the rule of rescue. *Journal of the American Medical Association*, *265*, 2218–2225.

Health Care Financing Administration. (1994). *Office of the Actuary report*. Washington, DC: US Government Printing Office.

Hillman, B. J. (1994). New imaging technology and cost containment. *American Journal of Roentgenology*, *162*, 503–506.

Hillman, B. J., Joseph, C. A., Mabry, M. R., Sunshine, J. H., Kennedy, S. D., & Noether, M. (1990). Frequency and costs of diagnostic imaging in office practice: A comparison of self-referring and radiologist-referring physicians. *New England Journal of Medicine*, *323*, 1604–1608.

Hodgkin, P. S. (1994). Health impact of endoscopic sinus surgery assessed by the Quality of Well-Being (QWB) Scale. Unpublished paper, University of California, San Diego.

Hudson, T. (1990). Experts disagree over the cost of defensive medicine. *Hospitals*, *64*, 74.

Kaplan, R. M. (1993a). Application of a general health policy model in the American health care crisis. *Journal of the Royal Society of Medicine*, *86*, 277–281.

Kaplan, R. M. (1993b). *Hippocratic predicament: Affordability, access, and accountability in health care*. San Diego: Academic Press.

Kaplan, R. M. (1993c). Quality of Life Assessment for Cost/Utility Studies in Cancer. *Cancer Treatment Reviews*, *19* (Suppl. A), 85–96.

Kaplan, R. M., & Anderson, J. P. (1988a). A general health policy model: Update and applications. *Health Services Research*, *23*, 203–234.

Kaplan, R. M., & Anderson, J. P. (1988b). The Quality of Well-Being Scale: Rationale for a single quality of life index. In S. R. Walker & R. Rosser

(Eds.), *Quality of life: Assessment and application* (pp. 51–77). London: MTP Press.

Kaplan, R. M., & Anderson, J. P. (1990). An integrated approach to quality of life assessment: The general health policy model. In B. Spilker (Ed.), *Quality of life in clinical studies* (pp. 131–149). New York: Raven.

Kaplan, R. M., Anderson, J. P., Patterson, T. L., McCutchan, J. A., Weinrich, J. D., Heaton, R. H., Atkinson, J. H., Thal, L., Chandler, J., & Grant, I. (1995). Validity of the Quality of Well-Being Scale for Persons with human immunodeficiency virus infection. *Psychosomatic Medicine, 57*, 138–147.

Kaplan, R. M., Atkins, C. J., & Timms, R. (1984). Validity of a quality of well-being scale as an outcome measure in chronic obstructive pulmonary disease. *Journal of Chronic Diseases, 37*, 85–95.

Kaplan, R. M., Bush, J. W., & Berry, C. C. (1978). The reliability, stability, and generalizability of a health status index (pp. 704–709). *Proceedings of the American Statistical Association, Social Status Section.* Washington DC.Kaplan, R. M., Bush, J. W., & Berry, C. C. (1979). Health Status Index: Category rating versus magnitude estimation for measuring levels of well-being. *Medical Care, 17*, 501–525.

Kaplan, R. M., Bush, J. W., & Berry, C. C. (1976). Health status: Types of validity and the index of well-being. *Health Services Research, 11*, 478–507.

Kaplan, R. M., Hartwell, S. L., Wilson, D. K., & Wallace, J. P. (1987). Effects of diet and exercise interventions on control and quality of life in non-insulin-dependent diabetes mellitus. *Journal of General Internal Medicine, 2*, 220–228.

Kaplan, R. M., Kozin, F., & Anderson, J.P. (1988). Measuring quality of life in arthritis patients (including discussion of a general health-decision model). *Quality of Life and Cardiovascular Care, 4*, 131–139.

Kattlove, H., Alessandro, L., Keeler, E., & Brook, R. H. (1995). Benefits and costs of screening and treatment for early breast cancer. *Journal of the American Medical Association, 273*, 142–148.

Kerlikowske, K., Grady, D., Rubin, S. M., Sandrock, C., & Ernster, V. L. (1995). Efficacy of screening mammography: A meta-analysis. *Journal of the American Medical Association, 273*, 149–154.

Kerner, D., Patterson, T. L., & Kaplan, R. M. (1996). Validity of the Quality of Well-being Scale in Alzheimer's disease. Manuscript submitted for publication.

Kirkman-Liff, B. L. (1991). Health insurance values and implementation in The Netherlands and the Federal Republic of Germany: An alternative path to universal coverage. *Journal of the American Medical Association, 265*, 2496–2502.

Kitzhaber, J. (1990). The Oregon Basic Health Services Act. Salem: Oregon State Senate.

Levit, K. R., Cowan, C. A., Lazenby, H. C., McDonnell, P. A., Sensenig, A. L., Stiller, J. M., & Won, D. K. (1994). National health spending trends, 1960–1993. *Health Affairs, 13*, 14–31.

Lewin, J. C., & Sybinsky, P. A. (1993). Hawaii's employer mandate and its contribution to universal access. *Journal of the American Medical Association, 269*, 2538–2543.

Miller, A. B. (1993). Canadian National Breast Screening Study: Public health implications. *Canadian Journal of Public Health.* (Revue Canadienne de Sante Publique), *84*, 14–16.

Miller, A. B. (1994). Mammography screening guidelines for women 40 to 49 and over 65 years old. *Annals of Epidemiology, 4*, 96–101.

Navarro, A. M., & Kaplan, R. M. (1997, in press). Mammography screening: Prospects and Opportunity Costs. *Women and Health.*

Neumann, P. J., & Johannesson, M. (1994). From principle to public policy: Using cost-effectiveness analysis. *Health Affairs, 13(3)*, 206–221.

Nystrom, L., Rutqvist, L. E., Wall, S., Lindgren, A., Lindqvist, M., Ryden, S., Andersson, I., Bjurstam, N., Fagerberg, G., Frisell, J., et al. (1993). Breast cancer screening with mammography: Overview of Swedish randomised trials. *Lancet, 341*, 973–978.

Orenstein, D. M., Nixon, P. A., Ross, E. A., & Kaplan, R. M. (1989). The quality of of well-being in cystic fibrosis: 1. *95*, 344–347.

Patterson, T. L., Kaplan, R. M., Grant, I., Semple, S. J., Moscona, S., Koch, W. L., Harris, M. J., & Jeste, D. V. (1996, in press). Quality of Well-being in late-life psychosis. *Psychiatry Research.*

Reinhardt, U. E. (1993). Do global budgets make sense for the U.S. health system? *Internist, 34*, 9–11.

Squier, H., Kaplan, R. M., Ries, A. L., Prewitt, L. M., Smith, C. M., Kriett, J. M., & Jamieson, S. W. (1995). Quality of well-being predicts survival in lung transplantation candidates. *American Journal of Respiratory and Critical Care Medicine, 152*, 2032–2036.

U.S. Department of Health and Human Services, Public Health Service. (1991). Healthy people 2000: National health promotion and disease prevention objectives (DHHS publication No. [PHS] 91-50213). Washington, DC: US Government Printing Office.

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): Conceptual framework and items selection. *Medical Care, 30*, 473–483.

Wenneberg, J. E. (1994). Health care reform and professionalism. *Inquiry, 31*, 296–302.

Warshawsky, M. J. (1994). Projections of health care expenditures as a share of the GDP: Actuarial and macroeconomic approaches. *Health Services Research, 29*, 293–313.

Williams, A. (1988). The importance of quality of life in policy decisions. In S. Walker and R. Rosser (Eds.), *Quality of life: assessment and application* (pp. 279–290). London: MTP Press.