Health-Related Quality of Life in Cardiovascular Disease

Robert M. Kaplan
Center for Behavioral Medicine, San Diego State University, and
Department of Community and Family Medicine
University of California, San Diego

Early mortality and reduced quality of life in the years prior to death are the most important health outcomes associated with cardiovascular disease. Other measures of cardiovascular status, including blood pressure, ejection fraction, and electrocardiogram (ECG) abnormalities, are only of interest because of their known association with poor health outcomes. Quality of life measures have gained increasing attention as outcome variables in studies of cardiovascular disease. This article reviews several current approaches to the assessment of health outcomes. A general health policy model is offered as a method for comparing program options in cardiovascular disease that may have very different objectives. Examples taken from the evaluation of hypertension screening and treatment, of heart transplantation programs, and of primary prevention of heart disease are offered. Methods for measuring the cost/utility of alternative procedures are also discussed.

Cardiovascular disease has a major impact on contemporary American society. Each year, approximately 6 million people have symptomatic coronary heart disease. Among them, over 700,000 are hospitalized for myocardial infarction, and more than 1.5 million die. Heart disease is a major concern for at least two reasons. First, it affects life expectancy. Second, it deteriorates quality of life. For example, 20% of Americans will develop coronary heart disease by their 60th birthdays. Among these, 11% will die suddenly and 44% will suffer debilitating heart attacks (Hartunian, Smart, & Thompson, 1981). The direct and indirect costs of heart disease in the United States have been estimated to exceed $100 billion per year (Cooper & Rice, 1976; Weinstein & Stason, 1985).

Cardiovascular Disease and Health Outcomes

Modern treatments for coronary heart disease focus on increasing life expectancy and on improving quality of life. Some treatments, such as coronary artery bypass graft (CABG) surgery, have produced a somewhat disappointing effect on increased life expectancy. A controlled clinical trial known as the Coronary Artery Surgery Study (CASS) showed no survival advantage of surgery in comparison with medical management for patients who were asymptomatic following a myocardial infarction. There was a nonsignificant trend toward improved survival for a subset of patients with significant three-vessel heart disease and with compromised cardiac function (i.e., reduced ejection fraction) (CASS Principal Investigators and Their Associates, 1983). Despite some skepticism about whether new treatments increase life expectancy, they have shown encouraging effects on the ability to perform daily activities and the relief of pain. For example, a national cooperative study suggested that surgically treated patients experience fewer symptoms than those who are managed with medicine (R. O. Russell et al., 1978). The various methods for the treatment and prevention of heart disease are all designed to increase life expectancy and to improve quality of life. With an appropriate conceptualization of these outcomes, approaches ranging from surgery to medical management to prevention can be compared directly with one another (Weinstein & Stason, 1985). We will return to some of these comparisons in the application section of this article. However, before we consider applications, we must review the complex issues relevant to the conceptualization and measurement of quality of life.

Health-Related Quality of Life

It is instructive to consider health-care measures in light of the objectives of prolonging life and enhancing life quality. Traditional biomedical indicators and diagnoses are important to us because they may be related to mortality or to quality of life. For example, elevated blood pressure over a sustained period of time may predict death due to heart disease, heart-disease-related disability, or stroke. If high blood pressure were unrelated to these outcomes, it would be unimportant. In other words, blood pressure is important because it correlates with health outcomes, although it is not an outcome itself.

We use the term **health-related quality of life** to refer to the impact of health conditions on function. Health often relates to quality of life independently of work, housing, air pollution, and so forth (Rice, 1984). Within the last few years, a substantial number of quality of life measures have been proposed. In the next sections, some issues relevant to assessing life quality will be addressed.
What is Health-Related Quality of Life?

Many new quality of life measurement systems have evolved during the last 20 years. These systems represent various traditions in measurement. At least two conceptual approaches have been represented in recent articles. One approach grows out of the tradition of health status measurement. In the late 1960s and early 1970s, the National Center for Health Services research funded several major projects to develop general measures of health status. All of the projects were guided by the World Health Organization (WHO) definition of health status, which states that "health is a complete state of physical, mental, and social well-being and not merely absence of disease" (World Health Organization, 1948, preamble). Those projects resulted in a variety of assessment tools, including the Sickness Impact Profile (Bergner, Bobbitt, & Pollard, 1976; Bergner, Bobbitt, Carter, & Gilson, 1981), the Quality of Well-Being Scale (Kaplan & Bush, 1982), the Rand Health Status Measure (Stewart, Ware, Brook, & Davies, 1978), and the McMaster Health Index Questionnaire (Chambers, 1988). Many of the measures examined how disease or disability affect performance of social role, ability to get around the community, and physical functioning. Some of the measures included separate components for the assessment of social and mental health.

Other authors have referred to quality of life as something that is independent of health status. For instance, Follick et al. (1988) suggested that quality of life represents a broader outcome than traditional measures of symptoms and mortality. Croog et al. (1986) used a wide variety of outcome measures and collectively referred to them as quality of life. These measures included the patient's subjective evaluation of well-being, physical symptoms, sexual function, work performance and satisfaction, emotional status, cognitive function, social participation, and life satisfaction. Other investigators, including Hunt and McEwen (1983), have regarded quality of life as a subjective appraisal of life satisfaction. In summary, a wide variety of dimensions has been used to describe quality of life. There is no agreement on which dimensions should be assessed in research studies. Yet, consideration of recurrent themes in the methodologic literature can assist in the evaluation of existent instruments.

Unidimensional Versus Multidimensional

Most researchers agree that quality of life is a multidimensional construct. Yet, there is considerable debate about whether outcome measures must necessarily represent this multidimensional structure. There are two major approaches to quality of life assessment: a psychometric approach and a decision theory approach. The psychometric approach attempts to provide separate measures for the many dimensions of quality of life. Perhaps the best known example of the psychometric tradition is the Sickness Impact Profile (SIP). The SIP is a 136-item measure that yields 12 scores. The scores are displayed in a format similar to that used for a Minnesota Multiphasic Personality Inventory (MMPI) profile.

The decision theory approach attempts to weight the dimensions of health to provide a single unitary expression of health status. Supporters of this approach have argued that psychometric methods fail to consider that different health problems are not of equal concern. One-hundred runny noses are not the same as 100 severe abdominal bleeds (Bush, 1984). In an experimental trial using the psychometric approach, it is not uncommon to find that some aspects of quality of life improve, whereas others get worse. For example, a medication might reduce high blood pressure but might also be associated with headaches and impotence. It is argued that the notion of quality is the subjective evaluation of observable or objective health states. The decision theory approach attempts to provide an overall summary measure of quality of life that integrates subjective function states, preferences for these states, morbidity, and mortality.

Citing studies on factor analysis, Ware, Brook, Davies, and Lohr (1981) suggested that different components of health, including mental, physical, and social aspects, might be statistically independent dimensions. Thus, an aggregate measure of health status might be the same as adding apples to oranges.

In rebuttal, Bush (1984) argued that, indeed, different components of quality of life might be considered analogous to very different pieces of fruit. However, it is the overall evaluation of the basket of fruit that is important. A fruit peddler who regularly delivers a full basket of fruit is preferred over one who delivers a half-empty basket. A basket of fruit in which all pieces are fresh and none are rotten is preferred over one in which some pieces are either missing or decayed. Bush argued that the psychometric approach is analogous to comparing one bowl of fruit containing an apple and missing an orange with a second bowl of fruit containing a rotted banana and missing a pear. Ultimately, our concern is with the overall desirability of the aggregate.

Disease-Specific Versus General

Most health-related quality of life measures are designed for use with any population. However, some investigators feel it is necessary to develop quality of life measures for specific diseases, such as cardiovascular disease. For example, the RAND Corporation has produced a series of booklets describing the conceptualization and measurement of physiologic health. Each booklet describes the measurement of a specific condition, such as coronary heart disease. The rationale underlying these measures is largely clinical, and the approach assumes that each medical condition has very specific outcomes. Heart patients are evaluated according to ejection fractions, blood gases, and so forth.

In contrast to disease-specific approaches, many investigators believe that all diseases and disabilities have a general effect on quality of life. In fact, the purpose of quality of life measurement is not to identify clinical information relevant to the disease: Instead, quality of life measurement seeks to determine the impact of the disease on general function. It is argued that, by focusing too specifically on clinical correlates of disease, the general impact is overlooked. A well-designed, general quality of life measure adequately captures a wide variety of dysfunctions associated with almost any disease. These dysfunctions might occur in many different systems and might be recognized
in symptoms such as confusion, tiredness, sexual impotence, and depression.

Common Methods for the Measurement of Quality of Life

A variety of methods has been proposed for the measurement of quality of life. It is not possible in the limited space provided here to review and critique all of these methods. Instead, some of the more widely used methods will be presented. Readers interested in more detailed reviews should consult Bergner (1985); Kaplan (1985a); and Wenger, Mattson, Furberg, and Elinson (1984).

Four of the more commonly used assessment methods include the Sickness Impact Profile (SIP), the Rand Health Insurance Measures, the Index of Activities of Daily Living Scales, and ad hoc measurement systems.

Sickness Impact Profile (SIP)

The SIP (Bergner et al., 1981) is one of the best-known and most widely used quality of life measures. It is a general measure that is applicable to any disease or disability group. Furthermore, the SIP has been successfully used with a variety of cultural subgroups.

The SIP includes 136 items that describe the effect of sickness on behavioral function. These items are divided into 12 categories. In turn, the 12 categories are further clustered into three groups: independent, physical, and psychosocial. The independent categories include sleep and rest, eating, work, home management, and recreation and pastimes. Physical categories include ambulation, mobility, and body care and movement. The psychosocial categories include social interaction, alertness behavior, emotional behavior, and communication.

Each SIP item has been evaluated by an independent group of judges on a 15-point scale of dysfunction. The judges' ratings determine the weighting of each item in the SIP scoring. The respondent does not consider the judges' weightings in deciding to endorse an item. The overall SIP percent score is obtained by separating the items endorsed by the respondent, summing their scale values, and dividing by the sum of all values for all items on the SIP. Then, this proportion is multiplied by 100 to compute a total percent-impaired score. Scores are obtained similarly for each category. Percent scores for each category can be plotted on a graphic display that looks similar to an MMPI profile. A variety of studies has attested to the reliability and validity of the SIP for various patient groups and clinical applications (see Bergner et al., 1981). An example of an application of the SIP in cardiovascular disease is summarized in Figure 1. The figure compares 50 patients with angina with 50 healthy volunteers. The angina patients had more impairment in all categories. However, work and recreation were most seriously affected. Angina had less impact on body care, mobility, and eating (Vandenberg, 1988).

Clearly, the SIP is one of the most important and best-developed quality of life measures. Two minor issues are relevant to its general use. First, the SIP is a measure of morbidity or illness but does not easily integrate with mortality or death outcomes.

Some medical studies simultaneously consider changes in death and changes in illness. The SIP may be less suited than other measures for these studies. The second problem is that the SIP is sometimes cumbersome to administer. With 136 items, it can be time consuming, and it requires alertness and attention by the respondent. However, the SIP is an example of a measurement system that has undergone systematic methodological refinements over many years. It has been widely used and well-evaluated.

Index of Activities of Daily Living

Perhaps the oldest general quality of life measure is the Index of Activities of Daily Living (ADL; Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963). Katz et al. were very early to argue that the major effects of disease and disability were exerted on function and on the ability to perform role activities in everyday life.

The ADL is most commonly used in studies of the elderly. The system includes six subscales: bathing, dressing, toileting, transfer, continence, and feeding. For each category, assessors judge whether the person is independent or dependent. For the category of bathing, people are judged to be independent if they need assistance only in bathing a single part of the body or can bathe themselves. They are judged to be dependent if they need assistance in bathing more than one part of the body. Once a judgment of dependence or independence is obtained for each of the six categories, an overall grade is assigned. To receive the top grade, or Grade A, the person must be independent in all six categories. Grade B is assigned for those who are independent in all but one of these functions. The bottom grade, G, is assigned to those who are dependent in all six functions. Several reliability and validity studies for the ADL have been reported (Katz, Downs, Cash, & Grotz, 1970). Despite its many important ap-

![Figure 1. Sickness Impact Profile impairment scores for angina and healthy patients.](image-url)
In studies of aging, the ADL has been criticized because it does not make distinctions towards the well end of the quality of life continuum. Stewart et al. (1978) suggested that nearly 80% of the noninstitutionalized population have no gross functional limitations and would obtain the top score in the ADL system. However, other population surveys have demonstrated that more than 50% of the population experience one or more symptoms on a given day (Kaplan, Bush, & Berry, 1976). Most of these cause minor irritation but do not cause the type of dysfunction detected by ADL measures. For example, chest pain, shortness of breath, and dizziness would not be recorded by the ADL scales if they did not disrupt major role activities. Given the wide array of behaviors and symptoms that are relevant to quality of life, the ADL has some significant limitations. Other measures are required to make discriminations among impaired individuals who score toward the healthy end of the functioning continuum.

**RAND**

Perhaps the most thorough review of the conceptualization of health status and quality of life measures has been conducted by the RAND Corporation. The RAND physical health status measures are adaptations of measures developed by Bush, Kaplan, and others to describe physical activity, social activity, and mobility. The social activity category was subdivided to include social activity, role activity, household activity, and leisure activity. The RAND group also adapted Dupey’s (1969) General Well-Being Index. In addition, they have included the General Health Perceptions Questionnaire (Ware & Karmos, 1976) as well as self-report questionnaires to assess the clinical status associated with a wide variety of medical conditions (Brook et al., 1979).

The RAND approach has the advantage of being very comprehensive. For example, the approach may demonstrate that patients with cardiovascular disease have shown minor improvements in certain aspects of mobility and role performance yet are experiencing the side-effects of mental confusion and headache. Unlike some other systems, however, the RAND approach does not allow a single summary statement about whether the patients are getting better or worse on a composite index.

**Ad Hoc Approaches**

Another approach is illustrated in a recent article published in the *New England Journal of Medicine*. In this multicenter clinical trial that compared medications to control mild-to-moderate high blood pressure, a variety of quality of life measures was used. The investigators used a wide range of measures in an attempt to characterize the benefits and side effects associated with using the medications. The measures included a sense of well-being scale and measures of physical symptoms, sexual function, work performance and satisfaction, emotional status, cognitive function, social participation, and life satisfaction. Many of these measures had multiple components. For example, cognitive function was evaluated using both the Wechsler Memory Scale and the Trailmaking Test from the Halstead-Reitan Neuropsychological Test Battery (Reitan, 1986). Physical symptoms were measured with the Physical Symptom Distress Index and the Sleep Dysfunction Scale (Reitan, 1986). This approach has the advantage of casting a wide net. Thus, it is possible to identify the advantages and disadvantages associated with specific treatments. The disadvantage is that there is no relative weighting system. If a medication reduces blood pressure while it increases headaches or sexual dysfunction, some judgment of quality is still required to determine whether the medication should be continued (Croog et al., 1986).

Another common problem in studies that use measures with multiple subscores is the subject-to-variable ratio. In smaller clinical trials, the number of subjects is sometimes approached by the number of variables. Under these circumstances, multivariate comparisons can be highly unreliable. For univariate comparisons, the multiple endpoints inflate the probability of a Type I error. Measures such as the SIP have a defined number of scales. Yet, some of the newer approaches appear to have many potential endpoints, which allows the investigator to select variables for discussion post hoc.

**Decision Theory Approaches**

Within the last few years, there has been growing interest in using quality of life data to help evaluate the cost/utility or cost effectiveness of health-care programs. Cost studies have gained in popularity because health-care costs have grown rapidly. Not all health-care interventions are equally efficient in returning benefits for the expended dollar. Objective cost studies might guide policy makers toward an optimal and equitable distribution of scarce resources. Cost/effectiveness analysis typically quantifies the benefits of a health-care intervention in years of life or quality-adjusted life years (QALYs). Cost/utility is a special use of cost/effectiveness that weights observable health states by preferences or utility judgments of quality (Kaplan & Bush, 1982). In cost/utility analysis, the benefits of medical care, behavioral interventions, or preventive programs are expressed in "well" years. These outcomes have also been described as quality-adjusted life years (Weinstein & Stason, 1976), discounted life years (Kaplan et al., 1976) or healthy years of life (L. B. Russell, 1986). Because the term quality-adjusted life years has become most popular, we will use it in this presentation. The QALYs integrate mortality and morbidity to express health status in equivalents of well years of life.

If a man dies of heart disease at the age of 50 years and we would have expected him to live to the age of 75 years, it might be concluded that the disease was associated with 25 lost life years. If 100 men died at the age of 50 years (and also had a life expectancy of 75 years) we might conclude that 2,500 (100 men × 25 years) life years had been lost. Yet, death is not the only outcome of concern in heart disease. Many adults suffer myocardial infarctions that leave them somewhat disabled over long periods of time. Although they are still alive, the quality of their lives has diminished. Quality-adjusted life years take into consideration the quality of life consequences of these illnesses. For example, a disease that reduces quality of life by one half will take away .5 QALYs over the course of each year. If it affects
To place the observable states on a preference continuum, with an anchor of 0 for death and 1.0 for completely well. In several studies, random samples of citizens from a metropolitan community evaluated the desirability of over 400 case descriptions. Using these ratings, we developed a preference structure that assigns weights to each combination of an observable state and a symptom/problem (Kaplan et al., 1976). Cross-validation studies have shown that the model can be used to assign weights to other states of functioning with a high degree of accuracy (R² = .96). The regression weights obtained in these studies are given in Tables 1 and 2. Studies have shown that the weights are highly stable over a 1-year period and that they are consistent across diverse groups of raters (Kaplan, Bush, & Berry, 1978).

Finally, it is necessary to consider the duration of stay in various health states. For example, 1 year in a state that has been assigned the weight of .5 is equivalent to .5 of a QALY.

The calculating formula for the point-in-time well-being score (W) for an individual is

\[ W = 1 + (CPXwt) + (MOBwt) + (PACwt) + (SACwt) \]

where wt is the preference-weighted measure for each factor, CPX is the symptom/problem complex, MOB is the mobility level, PAC is the physical activity level, and SAC is the social activity level.

Consider, for example, the person described by the following condition for a particular day: a CPX level of 11 (cough, wheezing, or shortness of breath with or without fever, chills, or aching all over) with a weight of −.257, a MOB level of 5 (no limitations) with a weight of .000, a PAC level of 1 (in bed, chair, or couch for most or all of day, health-related) with a weight of −.077, and a SAC level of 2 (performed no major role activity, health-related, but did perform self-care) with a weight of −.061.

The W value for this particular day is 1 + (−.257) + (.000) + (−.077) + (−.061) = .605.

This suggests that, for a particular day, a person in this condition is in an observable state and is judged to be about 60% of the distance between death and optimal function. If restricted to this state for an entire year, the person would have lost the equivalent of about .4 well years.

In addition to point-in-time well-being, it is also crucial to consider the duration of stay in well states. The person described in this example could be suffering from the flu or from disabling heart disease. Yet, the flu might last only a few days, whereas heart disease is chronic. Thus, in order to accurately conceptualize health status, we need to consider the duration of time in the state or the probability of movement between well states over the course of time. Conceptually, we express this as well years (WY) through the following formula:

\[ WY = \text{[Number of Persons} \times (CPXwt + MOBwt + PACwt + SACwt)] \times \text{Time}. \]
## Table 1
### Quality of Well-Being General Health Policy Model

<table>
<thead>
<tr>
<th>Step no.</th>
<th>Step definition</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>No limitations for health reasons</td>
<td>-.000</td>
</tr>
<tr>
<td>4</td>
<td>Did not drive a car, health related; did not ride in a car as usual for age (younger than 15 years), health related; and/or did not use public transportation, health related; or had or would have used more help than usual for age to use public transportation, health related</td>
<td>-.062</td>
</tr>
<tr>
<td>2</td>
<td>In hospital, health related</td>
<td>-.090</td>
</tr>
</tbody>
</table>

### Mobility Scale

<table>
<thead>
<tr>
<th>Step no.</th>
<th>Step definition</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>No limitations for health reasons</td>
<td>-.000</td>
</tr>
<tr>
<td>4</td>
<td>Did not drive a car, health related; did not ride in a car as usual for age (younger than 15 years), health related; and/or did not use public transportation, health related; or had or would have used more help than usual for age to use public transportation, health related</td>
<td>-.062</td>
</tr>
<tr>
<td>2</td>
<td>In hospital, health related</td>
<td>-.090</td>
</tr>
</tbody>
</table>

### Physical Activity Scale

<table>
<thead>
<tr>
<th>Step no.</th>
<th>Step definition</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>No limitations for health reasons</td>
<td>-.000</td>
</tr>
<tr>
<td>3</td>
<td>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; and/or limped, used a cane, crutches, or walker, health related; and/or had any other physical limitation in walking or did not try to walk as far as others the same age are able, health related</td>
<td>-.060</td>
</tr>
<tr>
<td>1</td>
<td>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</td>
<td>-.077</td>
</tr>
</tbody>
</table>

### Social Activity Scale

<table>
<thead>
<tr>
<th>Step no.</th>
<th>Step definition</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>No limitations for health reasons</td>
<td>-.000</td>
</tr>
<tr>
<td>4</td>
<td>Limited in other (e.g., recreational) role activity, health related</td>
<td>-.061</td>
</tr>
<tr>
<td>3</td>
<td>Limited in major (primary) role activity, health related</td>
<td>-.061</td>
</tr>
<tr>
<td>2</td>
<td>Performed no major role activity, health related, but did perform self-care activities</td>
<td>-.061</td>
</tr>
<tr>
<td>1</td>
<td>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</td>
<td>-.106</td>
</tr>
</tbody>
</table>

## Table 2
### Quality of Well-Being General Health Policy Model Symptom/Problem Complexes (CPX) With Calculating Weights

<table>
<thead>
<tr>
<th>CPX no.</th>
<th>Description</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Death (not on respondent’s card)</td>
<td>-.727</td>
</tr>
<tr>
<td>2</td>
<td>Loss of consciousness such as seizure (fits), fainting, or coma (out cold or knocked out)</td>
<td>-.407</td>
</tr>
<tr>
<td>3</td>
<td>Burn over large areas of face, body, arms, or legs</td>
<td>-.387</td>
</tr>
<tr>
<td>4</td>
<td>Pain, bleeding, itching, or discharge (drainage) from sexual organs—does not include normal menstrual (monthly) bleeding</td>
<td>-.349</td>
</tr>
<tr>
<td>5</td>
<td>Trouble learning, remembering, or thinking clearly</td>
<td>-.340</td>
</tr>
<tr>
<td>6</td>
<td>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</td>
<td>-.333</td>
</tr>
<tr>
<td>7</td>
<td>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</td>
<td>-.299</td>
</tr>
<tr>
<td>8</td>
<td>Pain, burning, bleeding, itching, or other difficulty with rectum, bowel movements, or urination (passing water)</td>
<td>-.292</td>
</tr>
<tr>
<td>9</td>
<td>Sick or upset stomach, vomiting or loose bowel movement, with or without fever, chills, or aching all over</td>
<td>-.290</td>
</tr>
<tr>
<td>10</td>
<td>General tiredness, weakness, or weight loss</td>
<td>-.259</td>
</tr>
<tr>
<td>11</td>
<td>Cough, wheezing, or shortness of breath with or without fever, chills, or aching all over</td>
<td>-.257</td>
</tr>
<tr>
<td>12</td>
<td>Spells of feeling upset, being depressed, or of crying</td>
<td>-.257</td>
</tr>
<tr>
<td>13</td>
<td>Headache, or dizziness, or ringing in ears, or spells of feeling hot, or nervous, or shaky</td>
<td>-.244</td>
</tr>
<tr>
<td>14</td>
<td>Burning or itching rash on large areas of face, body, arms, or legs</td>
<td>-.240</td>
</tr>
<tr>
<td>15</td>
<td>Trouble talking, such as lisp, stuttering, hoarseness, or being unable to speak</td>
<td>-.237</td>
</tr>
<tr>
<td>16</td>
<td>Pain or discomfort in one or both eyes (such as burning or itching) or any trouble seeing after correction</td>
<td>-.230</td>
</tr>
<tr>
<td>17</td>
<td>Overweight for age and height or skin defect of face, body, arms, or legs, such as scars, pimpls, warts, bruises, or changes in color</td>
<td>-.188</td>
</tr>
<tr>
<td>18</td>
<td>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</td>
<td>-.170</td>
</tr>
<tr>
<td>19</td>
<td>Taking medication or staying on a prescribed diet for health reasons</td>
<td>-.144</td>
</tr>
<tr>
<td>20</td>
<td>Eyeglasses or contact lenses</td>
<td>-.101</td>
</tr>
<tr>
<td>21</td>
<td>Breathing smog or unpleasant air</td>
<td>-.101</td>
</tr>
<tr>
<td>22</td>
<td>No symptoms or problem (not on respondent’s card)</td>
<td>-.000</td>
</tr>
<tr>
<td>23</td>
<td>Standard symptom/problem</td>
<td>-.257</td>
</tr>
</tbody>
</table>

Note: CPX = symptom/problem complex.
states of function. When the proper steps have been followed, the model quantifies the health activity or treatment program in the QALY's that it produces or saves. A QALY is defined conceptually as the equivalent of a completely well year of life, or as a year of life free of any symptoms, problems, or health-related disabilities.

More detailed descriptions of this system are available in other publications (Kaplan, 1985a, 1985b; Kaplan & Bush, 1982). In the following sections, we will illustrate applications of the QALY concept.

Applications

There is a wide variety of quality of life measurement applications for studies in cardiovascular disease. First, quality of life measures are being used more frequently as outcome measures in clinical trials. A clinical trial is an experimental evaluation of a medical or behavioral treatment conducted in a representative clinical setting. Second, major epidemiologic studies are beginning to use quality of life measures as covariates and as health outcomes. Third, clinicians are becoming increasingly concerned about the quality of life implications of their interventions strategies. Indeed, there is now an entire journal published by LaJoaq entitled Quality of Life and Cardiovascular Care. For most purposes, quality of life measures are used as dependent variables in research and clinical settings. However, quality of life measures are also becoming increasingly popular in policy analysis. In this section, a variety of applications of the QALY approach will be presented. These applications include discussions of adherence to antihypertensive medications, exercise, and heart transplantation.

Adherence to Antihypertensive Medications

Hypertension is a major public health problem because of its high prevalence and its association with heart disease and stroke. Many people are unaware that they have hypertension, and many of those who are aware are unwilling to take the necessary actions to control the condition.

Weinstein and Stason (1976) estimated the cost/utility for programs that screen severe hypertension (diastolic blood pressure > 105 mm/Hg) at $4,850 per well year, whereas the corresponding figure for mild hypertension screening programs (diastolic blood pressure = 95–104 mm/Hg) was $9,800 per year in 1976 dollars. The costs and effects were considered according to the prevalence of high blood pressure in the community, the costs of finding and treating new cases, and the expected benefits of treatment.

Weinstein and Stason's (1976) analysis also considered a variety of factors that influence these cost/utility ratios. One of the most important factors is adherence to the prescribed medical regimen after cases have been detected. The figures we discuss assume full adherence to the regimen. Yet, substantial evidence has revealed that full or 100% adherence is rare (see DiMatteo & DiNicola, 1982). Compliance with antihypertensive medications is of particular interest because taking the medication does not relieve symptoms. In fact, medication adherence can increase rather than decrease somatic complaints. Some studies have suggested that behavioral intervention can be very useful in increasing adherence to prescribed regimens (Haynes et al., 1976).

In their analysis, Weinstein and Stason (1976) considered the value of programs designed to increase adherence to antihypertensive medication. For their analyses, they decided that failure to adhere to treatment had occurred when patients took only 50% of the prescribed medicine and obtained only 50% of the benefit. Two cases of incomplete adherence were considered. On one extreme is the patient who fails to see a physician and purchases more medication than the 50% that was used. Here the cost is low. On the other extreme is the patient who remains under medical care and purchases medications but uses only 50% of them. In this case, the costs are high. Weinstein and Stason (1976) referred to these as the minimum cost assumption and the maximum cost assumption. Under the minimum cost assumption, patients do not receive the full benefits of medication because of incomplete adherence. Yet, they also do not spend their money to fill all prescriptions. According to Weinstein and Stason (1976), the cost effectiveness under this assumption is very similar to the full adherence in which patients receive the benefits of medication but make full expenditures. Under the maximum cost assumption, the effect of incomplete adherence is substantial, particularly for those who begin therapy beyond the age of 50. Earlier, it was noted that the cost to produce a well year using antihypertensive medications was $4,850 for those in a U.S. national sample who had pretreatment diastolic blood pressure greater than 105 mm/Hg. With incomplete adherence, these values would increase to $6,400 under the minimum cost assumption and to $10,500 under the maximum cost assumption. In screening for mild hypertension (diastolic blood pressure = 95–104 mm/Hg), the $9,880 per well year under the full adherence assumption would rise to $12,500 under the minimum cost assumption and to $20,400 under the maximum cost assumption.

Because adherence under the maximum cost assumption appears to have a strong effect on cost/utility, it is interesting to consider the value of behavioral interventions to improve adherence. Several studies have shown that behavioral interventions are effective (Shapiro & Goldstein, 1982), and it is reasonable to assume that a successful behavioral intervention will improve adherence rates by 50% (Haynes et al., 1976). Weinstein and Stason (1976) considered the cost/utility of interventions that would improve adherence by 50% under the maximum cost assumption. Their analysis of the hypothetical programs that would reduce diastolic blood pressure from 110 mm/Hg to 90 mm/Hg suggested a differential expected cost/utility for programs designed for men and for women. The results are summarized in Figure 2. The figure suggests that blood pressure screening and treatment programs are more cost effective for men if they are implemented early in the life cycle. Conversely, for women, the programs are most cost effective if they begin after the age of 40 years because women rarely die of heart disease before menopause. The most important aspect of the figure is that the adherence programs produce QALYS at a cost comparable to other programs. For example, an adherence program for men with severe diastolic hypertension would produce a QALY for less than $20,000 in any age group.
In the Weinstein and Stason (1976) monograph, a variety of other hypothetical conditions were considered. Under the assumption that the program would improve adherence by 50%, a significant benefit of the program resulted under the maximum cost assumption. However, under the minimum cost assumption, the hypothetical adherence intervention would have produced a significant benefit if it increased adherence by 50% but no significant effect if it increased adherence by only 20%.

Heart Transplantation

Heart transplantation is becoming a relatively common procedure in both the United Kingdom and the United States. Despite the notion that the procedure is extremely specialized, most major cities now have a heart transplantation program. Yet, because of the enormous costs associated with heart transplantation, there has been growing interest in the value of the procedure. Both the United States and the United Kingdom have conducted studies emphasizing the quantification of health benefits of cardiac transplantation. These studies have considered both survival and quality of life.

An example of a study quantifying the health benefits of a transplantation was reported by O'Brien, Buxton, and Ferguson (1987). These investigators performed a nonrandomized trial with 221 patients undergoing heart transplantation in the United Kingdom. The data suggested that there was a significant increase in survival for transplant patients in comparison with eras when the procedure was not available. In addition, the introduction of the immunosuppressive drug cyclosporine A in 1982 produced a significant increment in survival after heart transplantation. The United Kingdom group used a psychometrically based profile approach to quality of life assessment. The measure, known as the Nottingham Health Profile (Hunt & McEwen, 1983), yields six dimensions of social functioning: pain, energy, physical mobility, sleep, social isolation, and emotional reactions. However, the Nottingham Health Profile does not provide an overall measure of quality of life appropriate for a policy analysis. Therefore, the investigators modified the measure to provide a single aggregate score.

The United Kingdom study was hindered by a variety of design problems. For example, techniques of heart transplantation improved over the course of the study. Subjects who entered into the study at different times received different operations. In addition, there may have been a tendency for surgeons to select the very sickest patients for the procedure, and it was impossible to obtain a true control group. However, the best evidence is that the patients who are selected for the procedure have an extremely poor prognosis and without the intervention gradually get worse until they die. For the most recent transplantation cases, evidence suggested that patients improved in all six of the Nottingham Quality of Life categories. Using an approximation of the QALY system, O'Brien et al. (1987) reported that the transplantation patients improved from a very poor health status prior to surgery to an estimated .9 on a 0–1.0 scale (1.0 = optimum health). In addition, they estimated that the average patient gained 8.3 life years as a function of the heart transplantation. The 8.3 years of added life, weighted at .9, translated into approximately 7.5 years adjusted for quality of life.

The O'Brien et al. (1987) study is problematic because the quality of life measurement system was not designed for the ultimate analyses the investigators preferred. Thus, as they clearly acknowledged in their own article, the choice of a more appropriate measurement system may have facilitated the conclusions that were drawn from the study.

Exercise

It is widely believed that regular exercise may prevent coronary heart disease. However, the epidemiologic literature does
not clearly confirm or disconfirm these beliefs. Studies of San Francisco longshoremen (Paffenbarger & Hale, 1975), of British civil servants (Morris, Everitt, Pollard, & Chave, 1980), and of Harvard alumni (Paffenbarger, Hyde, Wing, & Jung, 1978) have all demonstrated that men who engage in vigorous physical activity over the course of years experience lower rates of heart disease mortality and morbidity than do men who are habitually less active. However, none of these studies involved the random assignment of men to various activity levels. Thus, it is not clear whether the results reflected selection or protection. In other words, it could be that healthier men prefer to be more physically active rather than that physical activity protects health. The study of Harvard alumni has been particularly influential because recent data have demonstrated that exercise is associated with lower death rates from all causes—not only heart disease. This implies that the protective effects relevant to heart disease are not offset by damaging effects that are reflected in other diseases (Paffenbarger, Hyde, & Wing, 1986).

Whether there are significant side effects associated with exercise has not clearly been established. A study of men and women in Seattle suggested that heart attacks occur at a higher rate during exercise than during periods of inactivity. However, the same study demonstrated that the protective effects of exercise outweigh the risks of having a heart attack during an exercise session (Sisovich, Weiss, Fletcher, & Lasky, 1984). Van Camp and Peterson (1986) reported that serious complications in supervised exercise programs for heart disease patients are rare. They observed only one fatality per 750,000 patient hours of supervised exercise and about nine cardiac arrests per million patient hours of exercise. These outcomes are not greatly different than those expected without exercise for patients in this age group.

On the other hand, various studies have suggested that regular exercise is associated with muscle strains, cramps, torn ligaments, and blisters. Among over 2,000 entrants who began the London Marathon, 26 participants were sent to the hospital and 4 remained under care overnight (Nicholl & Williams, 1983). A follow-up study of participants in an Atlanta recreational run revealed that more than a third of the regular runners suffered some temporary setback due to injury within a year. Fifteen percent consulted a doctor about their injuries and a remarkable 7% had collisions with bicycles or cars or were bitten by dogs (Koplan, Powell, Sikes, Shirley, & Campbell, 1982). In summary, neither the risks nor the benefits of exercise have been clearly demonstrated.

In her analysis of exercise, L. B. Russell (1986) questioned whether the costs of exercise programs have been thoroughly considered. For example, adults who begin an exercise program accrue a variety of expenses, including medical expenses. The American Heart Association recommends that physicians perform thorough and costly evaluations of adults considering exercise programs (American Heart Association, 1972). In England, where the existence of socialized medicine may influence such policies, the Royal College of Physicians suggested that most people do not need a medical exam before starting an exercise program. Instead, they recommend that people begin their programs slowly and seek medical advice if they have unexpected symptoms (Royal College of Physicians, 1983).

Worksite health promotion programs have also gained considerable attention. However, L. B. Russell (1986) suggested that the costs of these programs are often not well-estimated. A company may need to pay for a training facility, exercise supervisors, locker rooms, and so forth. The exercisers may need to spend money on running shoes, extra soap and water for additional showers, additional medical evaluations, medical costs for care of injuries, and other expenses. Given the data currently available, L. B. Russell (1986) was unable to estimate either the costs or the benefits of an exercise program.

Conclusion

Using quality of life measures and a general health policy model, we can make comparisons between very different alternatives in health care. The cost/utilities of the behavioral programs to improve adherence with blood pressure medication are more favorable for severe than for mild hypertension. In addition, the programs are more cost effective if they begin early in life for men but later in life for women. When they are compared in the same units, the behavioral programs compare quite favorably with surgery. Using the same system, Weinstein and Stason (1982) found that coronary artery bypass surgery for a three-vessel disease cost about $30,000 to produce a QALY. The same surgery for a one-vessel disease cost nearly $600,000 to produce a QALY. Williams (1988) argued that action to lower total serum cholesterol might produce a QALY for as little as $2,600. Thus, preventive actions may be considerably more cost effective than medical and surgical remedies for established problems. Over the next few years, the number of interventions analyzed using this system is expected to grow.

In this article a variety of approaches to quality of life assessment has been discussed. Quality of life is clearly a multidimensional construct. However, there is still considerable debate about whether or not multidimensional measures are required. Some approaches attempt to represent the multidimensionality by providing quality of life profiles. Other approaches attempt to scale the multidimensional construct with a unidimensional preference weighting system. The choice of a unidimensional versus a multidimensional approach depends on the purpose of the study. Multidimensional approaches provide more clinical diagnostic information about areas in which there are deficits. The unidimensional approach is better suited for policy analysis and comparisons of very different alternatives in health care. A related issue is the specificity of the measure for the disease being considered. Some investigators prefer measures specific to cardiovascular disease, whereas others argue that comprehensive but general measures are more desirable. Future research will be required to identify the incremental validity of using disease-specific measures.

Although there is still considerable debate about the definition of quality of life, policy analysts have been active in the use of the QALY concept. In some areas, particularly those relevant to prevention, there is a serious lack of research data needed for the computation of QALY indices. Often, the quality of life impact of an intervention is estimated rather than measured. More studies are required to identify the cost/utility of many
preventive measures, including exercise, diet, and coping with stress.

References


Received July 20, 1987
Revision received September 15, 1987
Accepted September 15, 1987