3 The Quality of Well-Being Scale: rationale for a single quality of life index

Robert M. Kaplan and John P. Anderson

This chapter provides an overview of the Quality of Well-Being (QWB) Scale and the General Health Policy Model from which it was derived. Our group argues that a single, comprehensive expression of quality of life has many desirable features for policy analysis, evaluation research and clinical investigation. A recurrent theme in several of our previous publications is that a single index of health status is both feasible and highly desirable.[1,2,15,16,24,34–43]

Health status measurement has been characterized by competing traditions; one of the major issues is disease specificity. Some investigators argue that specific measures are required for each disease category, while others, including our group, believe there are many advantages to a general approach. Among those favouring the general approach to health status measurement, some groups have focused on mortality while others have focused on morbidity. Our approach is to integrate morbidity and mortality into common units of health status. In this chapter we will elaborate each of these issues.

General versus disease-specific health measures

Most health-related quality of life measures are designed for use with any population. However, some investigators feel that it is necessary to develop quality of life measures for specific diseases, for example, the RAND Corporation has produced a series of booklets describing the conceptualization and measurement of 'physiologic health' where each booklet describes the problems in conceptualization and measurement of a specific condition, such as anaemia, acne and vision impairment. The rationale underlying the development of these measures is largely clinical. It suggests that medical conditions have very specific outcomes. Thus, for example, diabetic patients are evaluated according to blood glucose, while chronic obstructive lung

disease patients are evaluated according to pulmonary function. Clearly there are advantages to the clinician in considering outcomes relative to specific diseases.

In addition to general physiological indicators, there are also quality of life measures designed specifically for particular disease groups. For instance, Meenan and colleagues have developed a specific quality of life measure for arthritis patients[50] and this is only one among many approaches to health status assessment for this particular disease.[47]

Disease-specific measures can produce follow-up information that will necessarily fall within the error of more general measures. For instance, it may be useful to know that a new treatment for burns generally allows hand burn victims to regain sufficient finger dexterity to open a bag of potato chips, while current treatments do not provide these benefits. Any general measure that attempted to gather this wealth of detail would soon overwhelm respondents with what might be considered ridiculously trivial questions.

Though useful for many purposes, these disease-specific measures have a weakness from the policy point of view: their use precludes the possibility of comparing programmes that are directed at different populations or groups suffering from different diseases. In addition, many health care interventions affect outcomes that are not system-specific. For example, cigarette smoking may increase the probability of coronary heart disease, peripheral artery occlusion, cerebrovascular disease, and cancers of the larynx, lung, mouth, oesophagus, bladder, pancreas and stomach; it is also the major cause of chronic obstructive pulmonary disease.[25] From this point of view, the impact of smoking upon health is truly overwhelming, yet general health status measures can provide a comprehensive summary of these heterogeneous health effects. Policy analysis also requires a more general approach to health status assessment.

Mortality

Mortality remains the major outcome measure in most epidemiological studies and clinical trials. Typically, mortality is expressed in units of time. In order for mortality data to be significant, they must be expressed in the form of a rate, that is the proportion of deaths from a particular cause occurring in some defined time interval (usually per year), and usually, mortality rates are age-adjusted. Case fatality rates express the proportion of persons who died of a particular disease divided by the total number with the disease (including those who live).

There are many advantages to reporting mortality rates. They are hard data (despite some misclassification bias),[57] and the significance of the outcome is not difficult to comprehend. However, there are also some obvious limitations. Mortality rates consider only the dead and ignore the

living, so many important health care services, including prevention, can be expected to have little or no impact upon them. For example, each year there are approximately 1.2 million cataract removal procedures performed in the United States.[32] Although the procedure is essentially non-controversial, cataract removal has little or no impact on mortality and is certainly unrelated to infant mortality. An outcome measure that focused only on mortality would miss the value of this surgery, which proves to have benefits in as many as 95% of the cases.

Morbidity

The most common approach to health status assessment is to measure morbidity in terms of function or role performance, and morbidity estimates often include work days missed or bed disability days. Most approaches to health status assessment are essentially morbidity indicators. Thus, the Sickness Impact Profile (SIP)[7] represents the effect of disease or disability upon a variety of categories of behavioural function, while the RAND Health Insurance measures include separate categories for the effects of disease or health states upon physical function, social function and mental function. These measures do not integrate morbidity and mortality, though as each birth cohort ages, there is accrual of mortality cases. Death is a health outcome and it is important that this outcome not be excluded from any expression of health status.

To illustrate this, consider evaluating the effect of an integrated support and treatment programme as opposed to no support or treatment, for randomly assigned groups of very ill, elderly, nursing-home residents. Let us suppose that the programme maintained them all at a very low level of function throughout the year, while in the comparison group, the sickest 10% died. Looking just at the living in the follow-up, one finds the comparison group to be healthier, since the sickest have been removed by mortality. By this standard, the programme of no support as treatment might be put forth as the better alternative. With a measure that combined morbidity and mortality, however, the story would be very different, with mortality effects dragging overall health of the comparison group to a very low level. The importance of mortality will be discussed more fully later in this chapter.

Well-years

Our approach is to express the benefits of medical care, behavioural intervention, or preventive programmes in terms of well-years. Others have chosen to describe the same outcome as Quality Adjusted Life Years (QALYs).[73] Well-years integrate mortality and morbidity to express health

status in terms of equivalents of well-years of life. If a cigarette smoker died of heart disease at age 50 and we would have expected him to live to age 75, it might be concluded that the disease caused him to lose 25 life-years. If 100 cigarette smokers died at age 50 (and also had life expectancies of 75 years), we might conclude that 2500 (100 men by 25 years) life-years had been lost.

Death is not the only outcome of concern in heart disease. Many adults suffer myocardial infarctions leaving them somewhat disabled over a longer period of time. Although they are still alive, the quality of their lives has diminished. Our model permits various degrees of disability to be compared to one another. A disease that reduces the quality of life by one half will take away 0.5 well-years over the course of one year. If it affects two people, it will take away 1.0 well-year (equal to 2×0.5) over a one-year period. A medical treatment that improves the quality of life by 0.2 for each of 5 individuals will result in a production of one well-year if the benefit is maintained over a one-year period. Using this system, it is possible to express the benefits of various programmes by showing how many equivalents of well-years they produce. However, not all programmes have equivalent costs. In periods of scarce resources, it is necessary to find the most efficient use of limited funds. The well-year approach provides a framework within which to make policy decisions that require selection between competing alternatives. Preventive services may in this way compete with traditional medical services for the scarce health care dollars and moreover can be competitive in such analyses. Performing such comparisons requires the use of a general health decision model. In the next section, the general model of health status assessment and benefit–cost/utility (BCU) analysis will be presented.[2]

THE GENERAL MODEL

Building a Health Decision Model

The Health Decision Model grew out of substantive theory in economics, psychology, medicine and public health. These theoretical linkages have been presented in several previous papers.[2,15,16,24] Building a health decision model requires at least five distinct steps.

Step 1: Defining a function status classification

During the early phases of the Health Index Project, a set of mutually exclusive and collectively exhaustive levels of functioning were defined. After an extensive, specialty-by-specialty review of medical reference works, all of the ways that disease and injuries can affect behaviour and role performance were listed. Without considering aetiology, it was possible to match a finite number of conditions to items appearing on standard health

surveys, such as the Health Interview Survey (National Center for Health Statistics), the Survey of the Disabled (Social Security Administration), and several rehabilitation scales and ongoing community surveys. These items fit conceptually into three scales representing related but distinct aspects of daily functioning: mobility, physical activity, and social activity. The mobility and physical activity scales have three levels, while social activity has five distinct levels (see Appendix, p. 305). Several investigators have used this functional status classification (or a modified version of it) as an outcome measure for health programme evaluation.[70] However, the development of a truly comprehensive health status indicator requires several more steps.

Step 2: Classifying symptoms and problems

There are many reasons a person may not be functioning at the optimum level. Subjective complaints are an important component of a general health measure because they relate dysfunction to a specific problem. Thus, in addition to function level classifications, an exhaustive list of symptoms and problems has been generated. Included in the list are 21 complexes of symptoms and problems representing all the possible symptomatic complaints that might inhibit function. These symptoms and problems are shown in the Appendix, p. 306.

Step 3: Preference weights to integrate the QWB Scale.

We now have described the three scales of function and 21 symptom/problem complexes. With these, all we can do is compare populations in terms of frequencies for each scale step (and, if necessary, symptom/problem complex) as shown for the mobility scale in Table 3.1. Although comparison of frequencies is common in health services research, our system offers a strategy for integrating the frequencies into a single comprehensive expression. Comparing frequency distributions for different groups can be difficult and confusing. If our intent is to say which of these distributions is 'better' off and which 'worse', simple frequency distributions may not be able to help much. Is Group 1 better or worse than Group 5, and by how much? Also, this is only one scale; how can one make such decisions when there are three scales and the symptom/problem complexes to consider?

Table 3.1 Example of mobility scale frequency distributions for QWB Scales

Scale/Step	*Group 1*	*Group 2*	*Group 3*	*Group 4*	*Group 5*
MOB					
5	80	75	75	78	85
4	15	20	22	22	5
2	5	5	3	0	10

Another step is necessary to integrate the three scales and the symptom/problem complexes in a manner that will allow a single numerical expression to represent each combination of steps on the scales and symptom/problem complexes. The empirical means of accomplishing this is measured preferences for the health states. As noted earlier, the Health Decision Model includes the impact of health conditions upon the quality of life. This requires that the desirability of health situations be evaluated on a continuum from death to completely well. An evaluation such as this is a matter of utility or preference, and thus, functional level-symptom/problem combinations are scaled to represent precise degrees of relative importance.

Human judgement studies are needed to determine weights for the different states. We have asked random samples of citizens from the community to evaluate the relative desirability of a number of health conditions. Random sample surveys were conducted in the San Diego community during two consecutive years. The probability sample included 867 respondents ethnically representative of the population. When necessary, interviews were conducted in Spanish. From a listing of all possible combinations of the scale (mobility, physical activity, and social activity), we drew a stratified random sample of 343 case descriptions (items) and divided them into eight sets of computer-generated booklets. All respondents were assigned randomly to one of the eight booklets, creating eight sub-groups of approximately 100 respondents each. In a series of studies, a mathematical model was developed to describe the consumer decision process. The validity of the model has been cross validated with an R^2 of 0.94 (Ref. 40). These weights, then, describe the relative desirability of all the function states on a scale from zero (for death) to 1.0 (for asymptomatic optimum function). Thus, a state with a weight of 0.50 is viewed by the members of the community as being about half as desirable as optimum function or about halfway between optimum function and death.

Using these weights, one component of the general model of health is defined. This is the QWB scale, which is the point in time component of the Health Status Index.[24,39] The QWB score for any individual can be obtained from values associated with his/her function level, adjusted for symptom or problem.

The example in the Appendix, p. 307, shows a state for which a weight of 0.605 has been obtained. Using the symtom/problem adjustment, the Index becomes very sensitive to minor top-end variations in health status. For example, there are symptom/problem complexes for wearing eyeglasses, having a runny nose, or breathing polluted air. These symptom adjustments apply even if a person is in the top step in the other three scales. For example, a person with a runny nose receives a score of 0.83 on the QWB scale when he is at the highest function level.[39] Thus, the Index can make fine as well as gross distinctions.

Mathematically, the QWB score may be expressed as

$$W = \frac{1}{N} \sum_{\kappa=1}^{l} W_\kappa N_\kappa$$

where

W is the symptom standardized time-specific QWB score,
κ indexes the function levels [$\kappa = 1, \ldots l$]
W_κ is the QWB (weight, utility, relative desirability, social preference) for each function level, standardized (adjusted) for all possible symptom/problem complexes,
N_κ is the number of persons in each function level, and
N is the total number of persons in the group, cohort, or population.

Thus, QWB is simply an average of the relative desirability scores assigned to a group of persons for a particular day or a defined interval of time.

Several of our studies attest to the reliability[40] and validity[39] of the QWB scale. For example, convergent evidence for validity is given by significant positive correlations with self-rated health and negative correlations with age, number of chronic illnesses, symptoms and physician visits. However, none of these other indicators was able to make the fine distinction between health states which characterize the QWB scale. These data support the convergent and discriminant validity of the scale.[39]

Step 4: The well-life expectancy

QWB is only one of the two major components of the Health Decision Model. The other component requires consideration of transitions among the levels over time. Consider the health situation described in the Appendix, p. 000. Suppose that this condition described two different individuals – one who was in this condition because of participation in a marathon race and another because of arthritis. The fact that these individuals are in these conditions for different reasons is reflected by different expected transitions to other levels over the course of time. The marathon runner probably is sore from her ordeal, but is expected to be off and running again within a few days. However, the arthritis sufferer will probably continue to convalesce at a low level of function. A Health Decision Model must consider both current functioning and probability of transition to other function levels over the course of time. When transition is considered and documented in empirical studies, the consideration of a particular diagnosis is no longer needed. We fear diseases because they affect our current functioning, symptoms, or pain either now or at some time in the future. A person at high risk for heart disease may be functioning very well at present, but may have a high probability of transition to a lower level (or death) in the future. Cancer would not be a concern if the disease did not affect current functioning or the probability that functioning would be affected at some future time.

When weights have been properly determined, health status can be expressed precisely as the expected value (product) of the preferences associated with the states of function at a point in time and the probabilities of transition to other states over the remainder of the life-expectancy. Quality of well-being (W) is a static or time-specific measure of function, while the well-life expectancy (E) also includes the dynamic or prognostic dimension. The well-life expectancy is the product of QWB times the expected duration of stay in each function level over a standard life period, and can be expressed as

$$E = \sum W_\kappa Y_\kappa$$

where

E is the symptom-standardized well-life expectancy in equivalents of completely well-years, and

Y is the expected duration of stay in each function level or case type estimated with an appropriate statistical (preferably stochastic) model.

An example computation of the well-life expectancy is shown in Table 3.2. Suppose that a group of individuals was in a well state for 65.2 years, in a state of non-bed disability for 4.5 years and in a state of bed disability for 1.9 years before their deaths at an average age or 71.6 calendar years. In order to make adjustments for the diminished quality of life they suffered in the disability states, the duration of stay in each state is multiplied by the preference associated with the state. Thus, the 4.5 years of non-bed disability become equivalents of 2.7 well-years when we adjust for the preferences associated with inhabiting that state. Overall, the well-life expectancy for this group is 68.5 years. In other words, disability has reduced the quality of their lives by an estimated 3.1 years.

Table 3.2 Illustrative computation of the weighted life expectancy

State	Y_κ	W_κ	W_κ	Y_κ
Well	A	65.2	1.00	65.2
Non-bed disability	B	4.5	0.59	2.7
Bed disability	C	1.9	0.34	0.6
Total		71.6		68.5

(Reproduced from Ref. 38, with permission.)

$$\text{Weighted life expectancy} = \sum_{\kappa=1}^{l} W_\kappa Y_\kappa = 68.5 \text{ well-years}$$

$$\text{Current life expectancy} = \sum_{\kappa=1}^{l} Y_\kappa = 71.6 \text{ calendar-years}$$

Step 5: Estimating the benefit–cost/utility ratio

The San Diego Health Index Group has shown in a variety of publications how the concept of a well or weighted life expectancy can be used to evaluate the effectiveness of programmes and health interventions. The output of a programme has been described in a variety of publications as Quality Adjusted Life Years,[11] Well-years, Equivalents of Well-years, or Discounted Well-years.[15] Weinstein[72,73] has popularized the concept and calls the same output Quality Adjusted Life Years (QALYs), and this has been adopted by the Congressional Office of Technology Assessment.[58] It is worth noting that the Quality Adjusted Life Years terminology was originally introduced by Bush, Patrick and Chen,[59] but later abandoned because it had surplus meaning. The term wellness or well-years implies a more direct linkage to health conditions. Whatever the term, the Index shows the output of a programme in years of life adjusted by the quality of life which has been lost because of disease or disability.

By comparing experimental and control groups on a general health status measure, it is possible to estimate the output of a programme in terms of the well-years it produces. This is shown as the area between curves representing the two groups in Figure 3.1. Dividing the cost of the programme by the well-years it yields, gives the BCU ratio.

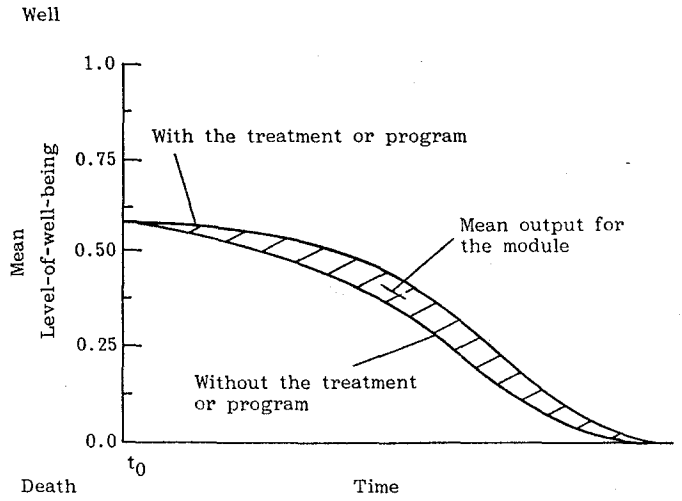


Figure 3.1 Theoretical comparison of treated and untreated groups. The area between the two curves is the output or benefit of a programme in well-year units. (Reproduced from Ref. 38 with permission).

There are many attractive elements of general health status measures. We argue that the ultimate purpose of health care and prevention is directed toward two simple objectives: investments in health care are aimed at

extending the duration of life, and health care programmes should improve the quality of life while individuals are alive. A comprehensive expression of health status can determine the effects of a programme using a unit that simultaneously considers risks and benefits. As Mosteller[54] has suggested, specific measures of health outcome often ignore the side-effects of treatment. A treatment for hypertension, for example, may cause gastric irritation, nausea and bed disability. Health benefits of treatment can be expressed in well-year units, as can health side-effects (costs).

CLINICAL EXAMPLES

This section provides several examples describing how this system might be used to evaluate pharmaceutical products or medical technology. Examples describing the use of the system for the evaluation of health policy alternatives are available.[2,11]

It is important to emphasize that the QWB scale is the measurement system for a General Health Policy Model. Ultimately, we hope that clinical trials will incorporate these measures so the estimates of treatment effects can be obtained in well-year units. Many of the analyses presented in this section depend upon estimates of QWB scores rather than the actual measurements, but are presented to emphasize the potential for utilizing quality of life measures for policy studies.

The tight control of insulin-dependent diabetes mellitus

Several studies have suggested that degree of hyperglycaemia is associated with the long-term risk of diabetic complications.[65] However, there is no strong experimental evidence confirming that reduction in blood sugar leads to a parallel reduction in diabetic complications. The most frequently cited study purporting to show the benefits of the tight control of diabetes[33] has been aggressively criticized because there were many therapeutic crossovers, data were incomplete, and the difference in blood glucose between experimental and control participants was not large.[61] Other studies have failed to show reversals of microvascular diabetic complications with intensive therapy.[6]

The question of tight control of diabetes was considered ambiguous enough for the US National Institutes of Health to begin a prospective clinical trial to evaluate the benefits of tight control versus ordinary care. The trial, known as the Diabetes Control and Complications Trial (DCCT), will include approximately 1400 subjects treated over a 10-year period. A portion of the DCCT subjects will be evaluated using the General Health Policy Model, which may well have substantial advantages for estimating treatment benefits. In addition to mortality, diabetes may be associated with

poor outcomes in a variety of organ systems. For example, poor control might lead to differential rates of retinopathy, kidney failure and foot infections. The difficulty is in finding one common expression for these outcomes, when some patients may have foot infections that result in amputations while others have eye problems that result in blindness. One purpose of our system is to aggregate these outcomes with death to provide a single expression of the impact of poor control. Diabetic coma receives a score of approximately 0.287 on our scale, while vision impairment that interferes with driving a car and work, but does not interfere with self-care might receive a score of 0.610. This tells us that two days of diabetic coma add up to less than one day of vision impairment, but a treatment that eliminates diabetic coma (averaged across the duration of the coma) might be considered more valuable than one that reduces vision impairment. The objective is to eliminate any sort of impairment; however, our system does provide for some weighting of the very different outcome measures used in the study.

The system also includes the capability of integrating side-effects and benefits of treatment in the same unit. For example, suppose that the treatment reduces the probability of retinopathy by 25%. We will assume that 40% of the patients will eventually get serious retinopathy.[44] Suppose further that the retinopathy begins at age 55 and continues until death at age 75, and the weight associated with blindness or serious vision impairment might be 0.5. Our sytem might suggest that the equivalent of 0.4 chances of developing serious retinopathy multiplied by 0.5 as the average decrease in well-being for 20 years, times 0.25 reduction in severity resulting from treatment would equal 1.0 well-year. In other words, the improved treatment of diabetes might add the equivalent of one healthy year to the life expectancy.

Now, we must consider the consequences or side-effects of tight control. For the sake of argument, assume that the intensive treatment begins at age 30. One-third of the patients experience nausea and weakness associated with tight control on half of the days. The duration is $75 - 30 = 45$ years, divided by the number of days in which there are symptoms, $0.5 \times 45 = 22.5$ years, multiplied by the weight associated with the symptom of sick or upset stomach which is 0.75. The net side-effects are 0.33 of all patients $\times$ 22.5 years $\times$ 0.25 average decrease in QWB (i.e. 1.0–0.75), which totals 1.87 years. In this example the side-effects might cause a loss of the equivalent of 1.87 years while the benefits produce a benefit of 1.0 year, although the benefits for other aspects of treatment must also be considered; so, for example, we would also consider the altered probability of kidney disease, heart disease, etc. With these added the benefits would most likely outweigh the side-effects.

Ultimately, the net effects of a treatment are expressed in QALY units. The next question concerns determination of the costs to produce a quality adjusted or well-year unit, from which comparison of health care programmes with very different specific objectives may be made.

Auranofin treatment for patients with rheumatoid arthritis

Clinical trials for treatments of rheumatoid arthritis have considered a wide variety of endpoints. The traditional approach has been to review clinical outcomes such as degree of synovitis. This is typically assessed through tender or swollen joints, grip strength, time to walk 50 feet, or duration of stiffness upon rising in the morning. At an international conference on outcome measurement in arthritis, it was suggested that comprehensive assessments of quality of life outcomes were highly desirable.[8] In a recent clinical trial involving 14 centres, more than 300 patients were randomly assigned to therapy with oral gold (auranofin) or a placebo. A wide variety of traditional and non-traditional measures were used to assess outcome, and among the non-traditional measures was the QWB scale. The outcome using the QWB is shown graphically in Figure 3.2. There was essentially no change in QWB function for the placebo group while the group receiving auranofin showed a mean improvement of 0.023 and this difference was statistically significant beyond the 0.005 level. Auranofin does not reach pharmacologically effective levels for about two months, and QWB scores for the treatment and placebo groups begin to diverge at about this time. Considering the many measures used in the trial, the percentage of variance accounted for with the QWB measure was among the most significant (see Figure 3.2). Outcomes measured using traditional, clinical measures, such as the 50-foot walk and duration of morning stiffness, were not statistically significant, although they did favour the auranofin group. In addition, simple self-ratings by both patients and physicians failed to detect the significant effect. However, a significant network of associations emerged suggesting that the QWB was associated with other similar measures of general function.[9]

The clinical significance of a difference of 0.023 must be considered, and although this appears to be small, the QWB provides a direct translation into clinically relevant units. A difference of 0.023 translates into 2.3 QALYs for each 100 patients who maintain the difference for 1 year. Also, although 0.023 appears to be a small change, the entire continuum from death to optimal health is represented on the 0 to 1.0 scale. The differences observed in the auranofin trial are quite respectable in comparison to those obtained through other medical treatments.

One of the most important aspects of the QWB is its capability of quantifying side-effects as well as benefits. Many of the specific scales used for the auranofin trial were not capable of detecting the general effect of the intervention upon health status, yet, gold preparations are known to cause significant adverse effects including diarrhoea, headache, rashes, digestive problems and abdominal pain. In fact, 59% of the auranofin treatment patients experienced diarrhoea at some point in comparison to 19% of tl placebo treated group. The General Quality of Life Assessment System

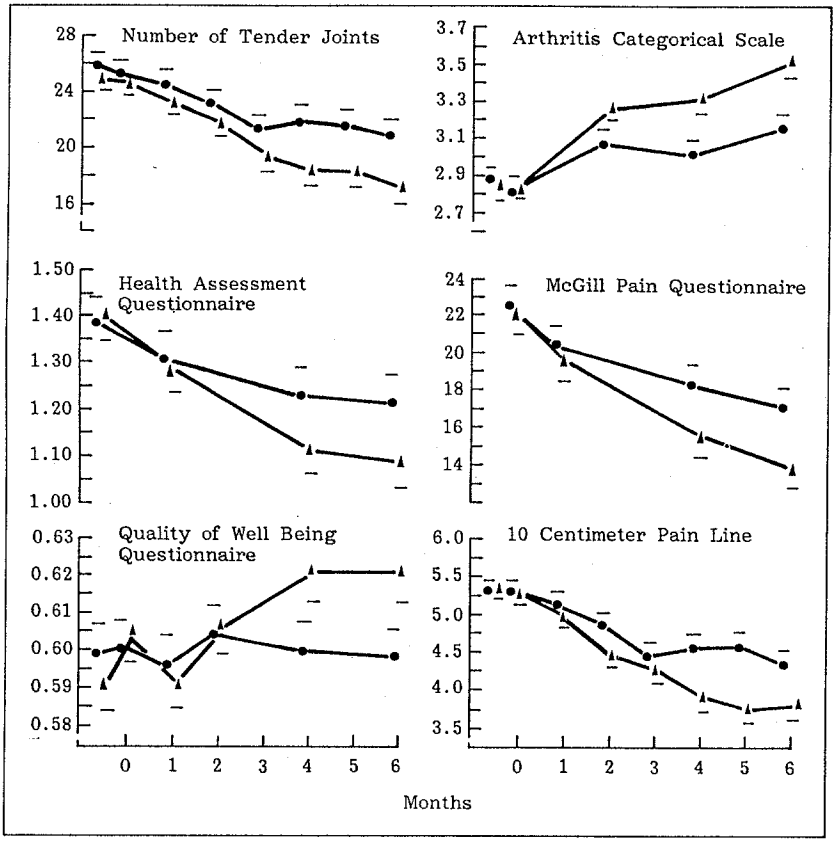


Figure 3.2 Outcome of patients treated with auranofin (▲——▲) or placebo (●——●) assessed by traditional and non-traditional measures. (Reproduced from Ref. 9 with permission)

allowed these side-effects to be integrated with benefits in order to provide a comprehensive expression of net treatment efficacy. In the near future, the auranofin cooperation group will release data on the BCU of auranofin therapy.

Coronary artery bypass grafts or CABG

Despite some controversy,[10,14,31] coronary artery bypass graft (CABG) has become a major treatment for symptomatic coronary artery disease. The number of procedures performed in the United States has steadily grown to an estimated 110,000 procedures in 1980 at an estimated cost of $15,220 per operation.[45] A 1981 forum sponsored by the National Center for Health

Care Technology concluded that the cost per surgery ranged from $11,000 to $25,000 and that the annual cost to the nation was near $2 billion.[56] The significance of the procedure and the expenses associated with it, led Weinstein and Stason[72] to conduct a systematic evaluation of the literature on CABG using a cost-utility model with the data provided by clinical reports, systematic longitudinal data banks, and clinical trials, including the major trials conducted by the European Coronary Surgery Group and the Veterans Administration (VA) Cooperative Study.

The analysis considered the benefit for a 55-year-old male population, since 55 years is approximately the median age for receipt of CABG. The analysis considered only those men who would be deemed operable by cardiologists on the basis of clinical characteristics and angiography, and was done separately for men with obstruction (defined as 50% or more) of 1, 2 or 3 coronary arteries or left main coronary artery disease. In each of the cases, ventricular function was good, with at least a 40% ejection fraction. The analysis for patients with poor ventricular function will not be considered here.

In order to calculate QALYs, Weinstein and Stason needed to integrate morbidity and mortality information. They used data about symptomatic relief from the European study[22,23] and from the Montreal Heart Institute,[13] and also simulated the benefit results, using a variety of quality judgements for observed levels of functioning and symptomatic angina.

The approach used by Weinstein and Stason uses data from different sources. Data from the VA study and the European trial differed in their evaluation of the benefits of surgery for one-vessel and two-vessel disease: the VA data suggest that surgery may be detrimental in these cases, while the European data indicate there will be benefits. Results from these two trials, and other data, were merged to obtain central assumptions that are operative in the analysis, although the analysis can also consider differing assumptions, and the impact these assumptions have upon quality adjusted life expectancy. Under the assumption that the preference for life with angina is 0.7 (on a scale from 0–1), Weinstein and Stason estimated the benefits of surgical treatment over medical treatment for the various conditions. They found that the benefits in quality adjusted life years would be 0.5, 1.1, 3.2 and 6.2 years for one, two and three vessels, and left main artery disease respectively.

Next, Weinstein and Stason estimated the cost of the surgery and evaluated cost–utility under the central assumptions. Assuming that surgery relieves severe angina, the estimates ranged from $30,000 per year for one vessel disease to $3,800 per year for left main artery disease. Weinstein and Stason performed these analyses under a variety of assumptions, and in doing so, they revealed the impact of considering quality of life. One assumption ignored quality of life and considered only life expectancy, but the cost-effectiveness of bypass surgery for one vessel disease under this assumptio cannot be estimated, since surgery has no effect upon survival. However,

many of the benefits of surgery are directed toward the quality of life rather than survival. If the surgery is performed to relieve mild angina, the cost–utility for vessel disease exceeds $500,000 (1986). A model that did not integrate mortality and morbidity would have missed the benefits for some types of surgery.

In summary, the Weinstein and Stason[72] analysis demonstrates that the BCU of CABG differs according to the characteristic of disease state, but the cost–utility figures compare favourably with those from other widely-advocated medical procedures and screening programmes.

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 have calculated the cost–utility for programmes screening severe hypertension (diastolic > 105 mmHg) to be $4850 per well-year, while the corresponding figure for mild hypertension screening programmes (diastolic 95–104 mmHg) was $9800 per year (1976).[71]

However, their 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 once cases have been detected. The figures given above assume full adherence to the regimen, but substantial evidence reveals that full (100%) adherence is rare.[19] 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. More studies have been devoted to compliance among hypertensive patients than to compliance in any other disease category, and some of these suggest that behavioural intervention can be very useful in increasing adherence to prescribed regimens.[27]

In their analysis, Weinstein and Stason considered the value of programmes designed to increase adherence to antihypertensive medication. Two separate problems were considered, namely there are drop-outs from treatment, and there is failure to adhere to treatments that have been prescribed. The two cases may differ in their cost. One extreme is the patient who fails to see a physician and purchase medication; here the cost would be very low. The other extreme would be the patient who remains under medical care, purchases medications, but does not take them; in this case, the costs would be high. Weinstein and Stason refer to these as the maximum cost assumption and the minimum cost assumption. Under the minimum cost assumption patients do not receive the full benefits of medication because of incomplete adherence, but they also do not spend money. Thus, according to

Weinstein and Stason, the cost-effectiveness under this assumption is very similar to 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 beginning therapy beyond the age of 50. Earlier, it was noted that the costs to produce a well-year for a national sample (US) were $4850 for those with pretreatment diastolic blood pressure greater than 105 mmHg. With incomplete adherence, these values increase to $6400 under the minimum cost assumption and $10,500 under the maximum cost assumption. For mild hypertensive screening (diastolic blood pressure 95–104 mmHg), the $9800 per well-year under the full adherence assumption rose to $12,500 under the minimum cost assumption and $20,400 under the maximum cost assumption.

Since adherence under the maximum cost assumption appears to have a strong effect upon cost–utility, it is interesting to consider the value of behavioural interventions to improve adherence. Several studies have shown the value of behavioural interventions and it is reasonable to assume that a successful behavioural intervention will improve adherence rates by 50% (Ref. 27). Weinstein and Stason considered the cost–utility of interventions that would improve adherence by 50% under the maximum cost assumption. Their analysis of hypothetical programmes that would reduce diastolic blood pressure from 110 mmHg to 90 mmHg suggests a differential expected cost–utility for programmes designed for males and for females. As Figure 3.3 shows, the intervention would improve the cost–utility for both males and females and at each age of therapy reveals the finding from epidemiological studies that blood pressure is better controlled in women than in men. In summary, the analysis demonstrates that even an expensive programme can improve cost–utility because it produces substantial improvements in outcome relative to its costs.

In the Weinstein and Stason monograph, a variety of other hypothetical conditions were considered. Under the assumption that the programme improves adherence by 50%, a significant benefit of the programme remained under the maximum cost assumption. However, under the minimum cost assumption the hypothetical adherence intervention would have a significant benefit if it increased adherence by 50% but no significant effect if it increased adherence by only 20%.

Other assumptions in the Weinstein and Stason analysis need to be considered. For example, they make (and discuss) many other assumptions about the relationship between hypertension and outcome, the linear relationship between adherence and outcome, and the effect of adherence programmes, but provide some data to support the reasonableness of each of these assumptions.

Figure 3.3 shows how very different programmes can be compared usir the system, and summarizes the two Weinstein and Stason studies discusse above.

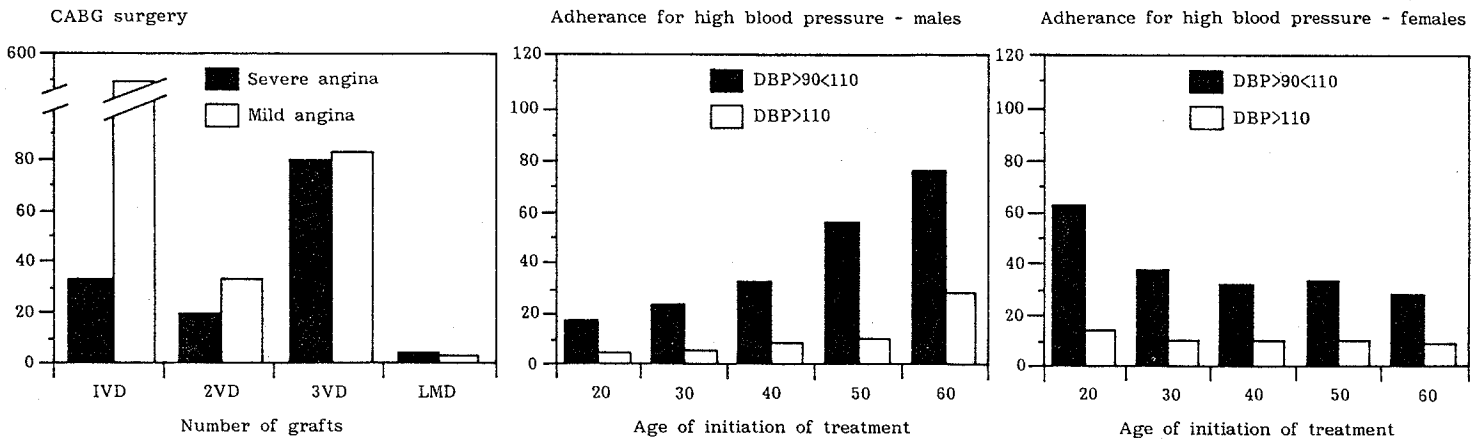


Figure 3.3 Comparison of cost/utility[a] of CABG: surgery and behavioural interventions for increasing adherence[b] with antihypertensive medications for men and women

(Data for comparisons come from Ref. 72, left panel and Ref. 71 for centre and right panels)

[a] All cost/utility figures adjusted to 1986 dollars

[b] Adherence analyses assume that the programme will increase medication adherence by 50%. These calculations were made under the maximum cost assumption

[c] Severe angina is defined as well-being score of 0.7 on 0–1.0 scale. Mild angina is defined as 0.9 on the same scale.

Screening for colon and rectal cancer

Colorectal cancer has become one of the most common cancers in the United States. As a cause of death from cancer, it is preceded only by lung cancer in men and by breast and lung cancer in women, and there will be approximately 140,000 new cases of colorectal cancer diagnosed with an estimated 60,000 deaths caused by this disease in the United States in 1986.[12] The relative 5-year survival rate is 40–50%, and this has not changed for the past 25 years.[12] It has been estimated that survival can be improved to 81% if the lesion is discovered early, while the patient is asymptomatic and the lesion is still localized.[62] Thus, as the 10-year survival approximates the 5-year survival, the value of early diagnosis and treatment is clearly indicated.

Two screening procedures are commonly in use for detection of asymptomatic cancer and precancerous polyps: (i) the faecal occult blood test (FOBT), where the individual is instructed to ingest high-fibre foods t irritate an existing but as yet asymptomatic tumour or polyp, producing bleeding which is subsequently detected in the stool; and (ii) flexible fibre-optic sigmoidoscopy, where the distal 35–65 cm of inner intestine wall is directly observed for lesions.

Part of the benefit of each procedure lies in early detection of cancerous lesions, and part in the identification of polyps for removal. Arguments for and against either procedure have been made on a number of grounds,[64] but no properly quantified assessment had yet been made regarding which of these screening methods was most cost-effective. In a recent paper, Anderson and colleagues used the QWB and General Health Policy Model (GHPM) to go beyond the verbal pro and con arguments for each type of screening.[3] They assessed the comparative BCU of the two screening methods using five different rates of positive findings: three taken from the reported literature for FOBT, three for sigmoidoscopy, with the fourth (artificially high) and the fifth (artificially low) shared by both methods (see Refs 18, 26, 28, 29, 46, 48, 49, 51, 60, 66, 74, 76, 77). The occult blood examples were (1) 3% positives, (2) 1% positives, (3) 5% positives, (4) 1% positives, all cancer, and (5) 7.5% polyps, no cancers. For flexible fibre-optic sigmoidoscopy, the examples involved (1) 7.5% polyps, 0.5% cancer, (2) 2% polyps, 0.5% cancer, (3) 12% polyps, 0.5% cancer, (4) 1% cancer, no polyps, and (5) 7.5% polyps, no cancer.

A number of assumptions were made about the imaginary population to be screened. It was assumed that a person who is 50 years old has the average remaining life-expectancy (28.1 years) for all Americans.[67] The average yearly salary (for calculation of salary savings) was taken from the mos current available figures for all races from the Bureau of Labor Statistics. Also, it was assumed that the age of retirement for all persons is 65 years, and

thus the number of earning years left for a person 50 years old was taken to be 15 years.

Several assumptions were made regarding the screening protocols and tests. First, it was assumed that flexible sigmoidoscopic examination by a primary care provider will cost $120. A $5 charge for the FOBT was kept constant whether the patient had one or six examinations. The frequency of true positives (50%) has been kept constant throughout the range (1–5%) of FOBT positive exams. Of the 50% true positives, 12% are assumed to be cancers and 38% polyps. This figure of 50% is based on the work of large prospective studies which show a 50–60% false-positive rate, and differs significantly from a previously published benefit–cost analysis since it reflects the false-positive rate for cancers and polyps rather than blood loss based on chromium-labelled red blood cells.[20,30,63] Some of these latter false-positives reflect real pathology (gastritis, ulcers, diverticulosis, etc.), but in at least 20% of patients who are FOBT-positive, no pathology is found despite an extensive examination.[30,68] The non-tumour pathology mentioned is, by definition, asymptomatic. No screening protocol has been shown to be effective for these conditions and they are neglected in this study.

Use of a 65-cm flexible sigmoidoscope by a skilled examiner is assumed, with a false-negative rate of zero. It is further assumed that there are no concurrent polyps or cancers should the sigmoidoscope find a single polyp. No out-of-hospital treatment costs are calculated. Most of the above assumptions bias the analysis in favour of FOBT.

The analysis assumes a single evaluation at age 50, and it further assumes the frequency of polyps and cancer and the positive and negative predictive values for FOBT are the same at age 50 as the average of the study populations (usually 40 to 70). It has been assumed 5% of all polyps will progress to become a cancer and that this polyp to cancer transition takes an average of 10 years.[7,21,52,53,55] The noted increase in the 5-year survival for patients diagnosed with colon cancer by screening, as opposed to those diagnosed at the times symptoms present, is not an artifact of lead-time bias. This is supported by the relative levelling-off in mortality after 5 years.[68,75]

The traditional method for comparing a value to be received in the future (e.g. medical expenses averted) with a value to be given at the present time (e.g. the cost of screening and treatment for colorectal neoplasms with occult blood or flexible sigmoidoscopy) is that of discounting. With the aim of looking at future gains and comparing them with the amount of money currently at hand, compound discount-rate multipliers were included in these analyses – for 6% at 10 years (0.558), and for 6% at 15 years (0.417). These discount multipliers were also used with the number of well-years produced (e.g. in looking at mortality prevented and well-years produced in persons with polyps and cancer), since cancer will not be the only cause of death in these persons.

Terms were included in the analysis for (a) screening and treatment costs

(dollars, well-years) for sigmoidoscopy and FOBT; (b) costs (dollars, well-years) of complications for persons undergoing colonoscopy and sigmoidoscopy; (c) FOBT screening costs (dollars, well-years) due to false-negative results; (d) symptoms prevented (well-years) for people with polyps and cancer; (e) prevented hospitalization costs (dollars, well-years) due to screening and polypectomy; (f) mortality prevented (well-years) and salary loss averted due to screening and polypectomy, and (g) mortality prevented (well-years) and salary loss averted owing to cancers discovered by screening.

Final results for all examples for both screening and treatment programmes are given in Table 3.3. For FOBT examples 1–3, well-years are being produced for between $114,103 per well-year and $133,710 per well-year. The figure for FOBT example 4 (1% positive, all cancers) is $41,884 per well-year. The final FOBT example loses both well-years and dollars.

For the flexible fibre-optic sigmoidoscope method, well-years are being produced for between $1489 per well-year and $5045 per well-year in examples 1–3. Example 4 was $5718 per well-year, while example 5 was producing both dollar savings and well-years.

Policy space for benefit–cost/utility (BCU) analysis

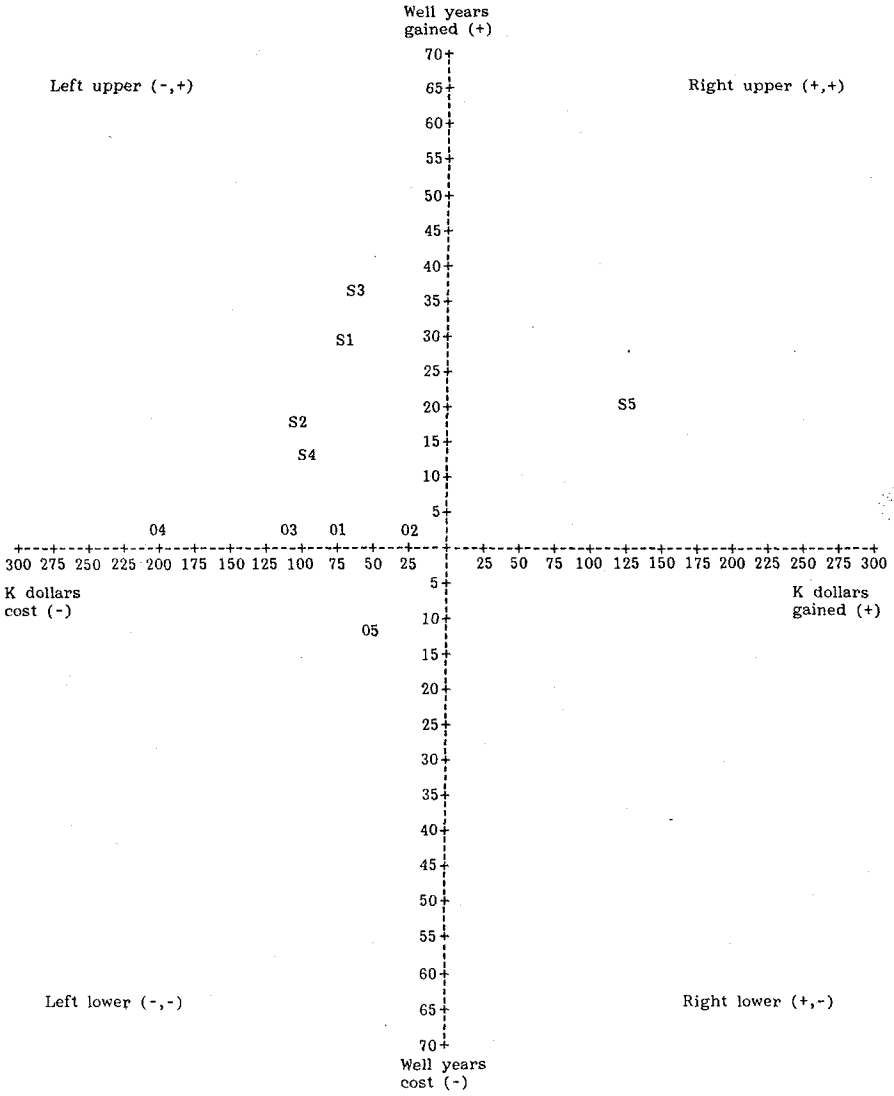
A two-dimensional health policy space (Figure 3.4), where costs (C) are subtracted from benefits (B) (to establish consistency in signs), provides a useful analytical framework for analysing resource allocation and related problems. Net dollar returns (R) per person, where ($R = B - C$) are plotted on the X-axis, and well-years gained or lost per person (Y) are plotted on the Y-axis. Any alternative action (whether treatment, programme, or policy) can be located in this space according to its dollar return and well-years produced or lost.

Alternatives with net economic benefits ($B > C$) fall in the right half of the plane, while programmes with net costs ($C > B$) fall in the left half. Medical treatment and social policies with positive effects ($+Y$) fall in the upper half of the plane, while alternatives (policies or practices) with negative health consequences ($-Y$) fall in the lower half.

The left upper quadrant ($-$, $+$) represents the area for standard BCU analysis, where dollars are being spent to produce well-years, and the negative ratios represent the relative efficiency of the programmes. In the absence of interdependencies and other non-linear constraints, the simple BCU algorithm identifies the optimal set of alternatives within a single budget limit. The left lower quadrant ($-$, $-$) represents unsatisfactory alternatives where dollars are being spent and well-years are being lost. While unusual, this would include ineffective treatments or practices that incur expense and actually do harm in some cases.

Table 3.3 Benefit–cost/utility results of colorectal cancer screening analysis

	Well-year benefits–costs total	*$ benefits minus $ costs total*	*Well-years gained and dollars gained*	*$/well-year benefit–cost utility ratio of programmes*	*Well-years lost and dollars lost by programme*
Occult					
1	0.6119143	−71,842		−117,405	
2	0.2039692	−27,281		−133,751	
3	1.021003	−116,005		−114,009	
4	−28.086800	−164,005		−6,373	
5	13.51552	−86,134			28.0869 + 165,130
Flex Sig					
1	29.16264	−73,114		−2,507	
2	18.55154	−93,592		−5,045	
3	37.84445	−56,360		−1,489	
4	19.98241	127.174	19.98241 + 127,174		
5	14.24024	−81,426		−5,718	



O1: Occult blood, 3% positives.
O2: Occult blood, 1% positives.
O3: Occult blood, 5% positives.
O4: Occult blood, 1% positives, all cancers.
O5: Occult bood, 7.5% positives all polyps.

S1: Sigmoidoscopy, 8% positive.
S2: Sigmoidoscopy, 2.5% positive.
S3: Sigmoidoscopy, 12.5% positive
S4: Sigmoidoscopy, 1% positive, all cancers.
S5: Sigmoidoscopy, 7.5% positives, all polyps.

Figure 3.4 Coordinates for health policy space to compare efficiency of alternative colorectal cancer screening programmes

The right lower quadrant (+, −) represents trade-offs of economic benefits for health, as in studies involving nuclear power, pollution control, occupational, environmental and consumer product safety, highway speed limits, the construction of overpasses and aircraft runways. Analytically, society is willing to sacrifice some well-years in return for substantial economic benefits. The General Health Policy Model can contribute to the analysis of such issues.

All alternatives in the right upper quadrant of policy space (+, +) produce not only well-years but also net dollar returns. The ratio of costs to benefits makes sense in the left upper and right lower quadrants of the policy space, but not in the right upper and left lower, where the outcomes are in a general sense additive (dollar returns plus well-years in the right upper quadrant, dollar costs in addition to well-year losses in the left lower).

Figure 3.4 shows that most options for colorectal cancer screening fall in the upper left quadrant; sigmoidoscopy with 7.5% positives, all polyps, falls in the upper right. Under these conditions, sigmoidoscopy not only produces health benefits, it also saves money. Under the same conditions, however, occult blood falls into the lower left quadrant. Here there are lost resources in addition to health consequences.

Though further study is needed to assess the effects of repeated (yearly) evaluations, the clinical implications of this study are fairly clear: Sigmoidoscopy should be the primary means for screening for colorectal cancer, with FOBT used in the patients with a negative sigmoidoscopy to check for cancers and polyps beyond the reach of the sigmoidoscope. Even though FOBT is the less expensive screening measure, the effects of occult blood false-negatives (missing existing polyps and cancers) is devastating from the BCU point of view.

SUMMARY

This chapter summarizes some of our current thinking on the potential for a general health policy model. We believe the system can be used as an aid for understanding clinical problems. In our previous work we have documented the validity[39] and the reliability[4] of the measurement system, and systematic evaluations of the question structure and reliability of administration have also been performed.[1] Although some authors[18] suggest that the interview procedure is long and tedious, there is substantial evidence that the extra effort results in greater precision of the instrument.[1]

Quality of life is clearly a multi-dimensional construct. However, there is still considerable debate about whether or not multi-dimensional measures are required. Some approaches, such as the SIP, represent the multi-dimensionality by providing quality of life profiles. Other approaches,

including our own, attempt to map the multi-dimensional contruct of well-being onto a uni-dimensional scale of preference or desirability. The choice of a uni-dimensional versus a multi-dimensional approach depends on the purpose of the study. Multi-dimensional approaches may provide more clinical diagnostic information about areas in which there are specific problems, for example, the clinician may learn that arthritis patients have difficulty in ambulation but not in sleep. The uni-dimensional 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. Some investigators prefer measures specific to the symptoms associated with a particular disease. However, we favour the more general approach because it captures both benefits and side-effects (both expected and unexpected) in a single, comprehensive unit. Clearly, debate about the specifics of measure construction will continue. Future research is required to identify the reliability of preference weights, the value of general versus specific measures, and the desirability of interviewer administered versus self-administered questionnaires. The inclusion of general quality of life measures in systematic clinical trials will help elucidate many of these issues.

In this chapter we offer several suggestions for the use of the QWB in clinical studies. Current research is often divided between measurement studies and policy analysis. The General Health Policy Model includes the measurement system described here. When taken in clinical studies, QWB measurements can be used directly in policy analysis; however, few clinical studies have taken QWB measures directly. We hope to see wider application of quality of life measurement in future clinical studies.

References

1. Anderson, J. P., Bush, J. W. and Berry, C. C. Classifying function for health outcome and quality-of-life evaluation. Self versus individual models. *Med. Care* 1986; 24: 454–469.
2. Anderson, J. P., Bush, J. W., Chen, M. and Dolenc, D. Policy space areas and properties of benefit–cost/utility analysis. *JAMA* 1986; 255: 794–795.
3. Anderson, J. P., Ganiats, T. G. and Kazemi, M. M. Screening and Treatment for Colorectal Cancer: A Benefit-Cost/Utility Comparison of Flexible Sigmoidoscopy and Fecal Occult Blood Methods Using the General Health Policy Model. Unpublished manuscript, University of California 1986.
4. Anderson, J. P., Kaplan, R. M., Berry, C. C. *et al*. Interday reliability of function assessment for a health status measure: The quality of well-being instrument. Unpublished manuscript, University of California, San Diego, 1987.
5. Anderson, J. P. and Moser, R. J. Parasite screening and treatment among Indochinese refugees: cost–benefit/utility and the General Health Policy Model. *JAMA* 1985; 253: 2229–2235.
6. Ballugooie, E. Hooymans, J. M. N. and Timmerman, Z. Rapid deterioration of diabetic retinopathy during treatment with continuous subcutaneous insulin infusion. *Diabetes Ca* 1984; 7: 236–242.

7. Bergner, M., Bobbit, R. A., Carter, W. B. and Gilson, B. S. The sickness impact profile; development and final revision of a health status measure. *Med. Care* 1981; 19: 787–806.
8. Bombardier, C., Tugwell, P., Sinclair, A. *et al.* Preference for endpoint measures in clinical trials: results of structured workshops. *J. Rheumatol.* 1982; 9: 798–801.
9. Bombardier, C., Ware, J., Russell, I. J. *et al.* Auranofin therapy and quality of life for patients with rheumatoid arthritis: results of a multicenter trial. *Am. J. Med.* 1986; 81: 565–578.
10. Braumwald, E. Coronary artery surgery at the crossroad. *N. Engl. J. Med.* 1977; 297: 661–663.
11. Bush, J. W., Chen, M. and Patrick, D. L. Cost-effectiveness using a health status index: analysis of the New York State PKU screening program. In Berg, R. (ed.) *Health Status Index*. Chicago, Hospital Research and Educational Trust, 1983, pp. 172–208.
12. Campeau, L., Lesteiance, J., Hermann, J. *et al.* Loss of the improvement of angina between one and seven years after aortocoronary bypass surgery. *Circulation,* 1979; 60 (Suppl. 1): 1–5.
13. Cancer Statistics, 1986, *CA* 1986; 36: 9–25.
14. CASS Principle Investigators and their Associates. Coronary artery surgery study (CASS): A randomized trial of coronary artery bypass survival data. *Circulation* 1983; 68: 939.
15. Chen, M. and Bush, J. W. Maximizing health system output with political and administrative constraints using mathematical programming. *Inquiry* 1976; 13: 215–227.
16. Chen, M. M., Bush, J. W. and Patrick, D. L. Social indicators for health planning and policy analysis. *Policy Sci.* 1975; 6: 71–89.
17. Day, D. W. and Morson, B. C. The adenoma–carcinoma sequence. In Morson B. C. (ed.) *The Pathogenesis of Colorectal Cancer*. Philadelphia, W. B. Saunders Co., 1978, pp. 58–71.
18. Demers, R. Y., Demers, P., Hoar, S. K. *et al.* Prevalence of colorectal polyps among Michigan pattern and model makers. *J. Occup. Med.* 1985; 27: 809–812.
19. DiMatteo, M. R. and DiNicola, D. E. *Achieving Patient Compliance*. Elmsford, Pergamon Press, 1982.
20. Eddy, D. M. *Screening for Cancer: Theory, Analysis and Design*. Englewood Cliffs, Prentice-Hall, 1980.
21. Enterline, H. T., Evans, G. W., Mercado-Lugo, M. D. *et al.* Malignant potential of adenomas of colon and rectum. *JAMA* 1962; 179: 322–330.
22. European Coronary Surgery Study Group. Coronary-artery bypass surgery in stable angina pectoris: survival at two years. *Lancet* 1979; *i*: 889–893.
23. European Coronary Surgery Study Group. Prospective randomized study of coronary artery bypass surgery in stable angina pectoris. *Lancet* 1980; *ii*: 491–495.
24. Fanshel, S. and Bush, J. W. A health status index and its applications to health-services outcomes. *Operations Res.* 1970; 18: 1021–1066.
25. Fielding, J. E. Smoking, health effects and control. *N. Engl. J. Med.* 1985; 313: 491–498.
26. Gibertsen, V. A. and Nelms, J. M. The prevention of invasive cancer of the rectum. *Cancer* 1978; 41: 1137–1139.
27. Haynes, R. B., Sackett, D. L., Gibson, E. S. *et al.* Improvement of medication compliance in uncontrolled hypertension. *Lancet* 1976; *i*: 1265–1268.
28. Hertz, R. E., Deddish, M. R. and Day, E. Value of periodic examinations in detecting cancer of the rectum and colon. *Postgrad. Med.* 1960; 27: 290–294.
29. Hilsabeck, J. R. Experience with routine office sigmoidoscopy using the 60-cm flexible colonoscope in private practice. *Dis. Colon Rectum* 1983; 26: 314–318.
30. Hoffman, A., Young, Q. and Bright-Asare, P. Early detection of bowel cancer at an urban public hospital: Demonstration Project. *CA* 1983; 33: 344–358.
31. Hultgren, H. N., Shettigar, R. and Miller, D. C. Medical versus surgical treatment of unstable angina. *Am. J. Cardiol.* 1982; 50: 663–670.
32. Jaffe, N. *Cataract Surgery and its Complications* (3rd edn). St. Louis, Mosby, 1981.

33. Job, D., Eschwege, B., Guyot-Argenton, C. *et al.* Effect of multiple daily injections on the course of diabetic retinopathy. *Diabetes* 1976; 25: 463–469.
34. Kaplan, R. M. Human preference measurement for health decisions and the evaluation of long-term care. In Kane, R. L. and Kane, R. A. (eds) *Values and Long-term Care.* Lexington, Lexington Books, 1982, pp. 157–188.
35. Kaplan, R. M. Quantification of health outcomes for policy studies in behavioral epidemiology. In Kaplan, R. M. and Criqui, M. H. (eds) *Behavioral Epidemiology and Disease Prevention*. New York, Plenum, 1985, 31–54.
36. Kaplan, R. M. Quality-of-life measurement. In Karoly, P. (ed.) *Measurement Strategies in Health Psychology*. New York, Wiley-Interscience, 1985, pp. 115–146.
37. Kaplan, R. M., Atkins, C. J. and Timms, R. M. Validity of a quality of well-being scale as an outcome measure in chronic obstructive pulmonary disease. *J. Chronic Dis.* 1984; 37: 85–95.
38. Kaplan, R. M. and Bush, J. W. Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychol.* 1982; 1: 61–80.
39. Kaplan, R. M., Bush, J. W. and Berry, C. C. Health status: types of validity for an index of well-being. *Health Serv. Res.* 1976; 11: 478–507.
40. Kaplan, R. M., Bush, J. W. and Berry, C. C. The reliability, stability, and generalizability of a health status index. *Am. Stat. Assoc., Proc. Soc. Stat. Section*, 1978, pp. 704–709.
41. Kaplan, R. M., Bush, J. W. and Berry, C. C. Health status index: category rating versus magnitude estimation for measuring levels of well-being. *Med. Care* 1979; 5: 501–523.
42. Kaplan, R. M. and Ernst, J. A. Do category scales produce biased preference weights for a health index? *Med. Care* 1983; 21: 193–207.
43. Kaplan, R. M., Ries, A. and Atkins, C. J. Behavioral management of chronic obstructive pulmonary disease. *Ann. Intern. Med.* 1985; 7: 5–10.
44. Klein, R. and Klein, B. Vision disorders in diabetes. In *National Diabetes Data Group – Diabetes in America*. NIH Publication, 1985, pp. 85–1468.
45. Kolata, G. D. Consensus on bypass surgery. *Science* 1981; 211: 42–43.
46. Kurnicke, J. E., Walley, L. B., Jacob, H. H. *et al.* Colorectal cancer detection in a community hospital screening program. *JAMA* 1980; 243: 2056–2057.
47. Liang, M. H., Cullen, K. and Larson, M. In search of a more perfect mousetrap (health status or a quality of life instrument). *J. Rheumatol.* 1982; 9: 775–779.
48. Lipshutz, G. R., Katon, R. M., McCool, M. F. *et al.* Flexible sigmoidoscopy as a screening procedure for neoplasia of the colon. *Surg. Gynecol. Obstet.* 1979; 148: 19–22.
49. McCallum, R. W., Meyer, C. T. *et al.* Flexible sigmoidoscopy: diagnostic yield in 1015 patients. *Am. J. Gastroent.* 1984; 79: 433–436.
50. Meenan, R. F. AIMS approach to health status measurement; conceptual background and measurement properties. *J. Rheumatol.* 1982; 9: 785–788.
51. Meyer, C. T., McBride, W., Goldblatt, R. S. *et al.* Clinical experience with flexible sigmoidoscopy in asymptomatic and symptomatic patients. *Yale J. Biol. Med.* 1980; 53: 345–352.
52. Morson, B. The polyp-cancer sequence in the large bowel. *Proc. R. Soc. Med.* 1974; 37: 451–457.
53. Morson, B. C. Evolution of cancer of the colon and rectum. *Cancer* 1974; 34: 845–849.
54. Mosteller, R. Innovation and evaluation. *Science* 1981; 211: 881–886.
55. Muto, T., Bussey, H. J. R. and Morson, B. C. The evolution of cancer of the colon and rectum. *Cancer* 1975; 36: 2251–2270.
56. National Center for Health Care Technology, Technology Assessment Forum. Coronary artery bypass surgery. *JAMA* 1981; 246: 1645–1649.
57. National Institutes of Health. *Epidemiology of Respiratory Diseases Task Force Report*. Washington, DC, US Government Printing Office, NIH Publication No. 81-2019, 1979.
58. Office of Technology Assessment, US Congress. *A Review of Selected Federal Vaccine and*

Immunication Policies: Bases on Case Studies of Pneumococcal Vaccine. Washington, DC, US Government Printing Office., 1979.
59. Patrick, D. L., Bush, J. W. and Chen, M. M. Toward an operational definition of health. *J. Health Soc. Behav.* 1973; 14: 6–23.
60. Ribet, A., Escourou, J., Frexinos, J. *et al.* Screening for colorectal tumors – results of two years experience. *Cancer Detect. Prev.* 1980; 3: 449–461.
61. Schade, D., Santiago, J., Skyler, J. *et al.* Effects of intensive treatment on long-term complications in intensive insulin therapy. In Schade, D., Santiago, J., Skyler, J. *et al.* Intensive Insulin Therapy. *Princeton Excerpta Medica*, 1983, Chapter 5.
62. Shahon, D. B. and Wangensteen, O. H. Early diagnosis of cancer of the gastrointestinal tract. *Postgrad. Med.* 1960; 27: 306–311.
63. Simon, J. B. Occult blood screening for colorectal carcinoma: a critical review. *Gastroenterology* 1985; 88: 820–837.
64. Sugarbaker, P. H., Gunderson, L. L. and Witties, R. E. Colorectal cancer. In DeVita, V. T. Jr., Hellman, S. and Rosenberg, S. A. (eds) *Cancer: Principles and Practice of Oncology.* Philadelphia, J.B. Lipincott Co., 1985, pp. 829–831.
65. Tchobroutsky, G. Relation of diabetic control to development of microvascular complications. *Diabetologia* 1978; 15: 143–152.
66. Traul, D. G., Davis, C. B., Pollock, J. C. and Scudamore, H. H. Flexible fiberoptic sigmoidoscopy – the Monroe Clinic experience: a prospective study of 5000 examinations. *Dis. Colon Rectum* 1983; 26: 161–166.
67. US Bureau of the Census. *Statistical Abstract of the United States.* Washington DC, Dept. of Commerce, 1985.
68. US Department of Health Education and Welfare. *Cancer Patient Survival, Report #5*, Washington DC, Public Health Service, National Institutes of Health, 1976.
69. US Department of Labor. Analyzing 1981 *Earnings Data From the Current Population Survey.* Bureau of Labor Statistics publication 2149, 1982.
70. Ware, J. E., Johnston, S. A., Davies-Avery, A. *et al.* Conceptualization and measurement of health for adults. In *The Health Insurance Study, Vol. 3, Mental Health.* Santa Monica, Rand, 1979.
71. Weinstein, M. C. and Stason, W. B. *Hypertension: A Policy Perspective.* Cambridge, Harvard University Press, 1976.
72. Weinstein, M. C. and Stason, W. B. Cost-effectiveness of coronary artery bypass surgery. *Circulation,* 1982; 66 (Suppl. 5, pt 2): III 56–66.
73. Weinstein, M. C. and Stason, W. B. Cost-effectiveness of interventions to prevent or treat coronary heart disease. *Annu. Rev. Public Health*, 1985; 6: 41–63.
74. Wherry, D. C. Screening for colorectal neoplasia in asymptomatic patients using flexible fiberoptic sigmoidoscopy. *Dis. Colon Rectum* 1981; 24: 521–522.
75. Winauer, S. J., Andrews, M., Miller, C. H. *et al.* Review of screening for colorectal cancer using fecal occult blood testing. In Winawer, S. J., Schottenfeld, D. and Sherlock, P. (eds) *Colorectal Cancer: Prevention, Epidemiology, and Screening.* New York, Raven Press, 1980, pp. 249–259.
76. Winnan, G., Berci, G., Pannish, J. *et al.* Superiority of the flexible to the rigid sigmoidoscope in routine proctosigmoidoscopy. *N. Engl. J. Med.* 1980; 302: 1011–1012.
77. Yarborough, G. W. and Waisbren, B. A. The benefits of sytematic fibreoptic sigmoidoscopy. *Arch. Intern. Med.* 1985; 145: 95–96.