Risk in Perspective

Measuring Quality-Adjusted Survival for Cancer Therapies

Jane C. Weeks, M.D.

...we explore scientific progress in the application of utility theory to measurement of quality of life in cancer patients.

Choosing between alternative cancer treatment strategies often involves weighing trade-offs between length and quality of life. For example, small gains in survival resulting from aggressive chemotherapy must often be weighed against potential quality-of-life costs to patients resulting from treatment-related toxicity. As a result, interest in assessing quality of life in oncology is long-standing. For example, Karnofsky's pioneering work on assessing the performance status of cancer patients was one of the earliest efforts to collect standardized descriptive health-related quality-of-life data. But such descriptive data are not ideal for evaluating trade-offs between length and quality of life. Another approach, less familiar to clinicians, addresses such trade-offs explicitly. Utilities, defined as the quantitative measure of the strength of a person's preference for an outcome, have their roots in economic theory and decision analysis. Like conventional health status measures, they assess health-related quality of life, but they differ from descriptive measures in that they reflect how a subject values a state of health, not just the characteristics of the health state. In this issue of **RISK IN PERSPECTIVE**, we explore scientific progress in the application of utility theory to measurement of quality of life in cancer patients.

QUALITY-ADJUSTED SURVIVAL

Quality-adjusted survival (QAS), a single measure that captures both length and quality of life, facilitates assessment of therapeutic trade-offs. It has clinical appeal because it makes explicit the factors that physicians and patients routinely consider implicitly in making treatment decisions. In addition, QAS, measured in Quality-Adjusted Life Years (QALYs), is the standard measure of outcome used in decision analysis and cost-effectiveness analysis. Therefore, there is growing interest by policy makers in evaluating the impact of alternative medical interventions on QAS.

The only appropriate measure of quality of life for use in calculating QAS is the utility. Utilities, unlike descriptive measures of quality of life, are suitable for calculating QAS because they are elicited in a way that ensures that the product of length of life and quality of life is a meaningful number. The measurement technique must guarantee, for example, that one year of life at a utility of x is exactly as desirable as 6 months of life at a utility of $2x$.

MEASUREMENT TECHNIQUES

In general, utilities are global measures of quality of life in a given state of health, and by convention, are measured on a scale from 0 to 1, where 0 represents death and 1 represents excellent health. Approaches to measurement of utilities have grown out of economic and decision theory rather than empiric research. These techniques tend to be relatively cumbersome and have not undergone the same rigorous psychometric validation as many well-accepted quality-of-life questionnaires.

The three most commonly used approaches to direct utility assessment include the standard (or reference) gamble,

the time trade-off, and the rating scale. The *standard gamble* assesses the respondent's utility for his or her own quality of life (or that of a hypothetical health state) by asking what probability of immediate death (s)he would be willing to risk in exchange for return to excellent health. The *time trade-off*, in contrast, asks how much life expectancy one would be willing to give up in exchange for a return to excellent health. Both techniques are conceptually complex and, even when administered by an interviewer using visual aids or by a specially-designed computer program, are difficult for some respondents to understand.

The *rating scale*, which asks the respondent to rate his/her quality of life on a visual or verbal numeric scale, is much simpler to administer, but does not generate a true utility. For example, there is no reason to believe that a respondent who assigns a state of health a score of 75 on a 100 point rating scale would be willing to give up one quarter of his/her life expectancy in exchange for a return to perfect health. In general, rating scales' values for a given health state tend to be lower than values obtained with standard gamble or time-trade-off techniques for the same health state. Empiric data suggest that the mean utility for a population may be predicted from the mean rating scale value by a "transformation" that adjusts the score upward.

SOURCES OF UTILITIES

The optimal source of utility data depends on the way it is to be used. If the goal is to identify the best treatment strategy for an individual patient, then it is clearly the patient's own preferences for the possible outcomes that matter. But because of the complexity of direct utility elicitation from patients and other practical obstacles, formal utility assessment to identify the course of treatment likely to maximize that patient's QAS is rarely a part of actual clinical decision-making.

The principal role of utility assessment is in identifying the optimal therapy for the "typical" patient (the clinical trial setting) and in determining how much value per

dollar is generated for a population in opting for that treatment over a less expensive alternative (cost-effectiveness analysis). In the past, the most frequently used source of utilities for both of these applications was "expert judgment," usually consisting of the analyst's own guesses or a survey of a few physicians. More recently there has been growing recognition that empiric utility data provide a much sounder basis for making these important decisions. Patients experiencing the health states of interest are an important source of utility data for use in analyzing clinical trials and decision analysis. They are the real "experts", of course. There is an emerging consensus, however, that for cost-effectiveness analysis, the optimal source of utility estimates is the general population, assuming that they are well-informed about the nature of the illness. The argument is that society's resources should be allocated according to the preferences of the general population.

A recent national panel on cost-effectiveness methods identified health indices as the preferred approach for obtaining societal utilities for health states. This hybrid between traditional quality of life measurement and utilities consists of two parts [see figure]. A quality-of-life survey is used to collect descriptive data from patients who are in the health state(s) of interest. The second component is a formula that is applied to these descriptive data to generate a utility. The formula comes from a one-time general population survey and reflects the relative importance assigned to different domains of quality of life by the survey population, measured using one of the direct utility elicitation techniques described above. Examples of such health indices include the Quality of Well-Being Index and the Health Utilities Index. Approaches currently undergoing validation include EuroQOL, a measure specifically designed for international use, and the Q-tility Index, a cancer-specific tool.

Some have argued for the adoption of a single, generic health index to be used by all analysts in calculating QALYs. The obvious appeal of this approach is that it

The Multiattribute Utility Approach
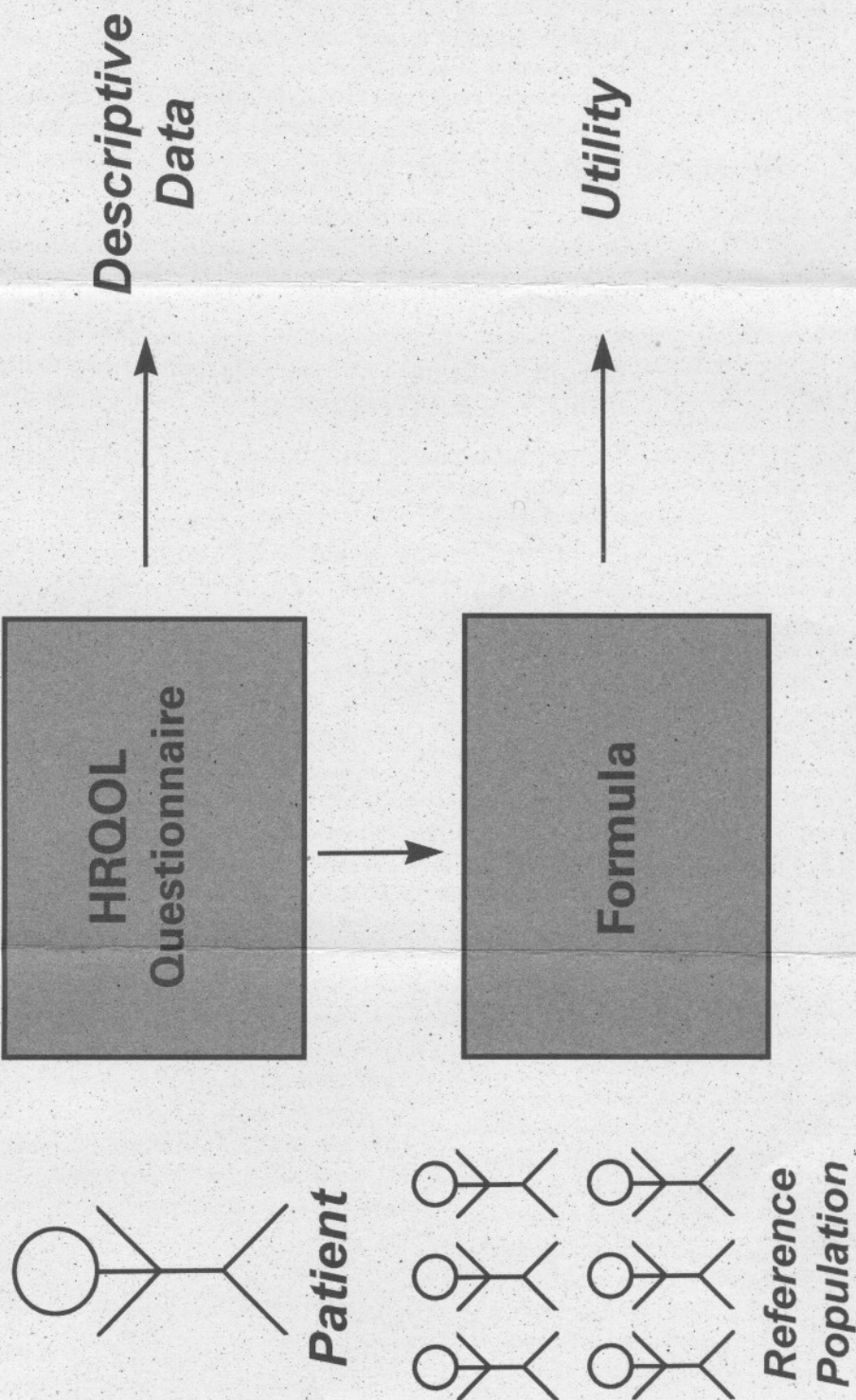
Patient

HRQOL Questionnaire → Descriptive Data

Formula → Utility

Reference Population

FURTHER READING

Tsevat J, Weeks JC,
Guadagnoli E, et al:
Using health-related
quality of life information:
Clinical encounters, clini-
cal trials, and health pol-
icy. *J Gen Intern Med*
9:576–582, 1994.

Torrance GW:
Measurement of health
state utilities for eco-
nomic appraisal. A
review. *J Health Econ*
5:1–30, 1985.

O'Leary JF, Fairclough
DL, Jankowski MK,
Weeks JC: Comparison
of time-tradeoff utilities
and rating scale values
in cancer patients and
their relatives: Evidence
for a possible plateau
relationship. *J Med Dec
Making* 15:132–137, 1995.

Weinstein MC, Siegel JE,
Gold MR, et al:
Recommendations of the
Panel on Cost-
Effectiveness in Health
and Medicine. *JAMA*
276:1253–1258, 1996.

Gelber RD, Cole BF,
Gelber S, Goldhirsch A.
The Q-TWiST method. In:
*Quality of Life and
Pharmacoeconomics in
Clinical Trials.* Second
Edition. B. Spilker Ed.
Lippincott-Raven
Publishers, Philadelphia,
1996, pp 437–444.

would ensure comparability of cost-effec-
tiveness ratios across conditions and thus
facilitate policy decisions about where to
invest limited health resources. But such a
generic scale is likely to be less sensitive
to the domains of quality of life known to
be important in a particular disease (such
as functional status and fatigue in cancer,
for example) than a disease-specific health
index. If disease-specific indices are to be
used they must place these domains
within the full spectrum of death to excel-
lent health so that the utilities they generate
are on the same scale as those produced
by generic measures.

CALCULATION OF
QUALITY-ADJUSTED SURVIVAL

The optimal approach to calculating QAS
depends on the source and nature of the
survival data. If survival is estimated in a
decision analytic model, QAS is calcu-
lated by simply assigning a utility weight
to each health state in the model. For
survival data observed in a clinical trial,
the conventional approach is to measure
utilities at regular intervals, and then find
the area under the curve of survival
versus mean utility.

Q-TWiST (Quality-adjusted Time
Without Symptoms of disease and
Toxicity of treatment) is an alternative
method of estimating QAS from clinical
trial data. In this approach, survival time
spent in defined clinical health states by

patients in a trial is weighted by a utility
specific to each health state. In contrast to
the calendar method, which requires a
mean utility value at regular intervals, the
Q-TWiST approach requires only one
utility value for a given clinical health
state. It is particularly helpful for analyz-
ing clinical trial data when the health
states involved are clearly defined,
expected to differ substantially in quality
of life, and last for more than a month or
two.

CONCLUSIONS

Recent changes in the health care envi-
ronment have created a new demand for
information on whether the benefits of
health interventions justify the costs, and
whether small gains in survival from new,
aggressive treatments justify the associ-
ated toxicity. QAS is a critical metric in
answering both these questions. Consensus
on how best to measure QAS has not yet
been reached, but progress is being made.
Generic and disease-specific health indices
for measuring societal preferences for
health states are being developed.
Practical strategies for measuring utilities
during clinical trials are being tested.
And, innovative approaches for calculating
QAS are being devised. For these reasons,
the assessment of patient well-being has
become one of the most dynamic fields in
decision making.