Risky Choices in Medical Decision Making: A Survey

by Milton C. Weinstein*

1. Introduction: Medical decision are risky decisions

Economic models of decision making under uncertainty presume five features of a decision-making situation. The first is a set of well-defined *alternatives* from which to choose. The second is *uncertainty* in the relation between the chosen alternative and the actual consequence. The third feature is a set *preferences* held by the decision maker and applied to the possible outcomes. Fourth, one or more *constraints* may limit the set of available alternatives. Finally, and most importantly, the *decision must be made*; to postpone a decision or to preserve the status quo is itself a decision with its attendant consequences.

While these features are shared by many economic decisions under uncertainty, such as in purchasing insurance or in assembling an investment portfolio — and it was surely these sorts of decisions that von Neumann and Morgenstern (1947) had in mind when they developed their now classical theory of economic choice under uncertainty — these are features also shared by many decisions that arise in clinical medicine and public health. Physicians must choose between alternatives: whether to perform surgery or to administer medication, which antibiotic to use, whether to perform a risky test. Consequences of action are uncertain: Will the patient survive the operation? What disease does the patient have? What will the test show? Value judgments arise in evaluating outcomes: Is the risk of surgical mortality worth taking to relieve the pain? Are the side effects of the medicine worse than the disease? Budgetary constraints limit today's health care provider in the same way that the classical economic consumer is limited by his or her income. Finally, the decision cannot be avoided, however painful the decision-making process may be.

By imposing various axioms of rational behavior on the decision maker — such as adherence to the laws of probability and transitivity of preferences — models such as the expected utility model of von Neumann and Morgenstern (1947) have been developed. These models can, and have, been used either to *describe* how decisions are made in fact, or to *prescribe* optimal decisions, that is, to help decision makers decide in a manner consistent with their own beliefs and preferences.

* Professor of Policy and Decision Sciences, Harvard School of Public Health, Boston, Massachusetts 02115 U.S.A. Presentation at the Ninth Annual Lecture of the Geneva Association, Faculté Universitaire Catholique de Mons, Belgium, 8 November 1985.

Researchers have found the expected utility model and its underlying axioms wanting as a *descriptive* model of economic decisions (Slovic et al., 1977; Kahneman and Tversky, 1979; Schoemaker, 1982). The same could be said of medical decisions, which are determined by the interplay of numerous factors, many of which defy rationality. Among these determinants of medical decisions are: tradition in the medical profession or habit of the individual clinician; the physician's knowledge — of lack of knowledge — of the effects of the clinical procedures being considered; patients' expressed or implied wishes or demands; physicians' preferences for, or aversion to, certain procedures or risk-taking situations; economic incentives faced by clinicians or the organizations in which they practice; legal considerations and the threat of malpractice suits; psychological factors such as curiosity, anxiety, or guilt; and socio-medical factors such as conformity to expectations of one's peers.

Despite the limitations of decision theory as a *descriptive* model, the *prescriptive* formulation of the theory of decision making under uncertainty, widely known as decision analysis (Raiffa, 1968), has carried over from its origins in business management to become a useful and practical approach to rational decision making in medicine (Weinstein et al., 1980). Decision analysis has also become a cornerstone in the theory and practice of technology assessment in the health care industry. The theory and application of decision analysis as a prescriptive tool in medical decision making and allocation of health care resources is the subject of my talk today.

2. Decision analysis as a prescriptive tool

Decision analysis was developed as a management tool in business administration, to offer guidance on decisions in production, marketing, and research and development. From modest early efforts around such questions as whether and where to drill for oil, the practice of decision analysis has grown to a multi-million dollar consulting industry, and many major corporations employ decision analysis among the many analytic techniques used to support investment and resource allocation decisions. Decision analysis is now taught in virtually every school of business administration in North America, and is regarded by students and graduates as a valuable part of their professional education.

Development and diffusion of decision analysis in the medical context has lagged behind business applications, but there has been a period of important theoretical and practical growth during the past decade. Professors of management have long found medical examples useful in their teaching, because medical decisions fit the decision-analytic paradigm so well. Indeed, one of the earliest applications of decision analysis in medicine was a doctoral thesis — at Harvard Business School — on how to manage the common sore throat.

During the 1970's, however, the medical community discovered decision analysis, with the result that medical researchers are now using decision-analytic concepts in their clinical publications, and it is now even possible for medical practitioners to perform decision analyses at the bedside, with the aid of a personal computer and software packages available for the cost of one or two days in the hospital.

Medical educators are using decision-analytic concepts to teach future doctors principles of clinical reasoning.

At the level of health care policy and planning, governmental agencies are combining methods of decision analysis and cost-effectiveness analysis as a basis for regulation and allocation of health resources.

Pharmaceutical and medical equipment manufacturers are using decision analysis to evaluate their products and as a medium for marketing them to clinicians and purchasers.

The insurance industry, too, is becoming increasingly aware of the need to make choices among potential medical technologies for reimbursement, in the presence of uncertainty.

The remainder of my talk will be in four sections. First, I will illustrate the process of clinical decision analysis using a simple decision of whether to perform a risky diagnostic test. This example will highlight four elements of decision analysis — structuring the problem with the aid of a decision tree, assigning probabilities with the aid of Bayes' Theorem, determining an optimal strategy by averaging out the decision tree, and performing sensitivity analysis to test whether the conclusions are stable when certain assumptions are changed. This first example will not introduce problems of values, preferences, or "utilities," nor will it specifically address the problem of medical choices under resource constraints. These two aspects of medical decision making — preferences under uncertainty, and resource constraints — will be considered in subsequent segments of my talk. Finally, I will comment on some of the limits and challenges for those who would like to see a more rational basis of medical resource allocation become a reality.

3. Decision analysis and evaluation of risky diagnostic tests

The first example is based on the following clinical scenario:

Scenario: Patients displaying signs of chronic progressive liver failure may have one of at least two conditions requiring different treatments: chronic hepatitis or cirrhosis. Assume that if a patient has chronic hepatitis, treatment with steroids will increase the probability of two-year survival from 67% to 85%. If, in fact, the patient has cirrhosis, treatment with steroids will do no good. All steroid treatment carries a risk of severe complications; assume this would lower the two-year survival of a cirrhotic from 50% to 48%. Assume that a liver biopsy can be performed to improve the diagnostic basis for treatment. Specifically, assume that the biopsy will indicate that hepatitis is present in 90% of patients who truly have hepatitis, and that it will indicate cirrhosis in 95% of patients who truly have cirrhosis. Despite the accuracy of the biopsy, there is a 1% chance of these sick patients dying from the biopsy itself. Consider only the alternative diagnoses of hepatitis and cirrhosis in a municipal hospital population where 80% of chronic progressive liver failure is caused by cirrhosis. If the outcome you care about is "survival at two years," should the patients undergo liver biopsy?

I have just presented you with quite a lot of data, more than the human mind can easily digest. Let us examine this information in a form that will help to analyze the decision problem (Table 1).

First, we note that there are two basic sets of choices. The first is whether to perform the biopsy. The second is whether to administer steroids; this decision might be faced in the absence of a biopsy, following a biopsy that shows hepatitis, or following a biopsy that shows cirrhosis.

Next, we review the probabilities. One set of probabilities describes the diagnostic accuracy of the liver biopsy: 90% accurate in the presence of hepatitis, and 95% accurate in the presence of cirrhosis. (The 90% figure is often called the *sensitivity* of the test for hepatitis, and the

Table 1: Liver biopsy — elements

- Starting point: patients at municipal hospital with chronic progressive liver failure.
 - Choices
 - Test: biopsy vs. no biopsy
 - Treat: steroids vs. no steroids
 - Probabilities
 - Information:
 - $P[\text{Biopsy shows hepatitis} \mid \text{hepatitis}] = 0.90$
 - $P[\text{Biopsy shows cirrhosis} \mid \text{cirrhosis}] = 0.95$
 - Clinical states:
 - $P[\text{Hepatitis}] = 0.2$, $P[\text{Cirrhosis}] = 0.8$
 - $P[\text{Survive biopsy}] = 0.99$
 - Outcome: survival at two years
 - = 85% with hepatitis and steroids
 - = 67% with hepatitis and no steroids
 - = 50% with cirrhosis and no steroids
 - = 48% with cirrhosis and steroids
-

95% figure is often called the test *specificity*.) Another set of probabilities relates to clinical states: a 20% chance that the patient has hepatitis, an 80% chance of cirrhosis, and a 99% chance of surviving the biopsy. The last set of probabilities relates to the ultimate outcome of interest, the chances of two-year survival with each possible combination of treatment and disease status:

- 85% with hepatitis and steroids,
- 67% with hepatitis and no steroids,
- 50% with cirrhosis and no steroids,
- 48% with cirrhosis and steroids.

Next, let us consider a decision tree, or decision-flow diagram, for this problem (Figure 1). The diagram is read from left to right; nodes indicated by a square are decisions controlled by the decision maker, nodes indicated by a circle are chance events governed by probabilities. Each path through the tree is associated with a probability of survival, such as 50% for the lower path of no biopsy, no steroids, and cirrhosis.

In order to analyze the tree, one must assign probabilities to each chance event. This is shown (in Figure 1) for the events “dies” or “survives” immediately following the biopsy, and for the disease states “hepatitis” and “cirrhosis” along the branch corresponding to “no biopsy”.

The remaining probabilities require additional analysis, because they are not available in the form required. For example, we know the probability that the biopsy will show hepatitis if the patient has hepatitis, but we do not know the probability that the patient has hepatitis if the biopsy shows hepatitis. We know the probability that the patient has hepatitis, but we do not know the probability that the biopsy will show hepatitis. So we need to apply the statistical

Figure 1: Decision tree for the liver biopsy

trick known as Bayes' Theorem. In our experience, this procedure is most clearly illustrated by using the display shown in Figure 2. As shown, the probability that the biopsy shows hepatitis is $(90\%)(20\%) + (5\%)(80\%) = 22\%$. The probability that such a patient actually has hepatitis is $18\%/22\%$, or 82% . The other "missing" probabilities in Figure 1 can be obtained analogously from Figure 2. All of the probabilities are shown in the completed decision tree (Figure 3).

Figure 2: Liver biopsy: use of 2x2 table

	H	C	
Biopsy says H	$0.9 \times 0.2 = 0.18$	0.04	0.22 (= 0.18 + 0.04)
Biopsy says C	0.02	$0.95 \times 0.8 = 0.76$	0.78 (= 0.02 + 0.76)
	0.20	0.80	1.00

$$P[H \mid \text{Biopsy says H}] = 0.18/0.22 = 0.82$$

$$P[C \mid \text{Biopsy says C}] = 0.76/0.78 = 0.97$$

Figure 3: Analysis of the liver biopsy

The analysis of the decision tree proceeds from right to left; we seek the value (in terms of two-year survival) of finding oneself at each node in the tree. Thus, for example, a patient who has had a liver biopsy, who survives the biopsy, whose biopsy shows hepatitis, and who receives steroids, has a 78.3% chance of survival ($78.3 = (0.82)(85) + (0.18)(48)$). A patient who has had a liver biopsy, who survives, and whose biopsy shows hepatitis, also has a 78.3% chance of survival, because the physician has the option of *choosing* steroids over no steroids. (Mathematically, the operation at decision nodes is maximization, and the operation at chance nodes is expectation.) Proceeding further to the left, a patient who has had a liver biopsy and survives, has a two-year survival of 56.6% ($56.6 = (0.22)(78.3) + (0.78)(50.5)$). Finally, a patient who has had a liver biopsy has a two-year survival of 56.03%. This is greater than the 55.4% survival if no biopsy is performed, so the optimal clinical decision is to perform the biopsy.

We can test the sensitivity of this conclusion to many variables, such as the risk of the biopsy or the *a priori* probability that the patient has hepatitis. We can, in fact, compute threshold values of these variables, above which the optimal decision switches to no biopsy.

Computation of the threshold for biopsy risk is easy. If the biopsy's mortality rate is m , then the percentage survival with the biopsy is $(1-m)(56.5)$, and this exceeds 55.4 whenever $m < 2.1\%$. As long as the biopsy results in the death of not more than 2.1% of patients, it is worth performing.

The threshold for the prior probability of hepatitis involves a less direct calculation. We note that the higher the probability of hepatitis, the less useful the biopsy, because the likelihood that the biopsy would cause a switch in therapy from steroids (the optimal choice without a biopsy) becomes smaller. It is possible to show that if the *a priori* probability of hepatitis is greater than 34.7%, the optimal strategy switches to no biopsy and treat with steroids.

(It is, of course, possible to compute thresholds for more than one variable simultaneously. If p is the prior probability of hepatitis, then the threshold surface for m and p is given by

$$m = (1.9 - 3.7p)/(33.3p + 49.9)$$

For lower m , biopsy is optimal; for higher m , no biopsy is optimal.)

Before turning to the next example, I would like to make four observations on this example. The first is that the example has been formulated in such a way as to avoid applying value judgments to the outcomes. By asserting that the objective is to maximize the probability of two-year survival, such issues as tradeoffs between short-term and long-term survival, or risk aversion, are avoided.

The second observation is that the example ignores resource constraints, viewing clinical procedures such as the biopsy as free goods with no opportunity costs. Explicit consideration of the opportunity cost of medical resources might lead to different conclusions.

My third observation on this example is that the unaided intuition of most persons — including experienced clinicians — who are exposed to this example is discrepant from the results of the analysis. This may be attributable, in part, to subconscious insertion of additional factors into the problem, such as the belief that death from the biopsy is worse than death from the disease. Or, it may be attributable to the fact that human beings are imperfect information processors (Tversky and Kahneman, 1974) who can benefit from formal decision-making models.

My fourth observation on this example concerns the probabilities. How are these probabilities to be estimated? For example, where do we get the estimate that 20 percent of patients presenting with chronic liver failure at a municipal hospital have hepatitis? Ultimately, the decision-theoretic model requires that such probabilities reflect the degree of belief held by the decision maker. However, such degrees of belief may, and should, be informed by data, such as the proportions of similar patients found in previous years at similar hospitals to have hepatitis. One problem is: What does one mean by *similar* patients at *similar* hospitals? Statistical techniques ranging from simple estimation from stratified data to multivariate regression models may be used to obtain probability estimates specific to particular patients or patient groups. In the absence of objective data, structured judgments of expert panels may be elicited, albeit at the risk of merely amplifying the biases in probability judgment just mentioned. From the clinician's point of view, decision analysis has both the advantage and disadvantage of forcing explicit attention on estimates that are often left implicit under the cloak of "clinical judgment".

4. Decision analysis and evaluation of risky treatments

My next two examples illustrate the role of preferences in medical decision making. The first of these is due to Barbara McNeil and her colleagues (McNeil et al., 1978); the second is based on my work with William Stason (Weinstein and Stason, 1981).

In the first example, we consider a patient with early-stage, resectable carcinoma of the lung. Two treatment alternatives, surgery and radiation therapy, are available. Since the quality of life, on average, is equivalent for the two treatments, we focus on survival, the probabilities of which are shown in Figure 4. After 2.5 years, the chance of survival with "typical"

Figure 4: Survival curves for operation and radiation therapy for a 60-year old (McNeil et al., 1978)

Figure 5: Construction of a utility curve (McNeil et al., 1978)

surgery is greater than with radiation therapy; a shorter-term horizon, however, favors radiation therapy because of the risk of surgical mortality. A patient whose overriding objective is to survive the next two years would clearly elect radiation therapy, while a patient whose overriding concern is to be alive in five years' time would elect surgery. In between, various shades of preference might lead to either alternative as optimal.

Dr. McNeil and her colleagues analyzed this problem by applying formal expected utility analysis. The utility values for survival to various points in times were obtained by interviewing patients with lung cancer. A typical utility schedule is shown in Figure 5. For this patient, survival to 5 years has a utility of 0.50, meaning that a certainty of 5 years is equally desirable as a 50-50 gamble between zero and 25 years survival. Survival to 1.5 years has a utility of 0.25, meaning that a certainty of 1.5 years is as desirable as a 25-75 gamble between zero and 25 years, or a 50-50 gamble between 0 and 5 years. The convex shape of this curve indicates that the patient is *risk averse* with respect to years of survival. This attitude may, in turn, reflect any or all of three components: diminishing marginal value of additional years of life, discounting the future, and pure risk aversion.

These researchers went on to calculate expected utilities for a number of patients, and found that a substantial number who had elected surgery might actually have preferred radiation therapy. It is noted that surgery offers the greater life expectancy (or expected utility if utility is linear) and also the greater five-year survival (or expected utility if utility is a step function), but is not optimal if the patient is sufficiently risk averse.

The second example of decision analysis involving preferences concerns coronary artery bypass surgery, a procedure now being performed 200,000 times per year in the United States. For most patients, the main reason for coronary artery bypass surgery is to relieve the pain and disability of angina pectoris. Except for those whose coronary artery disease involves the largest coronary artery — the left main coronary artery — or those with disease involving all three major arteries, a tradeoff exists between improved quality of life and somewhat reduced life expectancy.

We performed an analysis in which utilities were assigned to each of several health states, defined by the level of pain. Angina could be absent (the best state), moderate, severe, or very severe. By making various assumptions about the structure of preferences for life years and health states (Pliskin et al., 1980), the utility function reduces to quality-adjusted life years: the number of years of survival, weighted by a fraction between 0 and 1 reflecting the quality of life during those years. High weights (such as 0.9) correspond to mild pain and disability; lower weights (such as 0.5) correspond to extreme pain or disability. The weights may be derived either from a standard von Newman-Morgenstern lottery, from an indifference relation or "time trade-off", or from a subjective category rating (Figure 6). The choice of method involves trade-offs between theoretical merit (high for the lottery method, low for category rating) and ease of administration and reliability (low for lottery, high for category rating). Torrance (1984) has suggested using the category rating method in practice, but adjusting the results according to an empirical relation he has found between category ratings and lottery probabilities for comparable outcomes.

The results of our analysis are presented in terms of gains in quality-adjusted life expectancy (Figure 7). As the weight assigned to life with symptoms drops, the advantage of surgery increases for patients with disease of only one or two vessels. Extensions of this analysis, in which time preference and risk aversion are also taken into account, have been developed.

*Figure 6: Three methods for assigning weights to health states.
 (In this illustration, the health state being evaluated is "severe angina".)*

Figure 7: Estimated differences in quality-adjusted life expectancy between surgery and medical management by number of diseased vessels and quality weights.

$Q = 1.0$ is equivalent to no angina or no concern over pain;

$Q = 0.9$ is equivalent to mild angina and a sedentary lifestyle;

$Q = 0.7$ is equivalent to severe angina and an active life style;

$Q = 0.5$ is equivalent to very severe angina or serious psychological effects.

1VD = one-vessel disease, etc.

(Source: Weinstein and Saron, 1981.)

The resulting utility metric—quality-and-risk-adjusted life years—has considerable theoretical appeal, but has not yet gained widespread use. The interaction between risk posture and time preference creates troublesome theoretical and practical problems in this regard, however.

Decision analyses, many involving explicit utility analysis, have been performed for dozens of risky treatment decisions and diagnostic tests, in virtually all medical specialties and subspecialties. A detailed, but somewhat dated, bibliography may be found in our text (Weinstein et al., 1980). Recent examples may be found in the medical literature, including the new international journal, *Medical Decision Making*. Among the clinical decisions for which published decision analysis exist are: coronary artery bypass surgery versus medical management (Pauker, 1976; Weinstein and Stason, 1981), prenatal diagnosis of Down's syndrome (Pauker and Pauker, 1977), diagnostic evaluation of jaundice (Richter et al., 1983); prophylactic cholecystectomy versus expectant management of silent gallstones (Ransohoff et al., 1983), management of Herpes simplex encephalitis (Braun, 1980; Barza and Pauker, 1980), treatment choice for ovarian carcinoma (Simes, 1985), management of dyspepsia (Read et al., 1982), use of estrogen replacement therapy in the menopause (Weinstein and Schiff, 1983), and management of urethritis in men (Braun et al., 1982).

5. Choices under resource constraints

Increased interest in structured approaches to medical decision making has been due, in part, to the exigencies of resource constraints. It is not possible, in the presence of a seemingly limitless flow of new medical technologies, to deliver every technology to every patient who might benefit. Therefore, risky medical choices must be made under resource constraints, and the opportunity costs of resources spent must be taken into account. This new

environment — new, at least, for the industrialized world — has led to increased attention to the adaptation of the economic techniques of cost-effectiveness analysis and cost-benefit analysis to the resource allocation problems of the health sector.

There are some people, however, who deny the need to take resource constraints into account in medical decision making. They claim that, if certain steps are taken, it will be possible to deliver all health-improving technologies to all people within a reasonable health-care budget. This position, I contend, is based upon two misguided aspirations.

The first misguided aspiration is that, since most medical intervention is probably ineffective, all we need to do is learn which interventions work and which ones do not, in each possible group of patients. By eliminating all ineffective medicine, the supply of possible intervention would shrink to the point where the resource constraint would no longer be binding. This aspiration is misguided, for several reasons. Clinical efficacy studies are expensive, time-consuming, and often inconclusive. Even as they are underway, the technologies being evaluated are changing. In the end, there is almost always residual uncertainty about the effect of the technology, sometimes even more than at the outset. It is impossible, at a finite cost, to evaluate all technologies in all plausible subpopulations. Therefore, it is unrealistic to anticipate perfect information about effectiveness of medical interventions, although the marginal value of such information may very well be seriously underrated at present.

The second misguided aspiration of those who would like to avoid resource constraints is that hospitals and doctors can reduce costs, without affecting outcomes, by improving the efficiency of their delivery of services. Such efficiencies may very well ease the pressure in the short run, but in the end the inexorable forces of technology and patient expectations will once again push demands on resources up to the limit.

Hence, the relation between health resources and health benefits reflects diminishing marginal benefits from increased expenditures of resources. The resource constraint prohibits achieving the point of maximal health benefit. Therefore, optimal resource allocation is ensured by ranking medical practices according to their incremental cost-effectiveness ratio — incremental net resource cost per incremental health effectiveness. A health system that adopts technologies and health practices with lower values of the cost-effectiveness ratio in preference to those with higher values will have maximized health effectiveness subject to the resource constraint. This is the underlying principle of cost-effective resource allocation in health (Weinstein and Stason, 1977).

The numerator of the cost-effectiveness ratio for a technology — its net health resource cost — includes the following components: the direct cost of the procedure itself; *plus* costs of any procedures (test or treatments) induced by the procedure; *less* savings for procedures avoided as a result of the procedure; *plus* costs of diagnosing and treating side effects and complications of the procedure and any induced procedures; *less* savings for diagnosis and treatment of morbidity averted as a result of the procedure; *plus* the net change in the health care resource pool resulting from changes in life expectancy or altered productivity.

The denominator of the cost-effectiveness ratio — the net health effectiveness — may be measured on any of a variety of scales. These scales may range from narrowly specialized, such as the number of cases of breast carcinoma detected, to broadly general, such as the number of quality-adjusted years of life gained. Specialized measures avoid value judgments in weighing different attributes of health against one another, but are applicable only to those programs whose output is described by those particular measures. Quality-adjusted life

expectancy is now becoming a standard measure for evaluation, as reflected in the numbers of studies employing this measure (Boyle et al., 1983; Willems et al., 1980).

As an example of the kinds of comparisons it is now possible to make by using cost-effectiveness techniques, Figure 8 displays estimated cost-effectiveness ratios (net resource cost

Figure 8: Cost-effectiveness ratios (1984 U.S. dollars) for interventions in coronary heart disease.

- Coronary bypass surgery for LMD = left-main disease; 3VD = three-vessel disease; 2VD = two-vessel disease; 1VD = one-vessel disease; ▨ severe angina; ▧ mild angina; ▩ very mild angina.
- Treatment of post-MI adult men with beta-blockers.
- Admission to CCU compared to intermediate care unit (IntCU) in patients with a 20% (▩) or 5% (▨) risk of MI.
- Mobile CCU compared to previously existing CPR education and EMT Program.
- Treatment of sustained diastolic blood pressure (mm Hg) in specified range, or screening followed by treatment of all patients with diastolic blood pressure ≥ 95 mm Hg; ▩ full compliance; ▨ expected compliance.
- Treatment of adult men with serum cholesterol ≥ 265 mg/dL, using cholestyramine resin.
- Cholesterol screening of ten-year-old children, followed by a dietary intervention that lowers serum cholesterol by 10% in hypercholesterolemic children. Screening ▨ for all 10-year-olds; ▩ targeted by family history.

(Source: Weinstein and Stason, 1985.)

per quality-adjusted year of life gained) for a variety of technologies in the treatment or prevention of coronary heart disease (Weinstein and Stason, 1985). Comparative presentations of this sort may be useful in setting priorities for the use of new and existing technologies.

Another example, from Eddy's analysis of screening for cervical carcinoma, illustrates the importance of examining incremental cost-effectiveness ratios when considering mutually exclusive alternative technologies, or uses of the same technology (Eddy, 1980). These results (Table 2) suggest that annual screening with the Papanicolaou smear is justified only if it is considered appropriate to pay \$23,907 for each additional year of life expectancy above and beyond the benefits of screening every two years. It would be inappropriate to base this allocation of resources only upon the average cost per year of life saved by annual screening (\$1715), because that would ignore the fact that less expensive means to similar ends are available in the form of less frequent screening. It would also be incorrect to infer that screening every five years is even more "cost-effective", because that would ignore the availability of alternatives which, at the margin, are rather cheap ways to increase health benefits at a modest cost.

It is possible to broaden the scope of cost-effectiveness analysis to compare health programs that range outside the domain of clinical medicine (Table 3). Based, in part, on a review by Graham and Vaupel (1981), this comparison of cost-effectiveness ratios suggests that "heroic" medical interventions such as organ transplants and hemodialysis compare favorably to certain environmental and occupational health regulations of carcinogens but unfavorably to several automobile safety and air pollution regulations.

How would information on comparative cost-effectiveness of medical procedures actually be used to affect resource allocation decisions? The answer to this question depends on the health care system in which the question is asked, and who controls the flow of resources in that system. The preconditions for use of this information are that a budget exists, and that it is not possible to perform all potentially beneficial medical services.

In the United States, at least four parties could, potentially, use cost-effectiveness information to guide resource allocations.

First, insurers can formulate cost-effective coverage packages that maximize health impact subject to a premium level. Historically, insurance companies have been reluctant to determine benefits on any basis other than "medical necessity", but this state of affairs is changing as the competitive pressures to reduce premiums increase. The government, which acts as an insurance provider for the elderly and indigent in the United States, has similar incentives.

Second, physicians and hospital administrators can work together to determine cost-effective allocations of hospital resources. With prospective reimbursement of hospitals in the United States, incentives exist to give highest priority to the most cost-effective services. Coronary care units, new pharmaceuticals such as cephalosporin antibiotics, and new diagnostic technologies such as magnetic resonance imaging are among the technologies for which guidelines for priority use are likely to be based, in part, on cost-effectiveness considerations.

Health maintenance organizations have a natural incentive to allocate a fixed budget efficiently. And, ultimately, individual patients face decisions about how to allocate their own health care dollars.

*Table 2: Benefits, costs, and cost-effectiveness ratios of screening for cervical cancer using the pap smear**

Interval Between Pap Smears (Years)	Increase in Life Expectancy (Days)	Program Cost (\$)	Cost-Effectiveness Ratio (\$/Year)	Incremental Cost-Effectiveness Ratio (\$/Year)
5	53.99	54.54	368.72	–
4	58.00	69.65	438.31	1375.35
3	61.56	95.61	566.89	2661.63
2	64.55	149.39	844.73	6565.12
1	67.08	315.10	1714.54	23906.78

* Source: Eddy (1980).

*Table 3: Cost per year of life saved for some health investments (Graham and Vaupel, 1981)**

* Mandatory Air Bags	\$ 540
* Mandatory Smoke Detectors (Bedrooms)	1300
* Air Pollution Controls – Stationary Sources	3800
* Motor Vehicle Safety Act (1966)	6300
Coronary Artery Bypass (Three Vessel)	7200
* 55 MPH Speed Limit	12000
Hemodialysis	25000
* Air Pollution Controls – Mobile Sources	105000
Aflatoxin – 10 ppb Tolerance	140000
* Water Carcinogens – 100 mcl	240000
Liver Transplant – 1 Year Survival	250000
* Vinyl Chloride – 1 ppm occupational standard	490000
* Acrylonitrile – 1 ppm occupational standard	1900000
* Benzene – 1 ppm occupational standard	6600000

The institutions that control health care resources vary from one nation to another. However, all nations have imposed some mechanism that ensures that resource constraints are felt at some level in the system, in contrast to the open-ended system of full insurance and fee-for-service payment that existed in the United States until recently.

6. Limits and challenges for medical decision making

Despite the progress in theory and application of clinical decision analysis during the past fifteen years, formal analysis of decision making under uncertainty has not yet been accepted by medical practitioners and health care administrators to the degree that it has been adopted by business managers. My concluding remarks will address some possible reasons for this limited acceptance and some challenges for the future. These challenges that I will address correspond to the three elements of decision analysis that I have discussed: probabilities, utilities, and resource constraints. Specifically, they stem from the limited quality of information about the effects of clinical procedures in heterogeneous populations; incomplete understanding of the value and preferences of patients and potential users of health care; and imperfectly structured incentives for cost-effective resource use in both private and public health care systems.

Clinicians become frustrated at the task of assigning probabilities explicitly, especially when the empirical data base is deficient. A methodology that exposes one's ignorance is not easily accepted. Three developments in medical decision research, however, offer new promise in this regard.

The first is a set of techniques whose common objective is to synthesize available data into estimates of desired probabilities. Statistical techniques of multiple regression analysis now make it possible for clinicians to obtain probability estimates as functions of a set of clinical variables, thereby providing individualized probabilities rather than relying on aggregate estimates. Methods known under the rubric of meta-analysis (Louis et al., 1985) combine information from diverse sources into point estimates and standard errors of estimates, based on models of variance within and between sources. Implementation of all of these methods is facilitated by the appearance of computer data bases for a variety of diseases and treatments.

A second development that will facilitate probability assessment in medicine is the increasing use of structured judgment techniques to elicit the judgments of experts. Growing out of a research tradition that has spawned such methods as the Delphi method, these methods have been applied in formal decision-analytic studies and could, potentially, be incorporated into the transmission of clinical information through such vehicles as consensus development panels and publications of medical professional organizations.

A third development is the increased emphasis on quantitative knowledge in medical education. Courses on probability and statistics are widespread, and many medical students are exposed to at least an introduction to decision analysis. Perhaps future generations of physicians will not only accept, but demand quantitative information about their practices.

The challenges and mysteries regarding values and preferences are equally, if not more, formidable. Evidence from clinical studies confirm the findings, from psychological experiments, that patients' preferences do not seem to fit simple models suggested by expected utility theory. Problems of sensitivity to framing of information, labile and unstable preferences,

apparently excessive valuation of diagnostic information, and concern for psychological consequences of decisions have all been found. Inconsistencies in the stated and revealed marginal willingness to pay for lifesaving interventions also raise questions about the meaning of preferences and the desire of decision makers to adhere to basic principles of "rationality". Solutions to these so-called paradoxes may lie, in part, in enriching the range of attributes of outcome that are valued by medical decision makers, including patients. More radical solutions may involve rejecting the expected utility model as a prescriptive basis for action, given that it does not seem to perform well descriptively.

Framing effects in medical decision making are best illustrated by an experiment by McNeil et al. (1982), in which subjects were asked to imagine that they had lung cancer, and were asked to choose between two sets of probabilities of death or survival, corresponding to surgery and radiation therapy. Half the subjects were given probabilities of *death*, and the other half were given equivalent probabilities of *survival* (i.e. one minus the probabilities of death). Thus, the substantive choices presented to the two groups were identical, but they were framed differently, for example, as a 10 percent chance of death by 1 year or as a 90 percent chance of survival to 1 year. These investigators found that 44 percent of the subjects who received the data as probabilities of *death* preferred the prognosis with radiation therapy. However, only 18 percent of the subjects in the *survival* frame preferred radiation therapy. (One hypothesis was that subjects gave more attention to a difference between 0 percent and 10 percent immediate mortality than to a difference between 100 percent and 90 percent immediate survival, a behavior predicted by prospect theory (Kahneman and Tversky, 1979).) If preferences are so sensitive to the way in which information is presented (a fact well known to the advertising industry), then what does it mean to speak of a patient's "true" preferences?

Another source of mystery about preferences comes from studies that suggest that preferences are labile, that they change with time and with circumstances. In one study, women entering labor and delivery were asked to rate their aversion to the adverse effects of labor and of anesthesia (Christensen-Szalanski, 1984). From one month prior to labor through early labor, subjects strongly considered anesthesia worse than labor pains. From active labor to delivery, these preferences were reversed: anesthesia was regarded as the lesser of the two evils. Post-partum, preferences returned toward the pre-labor levels, through somewhat attenuated (presumably by the experience of labor). What is the physician to do if a patient, in labor, cries out for anesthesia, but the physician knows that the patient would regret it later? Which of these "multiple selves" represents the real patient? Can one model this kind of behavior in a rational utility model? The author of this study conjectures a descriptive model based on discounting at a non-constant rate, backward in time from a critical event, but this model stops short of prescribing a method for aggregating across preferences that shift over time.

A third finding about patient preferences would seem to challenge the decision-theoretic model of the value of diagnostic information. Decision theory says that the value of test information lies in its ability to affect subsequent decisions and, thereby, improve expected outcomes. In clinical applications of this principle, "test information" is usually defined to be "clinical information", and subsequent decisions are usually restricted to clinical decisions made by the physician (Pauker and Kassirer, 1980; Eeckhoudt et al., 1984). In a survey of patients' willingness to pay for the components of information contained in an ultrasound examination during apparently normal pregnancies (Berwick and Weinstein, 1985) we found that, of the total willingness to pay, 40 percent would be paid for the information even if the physician were unable to learn the results and, therefore, could not alter decisions accordingly.

Furthermore, 29 percent of the value was unrelated to the ability of either the physician or the patient to alter subsequent actions on the basis of the information — this value lay in “just knowing” whether the fetus was healthy, or due within three months, or not twins. Overall, only 52 percent of the stated value of information related to medical-decisional consequences of the information; the rest related to non-medical decisions, or just having the knowledge. While the decision-theoretic framework is surely robust enough to incorporate non-medical decisions and psychological peace of mind into an analysis, applications typically ignore these.

A fourth set of “paradoxes” in health-related preferences suggest that people attach greater value to health risk reductions that eliminate or avoid small risks that are attributable to an identifiable cause than to those that reduce greater statistical risks. The “certainty effect” and “pseudocertainty effect” of Tversky and Kahneman (1981) are among several plausible descriptive constructs for these phenomena. Bell (1982), and Loomes and Sugden (1982) have independently developed theories of regret, which may be viewed as prescriptive models (that is, refinements of utility theory) in which an outcome called “decision regret” is explicitly valued by the decision maker. As pointed out by Brett (1982), physicians may prefer to avoid actions that may result in adverse outcomes attributable to those actions (such as operative mortality), even if the expected outcome for the patient is worse. In Brett’s example, a patient has a 20 percent chance of having disease X, which carries a 60 percent chance of imminent mortality without surgery. Surgery cures disease X, but carries a 10 percent risk of operative mortality for all patients. Thus, the patient has a 12 percent chance of death without surgery and a 10 percent chance of death with surgery. But, of every 10 deaths with surgery, 8 (or 80 percent) are caused by the surgery, thus creating the opportunity for post-decisional regret. Death by act of God may carry less regret than death under the knife. The question would remain, however, whether the *physician’s* regret ought to play a role in determining the care of patients. Weinstein and Quinn (1983) have proposed a model of post-decisional “blame” which generalizes the regret model to cases where the post-decisional attribution of cause is probabilistic rather than deterministic.

This challenge, then, is to understand more about what patients, and potential patients, value about health care. It is clearly naive to assume that patients wish to maximize quality-adjusted life expectancy. But what do they wish to maximize? The more we learn about this question, the more acceptable will be the prescriptive models that seek to guide allocations of medical resources.

The final challenge is related to the incentives facing health-care decision makers to allocate resources. Every health-care system needs to consider how to transfer the societal objective of maximizing health-related utility within global resource limits into a system of decentralized incentives for individual providers and institutions. With such incentives in place, decision analysis and cost-effectiveness analysis will become important management tools at all levels of health care.

REFERENCES

- BARZA, M., PAUKER, S.G.: "The decision to biopsy, treat, or wait in suspected herpes encephalitis". *Ann. Intern. Med.*, 1980, 92:641-649.
- BELL, D.E.: "Regret in decision making under uncertainty". *Opns. Res.*, 1982, 30:961-981.
- BERWICK, D.M., WEINSTEIN, M.C.: "What do patients value? Willingness to pay for ultrasound in normal pregnancy". *Med. Care*, 1985, 23:881-893.
- BOYLE, M.H., TORRANCE, G.W., SINCLAIR, J.C., HORWOOD, S.P.: "Economic evaluation of neonatal intensive care for very-low-birth-weight infants". *N. Engl. J. Med.*, 1983, 308:1330-1337.
- BRAUN, P.: "The clinical management of suspected Herpes virus encephalitis: A decision-analytic view". *Am J. Med.*, 1980, 895-902.
- BRAUN, P., SHERMAN, H., KOMAROFF, A.L.: "Urethritis in men: Benefits, risks, and costs of alternative strategies of management. *Sexually Transmitted Diseases*", 1982, 9:188-199.
- BRETT, A.S.: "Hidden ethical issues in clinical decision analysis". *N. Engl. J. Med.*, 1982, 305:1150-1152.
- CHRISTENSEN-SZALANSKI, J.J.J.: "Discount functions and the measurement of patients' values: Women's decisions during childbirth". *Med. Decis. Making*, 1984, 4:47-58.
- EDDY, D.M.: *Screening for Cancer: Theory, Analysis, and Design*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.
- ECKHOUDT, L.R., LEBRUN, T.C., SAILLY, J.-Cl.: "The informative content of diagnostic tests: An economic analysis. *Soc. Sci. Med.*, 1984, 18:873-880.
- GRAHAM, J.D., VAUPEL, J.W.: "Value of a life: What difference does it make?", *Risk Analysis*, 1981, 1:89-95.
- KAHNEMAN, D., TVERSKY, A.: "Prospect theory: An analysis of decision under risk". *Econometrica*, 1979, 47:263-291.
- LOOMES, G., SUGDEN, R.: "Regret theory: An alternative theory of rational choice under uncertainty". *Econ. J.*, 1982, 92:805-824.
- LOUIS, T.A., FINEBERG, H.V., MOSTELLER, F.: "Findings for public health from meta-analyses". *Annu. Rev. Public Health*, 1985, 6:1-20.
- McNEIL B.J., PAUKER S.G., SOX, H.C. Jr., TVERSKY, A.: "On the elicitation of preferences for alternative therapies". *N. Engl. J. Med.*, 1982, 306:1259-1262.
- McNEIL, B.J., WEICHSSELBAUM, R., PAUKER, S.G.: "Fallacy of the five-year survival in lung cancer". *N. Engl. J. Med.*, 1978, 299:1397-1401.
- VON NEUMANN, J., MORGENSTERN, O.: *Theory of Games and Economic Behavior*, Second edition. Princeton, N.J.: Princeton University Press, 1947.
- PAUKER, S.G.: "Coronary artery surgery: The use of decision analysis". *Ann. Intern. Med.*, 1976, 85:8-18.
- PAUKER, S.G., KASSIRER, J.P.: "The threshold approach to clinical decision making. *N. Engl. J. Med.*, 1980, 302:1109-1116.
- PAUKER, S.P., PAUKER, S.G.: "Prenatal diagnosis: A directive approach to genetic counseling using decision analysis. *Yale J. Biol. Med.*, 1977, 50:275-289.
- PLISKIN, J.S., SHEPARD D.S., WEINSTEIN, M.C.: "Utility functions for life years and health status". *Opns. Res.*, 1980, 28:206-224.
- RAIFFA, H.: *Decision Analysis: Introductory Lectures on Choice Under Uncertainty*. Reading MA: Addison-Wesley, 1968.
- RANSOHOFF, D.F., GRACIE, W.A., WOLFENSON, L.B., NEUHAUSER, D.: "Prophylactic cholecystectomy or expectant management of silent gallstones". *Ann. Intern. Med.*, 1983, 99:199-203.

- READ, J.L., PASS, T.M., KOMAROFF, A.L.: "Diagnosis and treatment of dyspepsia: A cost-effectiveness analysis". *Med. Decis. Making*, 1982, 2:415-438.
- RICHTER, J.M., SILVERSTEIN, M.D., SCHAPIRO, R.: "Suspected obstructive jaundice: A decision analysis of diagnostic strategies". *Ann. Intern. Med.*, 1983, 99:46.
- SCHOEMAKER, P.J.H.: "The expected utility model: Its variants, purposes, evidence and limitations". *J. Econ. Lit.*, 1982, 20:529-563.
- SIMES, R.J.: "Treatment selection for cancer patients: Application of statistical decision theory to the treatment of advanced ovarian cancer". *J. Chron. Dis.*, 1985, 38:171-186.
- SLOVIC, P., FISCHHOFF, B., LICHTENSTEIN, S.: "Behavioral decision theory". *Ann. Rev. Psych.*, 1977, 28:1-39.
- TORRANCE, G.W., BOYLE, M.H., HORWOOD, S.P.: "Application of multi-attribute utility theory to measure social preferences for health states". *Opns. Res.*, 1982, 30:1043-1069.
- TVERSKY, A., KAHNEMAN, D.: "Judgment under uncertainty: heuristics and biases". *Science*, 1974, 85:1124-1131.
- TVERSKY, A., KAHNEMAN, D.: "The framing of decisions and the psychology of choice". *Science*, 1981, 211:453-458.
- WEINSTEIN, M.C., FINEBERG, H.V., ELSTEIN, A.S., FRAZIER, H.S., NEUHAUSER, D., NEUTRA, R.R., McNEIL, B.J.: *Clinical Decision Analysis*. Philadelphia: W.B. Saunders, 1980.
- WEINSTEIN, M.C., QUINN, R.J.: "Psychological considerations in valuing health risk reductions". *Natural Resources J.*, 1983, 23:659-673.
- WEINSTEIN, M.C., STASON, W.B.: "Cost-effectiveness of coronary artery bypass surgery". *Circulation*, 1981, 66 (Suppl. III):III-56-III-66.
- WEINSTEIN, M.C., STASON W.B.: "Cost-effectiveness of interventions to prevent or treat coronary heart disease". *Annu. Rev. Public Health*, 1985, 6:41-63.
- WEINSTEIN, M.C., STASON, W.B.: "Foundations of cost-effectiveness analysis for health and medical practices". *N. Engl. J. Med.*, 1977, 296:716-721.
- WILLEMS, J.S., SANDERS, C.R., RIDDIOUGH, M.A., BELL, J.C.: "Cost-effectiveness of vaccination against pneumococcal pneumonia". *N. Engl. J. Med.*, 1980, 303:553-559.