Milbank Memorial Fund Quarterly/*Health and Society, Vol. 56, No. 3, 1978*

Decision Theory in Medicine

A Review and Critique

DANIEL A. ALBERT

Department of Internal Medicine,
The University of North Carolina at Chapel Hill

DECISION THEORY IS A GROUP OF RELATED CONSTRUCTS that seek to describe or prescribe how individuals or groups of people choose a course of action when faced with several alternatives and a variable amount of knowledge about the determinants of the outcomes of those alternatives. As such, it has been divided into theories concerned with either individual or group decisions and either descriptive (how people *do* behave) or prescriptive (how people *should* behave) decisions (Suppes, 1967):

	Individual	*Group*
Prescriptive (or Normative)	Classical Economics (Statistical) Decision Theory Moral Philosophy	Game Theory Welfare Economics (including Cost-Benefit Analysis) Political Theory
Descriptive	Experimental Decision Studies Learning Theory Surveys of Voter Behavior	Social Psychology Political Science

Although many health planning decisions are group normative decisions (especially cost-benefit analysis), applications to clinical medicine have been primarily in the area of individual prescriptive decision analysis—that is, how an individual faced with alternatives

002-3745-78-5603-0362-39/$01.00/0

should make a choice (Gorry, Kassirer, Essig, et al., 1973; Lusted, 1971; Pauker and Kassirer, 1975; 229; Schwartz, Gorry, Kassirer, et al., 1973; Sisson, Schoomacker, and Ross, 1976).

Another categorization of decision theory is by the amount of knowledge the decision maker has about the determinants of various outcomes for a given action. Thus, there are: 1) decisions of certainty, where each alternative course of action has a single well-specified outcome; 2) decisions of risk or uncertainty, where each alternative course of action has a well-defined set of possible outcomes each with a probability of occurrence; and 3) decisions of ignorance, where each action results in a range of possible outcomes but the probability of occurrence of each outcome is unknown (Luce and Raiffa, 1957).

Examples of individual normative decisions *under certainty* include linear programming problems where there is a set of possible acts that will maximize a given index while satisfying a limiting condition (feasibility criterion). An illustration of this type of problem is that of choosing a diet that satisfies minimum daily requirements and minimizes cost. Routing and scheduling problems often fall into this category and have great administrative importance. This class of problems requires: 1) acts, each of which is associated with a number; 2) a feasibility criterion or constraints on the acts; and 3) an index associated with each act that reflects a property of the act, *i.e.*, cost. The problem is usually to minimize or maximize item 3 above while satisfying item 2.

Individual decision-making *under risk* involves the analysis of a gamble. The decision involves a choice between members of a set of acts whose outcomes depend on chance—the "state of nature." The expected outcome value for an action is determined by the probability of its occurrence and its value to the decision maker. This value, attached to an outcome, is a measure of the decision maker's preference for that outcome. Thus, each action determines a set of possible outcomes. Each expected outcome is the product of the probability of occurrence times the utility of the outcome. The solution to the problem is to find which action's total expected outcome value is maximal. Examples of decisions of risk include whether to carry an umbrella given a probability of rain; the choice of an investment given known risks and assumed payoffs; and the choice of therapies given probabilities of diseases and known outcomes of those therapies (Winkler, 1972).

Choices between actions with a known set of possible outcomes but no assignment of probabilities constitute decisions *under ignorance.* To illustrate from our previous examples, the choice of whether to take an umbrella without a forecast, the choice of an investment with no idea of the risks, or the choice of therapy without any idea of the probability of the diseases in question all constitute decisions under ignorance.

In most medical problems, the choice of actions, the range of outcomes, and some estimation (more or less precise) of the probabilities are available. Thus, medical decision problems are usually in the class of decision problems under risk.

Historical Perspective on Decision Theory

The historical roots of decision theory are difficult to trace, and probably some originate in antiquity (Savage, 1972: 91). Recognizable traces of decision theory concepts are found in medieval ethics, as, for example, the quote attributed to the Port Royal School by Keynes (1956: 1360): "In order to judge of what we ought to do in order to obtain a good and to avoid an evil, it is necessary to consider not only the good and evil in themselves, but also the probability of their happening and not happening, and to regard geometrically the proportion which all these things have taken together." This statement expresses the fundamental components of a decision—that each alternative must be examined with regard to its value and the probability of its occurrence. Variants of this proposition were expressed by Leibnitz, Locke, and Butler (Suppes, 1967).

Other mathematical roots arose as probability theory evolved from the gaming tables of Monte Carlo to a real-world description. Two concepts generated during this evolution were crucial to the creation of decision theory. First, the notion of probability was extended from the narrow notion of the frequency of independent repetitive events, such as rolls of the dice, to a measure of the confidence of the truth of a particular proposition, such as the probability of the occurrence of rain. The second crucial notion was the dissociation of price or monetary value of an object from the value of utility of that object for an individual. Thus, although $100 has a fixed monetary value, its utility or value is quite different for a rich and a poor man. These concepts were refined and advanced by Bernoulli, Keynes, Ramsey and DeFinetti (Suppes, 1967).

In 1944, the concepts of value and probability were fused into a unified theory of decision making by Von Neumann and Morgenstern, and published in 1947 in their historic work, *Theory of Games and Economic Behavior,* which laid a mathematical foundation for decision theory. Since then, the field has burgeoned with multiple applications to policy making, sociology, economics, and psychology. Since Ledley and Lusted's initial applications (1959) to clinical problems there has been a flurry of interest in medical decision making.

Formulation of the Problem

Decision problems under risk are generally formulated in either a matrix or a tree diagram. A decision matrix enumerates the possible actions on the ordinate and the states of nature (*i.e.*, the situational determinants of outcomes for a given action) on the abscissa. The outcomes are assigned to the intersection lines and are labeled by the coordinates. Each outcome has a value or utility assignment and a probability of becoming existent. The probability assignments are the same for each outcome of a given state of nature, that is, all outcomes in a row have the same probability assignment. Diagrammatically, this may be expressed as:

$$
\begin{array}{lcccc}
 & S_1/P_1 & S_2/P_2 & S_3/P_3 \;\ldots & S_n/P_n \\
A_1 & O_{1,1}/P_1, U_{1,1} & O_{1,2}/P_2, U_{1,2} & O_{1,3}/P_3, U_{1,3} & O_{1,n}/P_n, U_{1,n} \\
A_2 & & & & \\
A_3 & & & & \\
\cdot & & & & \\
\cdot & & & & \\
\cdot & & & & \\
A_m & & & & O_{m,n}/P_n, U_{m,n}
\end{array}
$$

where $A_{1\to m}$ is the set of actions, $S_{1\to n}$ is the set of states of nature, and $O_{1\to m, 1\to n}$ is the set of outcomes, each with an associated probability and utility value.[1] For example, in the umbrella problem

[1]In this paper the notation $X_{1\to k}$ is used to denote the set $\{X_1, X_2, \ldots X_k\}$.

the matrix would appear:

	rain	no rain
A_1 take umbrella	stay dry	useless umbrella
A_2 no umbrella	get wet	walk without bother

Each outcome (stay dry, get wet, etc.) would be assigned a value by the decision maker and a probability that would be uniform throughout each column.

The expected utility of each action, $A_{1\to m}$, is the sum of the products of the probability and utility assignments for all possible outcomes of that action. For example, for action A_1 the expected utility is:

$$U_{1,1} \times P_1 + U_{1,2} \times P_2 + U_{1,3} \times P_3 + \ldots U_{1,n} \times P_n.$$

The solution to the decision problem is to find which action's expected utility is maximal. In the umbrella question, let us say that the probability of rain is 50%, and the values have been assigned as follows: stay dry = 1.0, get wet = 0, useless umbrella = 0.9, and walk without bother = 1.0. In this case, the solution to the problem would be:

	rain 0.5	no rain 0.5
take umbrella	stay dry 1.0	useless umbrella 0.9
no umbrella	get wet 0.0	walk without bother 1.0

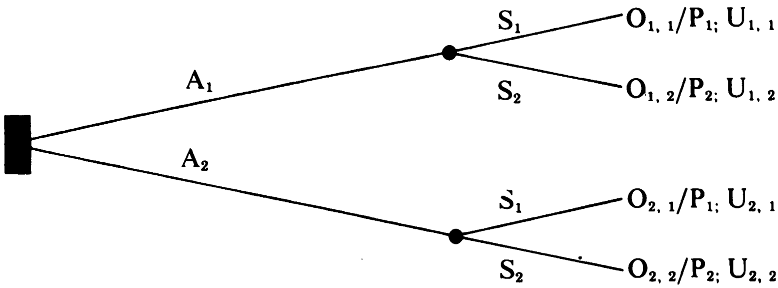
The equations may be expressed as:

take umbrella = $1.0 \times 0.5 + 0.9 \times 0.5 = 0.95$;

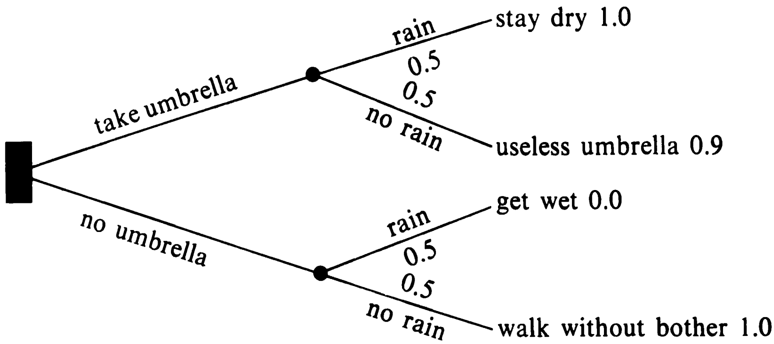
no umbrella = $0.0 \times 0.5 + 1.0 \times 0.5 = 0.50$.

Thus, take umbrella would be the preferred action.

Another useful representation of the problem is a "tree" diagram. The initial branches of the tree represent the possible actions, and the subsequent branches lead to the possible outcomes dependent on the various states of nature. Willful decisions are identified by squares at the branch points, and chance nodes are denoted by circles. For a decision problem involving two possible actions and two states of nature, the tree would appear:



In the umbrella problem the tree diagram would appear:



For multiple possible actions and states of nature, the tree becomes very "bushy" and unwieldy. Pruning of the tree is an entirely judgmental process that involves eliminating branches of either low probability or low utility or both (Raiffa, 1968). The tree diagram is very convenient for problems that involve several separate decisions over time such that the existence or character of one decision is dependent on the outcome of the previous decision.

Components

Decision theory rests on five main conceptual components: states of nature, actions, outcomes, probability, and utility functions. We will briefly explore each concept individually.

States of Nature

"State of nature" is the nomenclature for the space-time determinants of an outcome for a given action. As Luce and Raiffa state (1957): "the set of 'states of nature' is assumed to form a mutually exclusive and exhaustive listing of those aspects of nature which are relevant to this particular choice problem and about which the decision maker is uncertain." Two states of the world could be very different, but, unless they generate or are related to two different outcomes for a given action, they will be labeled the same state of nature. This viewpoint is convenient for medical problems where actions are often choices of intervention and outcomes can be thought of as the result of interaction between the intervention and the state of the patient. Fortunately, one need only be able to specify a set of outcomes for each action and label them in a one-to-one fashion with hypothesized states of nature. It is not necessary to express explicitly the characteristics of the state of nature that are causally related to the generation of a given outcome from an action. In our umbrella problem, the states of nature are rain or no rain.

Actions and Outcomes

Actions, as we indicated in the previous section, are those options available to the decision maker that generate outcomes. In the umbrella question the actions were to take the umbrella or not, and the outcomes were staying dry, getting wet, taking a useless umbrella, and walking without the bother of an umbrella. The scheme is said to be well specified in that it does not permit unstated actions, unknown outcomes, or outcomes unrelated to specific actions. Both sets of possible actions and outcomes are finite, complete, and invariant for a given decision problem. Even these modest completeness requirements are beyond specification for many medical problems.

Probabilities

Although there is little explicit restriction on the probabilities assigned to the various states of nature, there are several implicit restrictions. First, the accuracy of the solution depends directly on the probability assignments. Second, the sum of the probabilities of all the outcomes for a given action must be unity. Third, the probability assignments reflect the confidence of the decision maker in the likelihood of outcomes (Good, 1959).

The accuracy of probability assignments is a major stumbling block of both medical and non-medical decision-making (Tversky and Kahneman, 1974: 1124). On the non-medical side, it has been shown in gambling situations that subjects consistently overestimate low probabilities and underestimate high probabilities. Furthermore, there appears to be some coloring of the probability estimates by the size of the stakes involved. When the stakes were of no value to the decision maker his probability estimates were improved, but this is inapplicable to most medical decision problems. The fact that people are not very good at estimating probabilities is one of the reasons for the financial success of gambling houses (Edwards, 1954; 1961).

Medical decision analyses have dealt with the problem of accuracy of probability assignments in two fashions. Some studies use available data from clinical studies to calculate probabilities (McNeil and Adelstein, 1975; Sisson, Schoomacker, and Ross, 1976). Here the inaccuracy is the abstraction from the frequency of a finding in a large population, to the likelihood of an event in an individual who may or may not differ from the characteristics of the norm of the population sampled. The other technique is the opinion of experts in the field (Gorry, Kassirer, Essig et al., 1973; Pauker and Kassirer, 1975; Schwartz, Gorry, Kassirer et al., 1973; Thornbury and Fryback, 1977). Some studies suggest considerable variability and inaccuracy among clinicians as predictors. However, investigation of the predictive ability of experienced clinicians suggests that their overall error rate may be quite low (King and Manegold, 1965; Shapiro, 1977). One technique for dealing with the variability in probability assignment is called "sensitivity analysis" (Pauker and Kassirer, 1975; Winkler, 1972). It involves calculating the expected utility of various outcomes using the upper and lower limits of the probability range. If the decision choice is preserved at both ends of the range of probabilities, then the analysis is not thought to depend on the specific probability assignments as long as they are within this range.

The sum of the probabilities of all outcomes of a given action must total 1.0. However, experimental subjects often overshoot or undershoot this mark when asked about multiple probabilities of mutually exclusive and exhaustive situations (Edwards, 1961; Kyburg and Smokler, 1964).

The third random requirement—that the probabilities represent the decision maker's confidence in the likelihood of an outcome—is opposed to the view that probabilities stand for the frequency of outcomes of random independent events. Because we are able to talk reasonably and intelligibly about the likelihood of non-random non-repetitive events (such as a weather forecast) in much the same way as we speak of rolls of dice, a subjective view, that probability represents the degree of belief in a statement, has become popular (Kyburg and Smokler, 1964; Menges, 1970). Thus, we can assign probabilities to outcomes in non-repetitive medical decisions where the availability of empirical evidence related to the actual frequency of outcomes may be lacking. This subjectivist viewpoint is very controversial (Mainland, 1967).

Bayesian Inference. In a medical decision under risk, the probabilities of the various diagnoses will change with the acquisition of new data. Thus, each test result forms a new branch of a decision tree, altering the probability of each prospective diagnosis. For example, in a work-up for tuberculosis (TB), a positive tuberculin test certainly alters the probability of active TB. This change in probability achieved by new information is calculated as follows using Bayes theorem (Hall, 1967):

$$P(D/S) = \frac{P(S/D)\,P(D)}{P(S/D)\,P(D) + P(S/\bar{D})\,P(\bar{D})},$$

where D is the disease,
S is the symptom (sign or laboratory test),
$P(D/S)$ is the probability of D given S (*a posteriori* probability),
$P(S/D)$ is the probability of S given D (known as the likelihood of S given D),
$P(D)$ is the probability of D (the *a priori* probability or prevalence of D),
$P(\bar{D})$ is the probability of not D, *i.e.*, the probability of the alternative diagnosis,
$P(S/\bar{D})$ is the probability of S given the alternative diagnosis (not D).

For example, let us say that from the history and physical exam we feel a patient has a 0.5 chance of having active TB. In addition, we estimate that the probability of a positive tuberculin test in active TB is 0.95 and in the absence of active TB is 0.8. The probability of active TB in our patient if he has a positive tuberculin test would be:

$$\frac{0.95 \times 0.5}{0.95 \times 0.5 + 0.8 \times 0.5}, \text{ or } 0.54 \text{ or } 54\%.$$

Therefore, if the probability of active TB before the test was 0.5, the probability after a positive tuberculin test (given all the other estimations) would be only 0.54.

The equation may be visualized by a Venn diagram (Fig. 1). A lucid explanation and proof of the Bayes theorem are contained in Hall (Hall, 1967: 555). A commonplace medical use of Bayes theorem is in genetic counseling where the probability that progeny will be affected is calculated using Bayesian analysis from data about the family (McKusick, 1964: 151; Murphy and Mutalik, 1969; Pauker and Pauker, 1977). The implications and ramifications of Bayesian inference in medical situations have been investigated extensively (Cornfield, 1964: 163; Elstein, 1976; Feinstein, 1977; Gustafson, Kestly, Griest et al., 1976; McNeil and Adelstein, 1975; Woodbury, 1963).

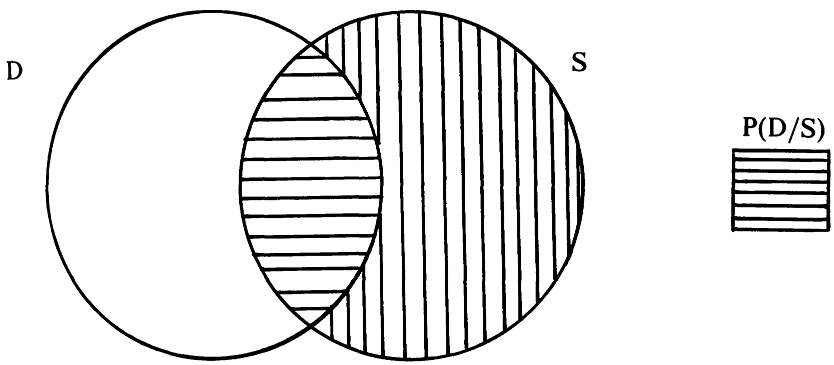


FIG. 1. Venn diagram illustrating the probability of active TB occurring in a patient based on the Bayes formula. The open circle represents the disease, D; the vertically ruled circle represents the symptoms, S; and the overlapping portion with the horizontal rules represents the probability of the disease occurring in the presence of symptoms, P(D/S).

Automated Diagnosis. Numerous mechanical, semi-mechanical, and automated systems of diagnosis have been proposed covering a wide range of disease categories, including thyroid disease, congenital heart disease, and major neurological diseases (Croft, 1972; Gorry and Barnett 1968; Overall and Williams, 1963; Safran, Tsichlis, Bluming et al., 1977; Warner, Toronto, Veasey et al., 1961). Although different mathematical and logical models underlie these attempts, there are certain unifying features (Gorry, 1970: 293;

Kayser, 1975: 76). These include the selection of the: 1) set of signs, symptoms, and lab tests to be used to discriminate among disease categories; 2) disease categories to be discriminated among; and 3) set of rules that assign a patient's data set to the disease classes. This process may be visualized as follows:

$$
\begin{array}{llcccc}
 & & & & \text{Diseases} & \\
 & & D_1 & D_2 & D_3 \; . \; . \; . & D_n \\
\text{Possible} & SC_1 & P_{1,1} & P_{1,2} & P_{1,3} & P_{1,n} \\
\text{Symptom} & SC_2 & & & & \\
\text{Complex} & SC_3 & & & & \\
 & \vdots & & & & \\
 & SC_m & & & & P_{m,n}
\end{array}
$$

where SC is a symptom complex composed of possible positive and negative clinical data. Thus, for each possible symptom complex, every disease considered has a probability. The decision rule is, for a given symptom complex (*i.e.*, for a given patient's symptoms), to choose the disease $D_{1 \to n}$ with the highest probability.

Another means of visualizing this process with those models that use a sequential Bayesian inference scheme is given in Fig. 2, which describes one patient's combination of three positive symptoms (s_1, s_3, s_4) and two negative symptoms (s_2, s_5). The probability of each disease being considered would be recalculated with the addition of a new value for each symptom, $S_{1 \to n}$. For each patient with 30 pieces of clinical data ($s_{1 \to 30}$) and 20 possible diseases ($D_{1 \to 20}$), there would be approximately 600 calculations. As the disease categories and patient load increase, this increased number of calculations can become unwieldy for even a modern computer (Croft, 1972).

Those automated diagnostic programs using Bayesian inference generate probabilities of each disease entity for each patient. The above formulation of probabilistic automated diagnostic devices clarifies their position within decision theory. They are individual normative decision devices under risk or uncertainty. That is, they generate a probability of a diagnosis for each patient given certain signs and symptoms. Since all diagnoses are pursued equally and ranked strictly according to their probability, we must infer that they are given the same value or utility. This equal ranking of diagnoses is

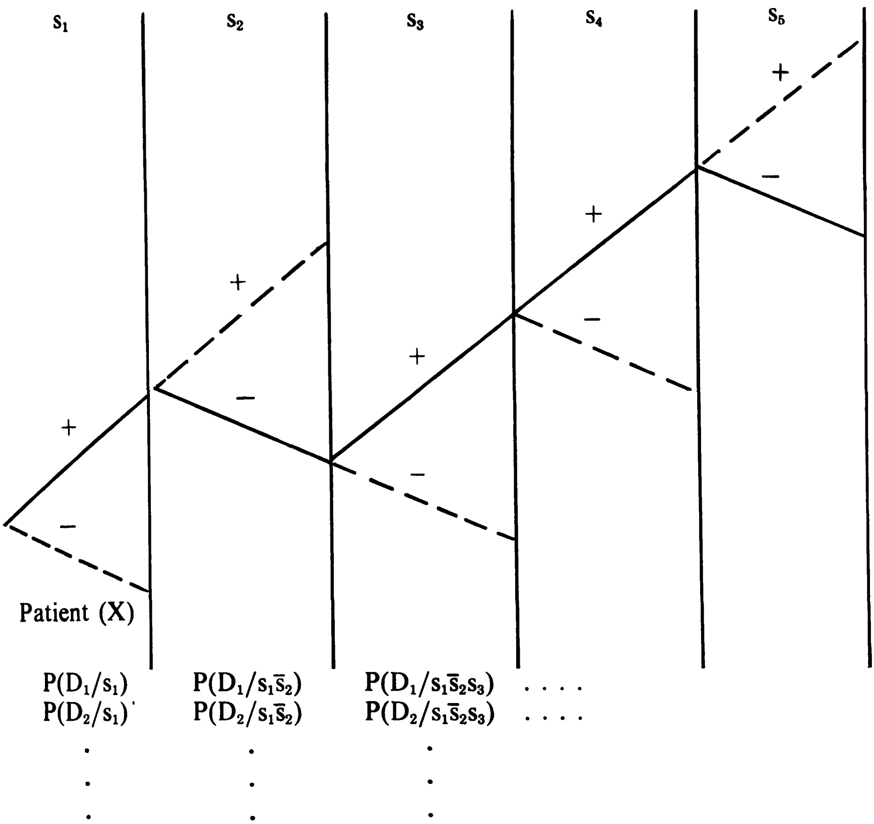


FIG. 2. Diagram illustrating the sequential steps necessary to calculate the probability of a disease occurring in a patient who has three positive symptoms for the disease (s_1, s_3, s_4) and two negative symptoms (s_2, s_5), using Bayesian inference.

neither necessary nor realistic—some diagnoses are of little importance to the patient and are unaffected by medical intervention, while others are of great importance. This distinction can only be generated by some measure of utility or value for a diagnosis. However, automated diagnostic programs often stop at the point of choosing a most likely diagnosis (Croft and Machol, 1974). Other problems in automated diagnosis include: 1) change in symptoms during the course of the disease; 2) change in *a priori* probabilities with seasonal or epidemic conditions; and 3) assumption of statistical independence of each datum (the information of each datum is not contained within any other data) which will consistently tend to overestimate the end probability when Bayes theorem is used

(Anderson and Boyle, 1968; Fryback, 1974). Programs utilizing different mathematical schemes instead of Bayesian inference have related problems. The progress in automated diagnosis has been recently reviewed (Ross, 1972).

Utility Theory

Utility (or value) assignments and determinations are the cornerstone of decision analysis (Lindley, 1975). Although there is an intuitive notion of the value of an outcome, the utility assignments must be quantitated in order to manipulate the variables in such a way as to generate comprehendible solutions to a decision problem. A linear scale of utilities is achieved by confronting the decision maker with gambles between the alternatives. Let us say that an individual is asked to express his preferences between three objects, A, B, and C, and at the outset he is able to state his preference of A to B, B to C, and A to C. To quantitate his utility scale he is offered a coin flip lottery between B for certain versus a gamble between A and C. He is asked what are the odds of getting A or C such that he would consider it a fair coin flip against B. Let us say that he responds that if you could guarantee him a 2/3 chance of getting A versus a 1/3 chance of getting C, then this would be a fair equivalent of B for sure. We arbitrarily assign the least preferred alternative C the value 0, and the most preferred A, 1.0; B would therefore be 2/3. If one continues this lottery process for all the alternatives available, a linear scale of utilities can be generated. In the umbrella problem, the decision maker (to obtain the previously mentioned utilities) would have had to claim that carrying an umbrella in the rain was equally desirable to not carrying an umbrella when it did not rain, and that carrying an umbrella when it did not rain was equivalent to a 90% chance of one of these against a 10% chance of getting wet.

These utility functions generate a rather peculiar scale that is cardinal in the sense of assigning numbers to the various alternatives, but it is not metric (like money) in the sense that the numbers do not represent true relative preferences. The most useful analogy is between a utility scale and a temperature scale. Although numbers are assigned to different degrees of temperature, there is no fixed end point nor a fixed zero point. A temperature scale will give us the relative positions of various temperatures, but they are not comparable in the same fashion as a metric scale (*i.e.*, it makes no sense to speak of 80°F as twice as hot as 40°F when an alternative scale, centigrade, yields numbers 27°C and 4.5°C) (Swalm, 1966). Another scale similar to utility is the calendar years.

Perhaps the clearest application of the lottery procedure to determine utilities is the generation of a utility scale for money. In this case, a person is asked to choose between X dollars for certain versus a gamble between Y and Z dollars. In most circumstances, people will choose a sum for certain that is lower than the dollar sum times the probabilities in the gamble. For example, if dollars equaled utility, $100 would be the fair equivalent of a 50/50 bet between $200 and $0.; however, most people would take the $100 for certain. This indicates that $100 has more utility for them than the 50/50 gamble between $200 and $0. A scale of utility created for money is shown in Fig. 3.

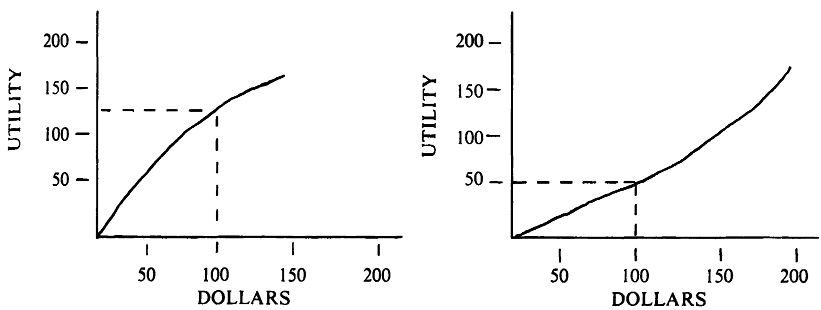


FIG. 3. Scale of utility created to illustrate the utility placed on money by people applying the lottery procedure. *Left:* Curve of a conservative non-risk-taking individual. *Right:* Curve of a risk-taking gambler.

The shape of one's utility curve for any set of alternatives is very personal and may change with time. Thus, even similarly ranked executives in the same company, when asked to make monetary decisions in the company's name, had dramatically different utility curves for money—however, the curves were predominantly conservative (Swalm, 1966).

Utility theory has been axiomatized by Luce and Raiffa (1957) as follows:

i. Any two alternatives shall be comparable, *i.e.*, given any two, the subject will prefer one to the other or he will be indifferent between them.
ii. Both the preference and indifference relations for lotteries are transitive, *i.e.*, given any three lotteries A, B, and C, if he prefers A to B and B to C then he prefers A to C; and if he is indifferent between A and B and between B and C, then he is indifferent between A and C.
iii. In case a lottery has as one of its alternatives (prizes) another lottery, then the first lottery is decomposable into the more basic alternatives through the use of the probability calculus.
iv. If two lotteries are indifferent to the subject, then they are interchangeable as alternatives in any compound lottery.
v. If two lotteries involve the same two alternatives, then the one in which the more preferred alternative has a higher probability of occurring is preferred.
vi. If A is preferred to B and B to C, then there exists a lottery involving A and C (with appropriate probabilities) which is indifferent to B.

One advantage of the axiomatic method is that it explicitly lays out the foundations of a theory so that it can be easily critiqued. Axioms i, ii, iv, and v seem intuitively reasonable. Axiom iii, after some reflection, seems reasonable, but axiom vi is somewhat controversial. The claim in axiom vi is that all outcomes can be rated with respect to each other. Is this truly the case? Medical decisions often deal with intangibles such as life, death, pain, and relief that seem to defy comparison with more mundane matters such as money, time lost from work, temporary functional impairment, etc. In the business world where decision theory has been applied more extensively, intangibles such as goodwill and community standing have been measured (Swalm, 1966). However, it is not clear that the intangibles involved in medical problems can be rated with respect to each other. A further claim of axiom vi is that one's betting behavior in a gamble or lottery truly reflects one's preferences. Is this always the case?

These axioms generate a series of utility functions that assign numbers to alternatives in a lottery, and those numbers reflect the decision maker's preferences in that lottery. If numerous alternatives are subjected to this lottery type expression of preference, a linear utility scale is derived. Thus, in most medical decisions this amounts to ranking preferences for diagnoses or therapies (in terms of benignity, treatability, cost, pain, etc.)

Problems with Utility Assignments. From the preceding discussion, several problems with utility assignments appear. The first and most obvious is that monetary value cannot be substituted for utility assignments. Although we will return to this later in a discussion of benefit-cost analysis, it should be clear that a substitution of monetary value for utility in a decision problem presumes that everyone has the same utility curve for money. This is not the case, as we saw in the last section. Although no medical decision analysis has expressed utilities in dollar terms, the choice of utility assignments often suggests a monetary equivalent (Gorry, Kassirer, Essig et al., 1973; Schwartz, 1973).

The second problem in utility assignment is designating utility functions for various alternatives without the lottery procedure. This is technically not legitimate because of person-to-person variability. However, arbitrary utility assignments are routinely reported in the literature because individual lotteries are unfeasible (Gorry, Kassirer, Essig et al., 1973; Schwartz, 1973). A surgical scar, the cost of an operation, pain, etc., all have personal values. Utility functions for one individual cannot be carried over to another individual or group. The fact that each utility function for each alternative for each person would have to be recalculated each time a decision was made seems to severely limit the technical feasibility of utilizing decision theory for individual patients.

Kassirer (1976) points out that some of these criticisms, although correct, miss the point that these functions will often be the same from person to person and from one point in time to another. Thus, with the understanding that the analysis is not rigorous but considerably better than informal clinical reasoning, we utilize this model to help us sort the issues involved in a complex clinical decision.

One issue we mentioned earlier is that of intangibles. How does one rate his preference with regard to life, death, and other abstract notions? One answer is that we do it all the time by risking death when we decide to go mountain climbing, kayaking, or even taking a drive in a car. In medical decisions, substantial mortality and morbidity may be present in one or all of the alternative therapies. This affects the utility of those therapies. The manner in which these factors go into generating an individual's utility for an alternative is still opaque and is part of the problem of multi-attribute utilities (Keeney and Raiffa, 1976). That is, a linear utility scale may fail to capture all the subtle complexities and tangled web of attributes that each alternative has in a clinical decision problem (Hausner, 1954: 167; Huber, 1974; Keeney, 1972). Clearly, viewing medical options simply in terms of life and death is inadequate (Barnoon and Wolfe, 1972). We need a utility measure that can integrate diverse attributes of therapeutic alternatives such as cost, morbidity, mortality, expected compliance, treatability, discomfort, loss of self-image, etc. Although there are mathematical tools for manipulating multidimensional utilities, only infrequently has this been shown to have practical application (Gardiner and Edwards, 1975: 2; Ginsberg, 1971: 42).

The problem of interpersonal comparisons of utility has not been solved satisfactorily from a theoretic standpoint, and certain aspects are important for medical decisions. Whenever physicians make clinical decisions they integrate their own value system with the patient's value system to generate preferences for alternative diagnoses or therapies. For example, when a physician recommends an operation to a patient, he has assessed the patient's value of pain and integrated it into his decision. When there is any question about this integration, the patient may express his doubt in terms like: "Doctor, I don't want this operation if there is a chance I might die." In most cases a discussion would result in a satisfactory mutual understanding of the relative benefits and detriments of each course of action. The presumption in medical decisions is that the decision is made with an interpersonally accurate utility function until conflict arises. Claims that the patient is being experimented upon are conflicts of this nature, *i.e.*, the patient claims that the doctor is not acting in his behalf, and that the doctor's utility assignments for the alternatives involve weighing them from a scientific or a broad social viewpoint and not from the patient's viewpoint. It is clear that either using the patient's or the physician's utility assignments alone is inadequate.

Basically, the physician is in the position of a benevolent dictator. He generates decisions in the patient's behalf with some appreciation of the patient's value system. The ability to assess and integrate patients' values is one of the subtle attributes of a good clinician. There is no formula that will act as a guide to the integration of the patient's values into the decision-making process. However, one must look for key indicators, primarily from the patient's past experience, that would lead him to believe the patient's values are the same or different than his. Of course, the greater the cultural gap between patient and physician, the more difficult this appreciation of values. In the end, when conflict is unresolvable, the patient has clear-cut veto power over every decision—this is inherent in patients' rights.

Decision Criteria Under Risk

For decision problems under risk, there is one accepted decision criterion—choosing the alternative that maximizes expected utility. That is, for each action there is a set of possible outcomes. Each outcome is associated with a probability of occurrence and, as we saw earlier, a value assignment or utility function. The expected utility for each outcome is the probability multiplied by the utility. The expected utility for each act is the sum of the expected utilities for all the possible outcomes of the action. The decision is generated by comparing the expected utilities for each act or alternative and choosing the maximal expected utility of those alternatives—a Bayes strategy. (See the umbrella problem, pp. 365–367.)

This principle may be sound as a *prescription*, but it is not an accurate *description* of decision behavior. A problem with maximized expected utility as a description of behavior, called the "sure thing" principle, was pointed out by Keynes (1956: 1360) and Savage (1972: 91). A medical variation on the "sure thing" principle is that a patient faced with the choice of living with an illness with which he is familiar versus opting for a medical intervention may opt to remain with the condition because of fear of the unknown. The option that is "for sure" is chosen even though it has lower expected utility. The patient who does this may be accused of being irrational, and he is, if the criterion of maximized expected utility is the criterion of rationality. Maximized expected utility is the decision criterion that embraces what we consider rational decision making, but it fails to describe the broad range of both rational and irrational human decision making behavior. This is a criticism of human decision-making rather than decision theory.

Individual Normative Decision Analyses Under Risk

A wide range of individual medical problems has been subjected to rigorous decision analysis, including coronary artery bypass surgery, right lower quadrant abdominal pain, radical neck dissection for oropharyngeal cancer, streptococcal sore throat therapy, post-myocardial-infarction prophylactic anticoagulation, swine flu inoculation, and the screening test in general (Anderson, Lierena, Davidson et al., 1976; Bay, Flatham, and Nestman, 1976; Bunch and Andrew, 1971; Emerson, 1975; Emerson, Teather, and Handley, 1974; Forst, 1971; Giauque, 1972; Hensche and Flehinger, 1967; MacRae, 1976; Pauker, 1976; Pauker and Pauker, 1977; Pliskin, 1975; Rubel, 1967; Safran, Tsichlis, Bluming et al., 1977; Schoenbaum, McNeil, and Kavet, 1976; Sisson, Schoomacker, and Ross, 1976; Teather, Emerson, and Wandley, 1974; Tompkins, Burnes, and Cable, 1977). The only apparent limitation on the range of these analyses is the time and effort involved with the increasing complexity of the problems undertaken. The decision analyses discussed in this section are management problems. Some are pure therapeutic decisions with known illnesses, such as a mid-shaft fracture, and some are combined diagnostic-therapeutic dilemmas. These combined analyses are necessarily more complex, but they open options to the decision maker such as halting a diagnostic evaluation to initiate a therapeutic trial. The staggering complexity of these analyses requires simplifying assumptions even in the automated programs (Betaque and Gorry, 1971; Ginsberg and Offendsend, 1968). Yet many of the analyses yield significant results with pencil-and-paper-type calculations (Hensche and Flehinger, 1967). The simplifying assumptions can occasionally lead to eliminating important options such as the possibility of an unsuspected diagnosis in an early analysis by Ginsberg (1971) or the introduction of a new and less risky diagnostic test, the gallium scan, as a replacement for the lymphangiogram test being analyzed (Safran, Desforges, Tsichlis et al., 1977). (This test was added in another publication (Safran, Tsichlis, Bluming et al., 1977).) Decision analyses are likely to eliminate important options when either the problem is ambiguous or the technology used to evaluate the problem is rapidly evolving.

From a theoretic standpoint, most of these analyses are relatively similar. The probability estimates are arrived at in one of two ways. Most commonly, they are the subjective estimates of an expert in the field (Pliskin, 1975). Less commonly they are frequency data obtained from the literature (Kassirer, 1976; Pauker, 1976). Rarely, probability estimates incorporate frequency data accumulated expressly for the purpose of the study (Emerson, 1975). For a diagnostic problem we may assume that the prior probabilities (before any tests were performed) are equal or that they reflect the prevalence of the disease in the population at risk (Hockstra and Miller, 1976). Some programs were designed to update the probability estimates as data accumulated (Safran, Tsichlis, Bluming et al., 1977). Other programs had cut-off probabilities beyond which the diagnoses were not pursued (Gorry, Kassirer, Essig et al., 1973). In recent analyses, the trend has been to subject the probability estimates to "sensitivity analysis," which involves calculating the expected utility for all the alternative courses of action using the upper and lower limits of the probability range for each of the outcomes (Kassirer, 1976; Pauker and Kassirer, 1975). Usually the probability estimates and utility estimates are varied separately during a sensitivity analysis. Conceivably, both estimates could be in error simultaneously, but this is rarely considered.

The utility assignments were obtained by a wider range of procedures than the probability estimates. In some of the less sophisticated analyses the utility assignments were the author's arbitrary value estimations (Hensche and Flehinger, 1967; Sisson, Schoomacker, and Ross, 1976). More rigorous analyses utilized the lottery process described in an earlier section (Betaque and Gorry, 1971; Schwartz, Gorry, Kassirer et al., 1973; Thornbury and Fryback, 1976). Most of the utility scale determinations were performed with the same group of expert physicians who generated the probability estimates. In the automated analyses it is difficult to envision a simple way of incorporating a patient's utility scale into the program. In a few of the recent nonautomated analyses, patients' utility scales for some of the major attributes of the alternatives were assessed (*i.e.*, pain and survival) and incorporated into the analysis (Pliskin and Beck, 1976). Pauker (1976) and Pauker and Pauker (1977) explain in some detail the application of the lottery process to an individual patient.

Certain studies, apparently realizing the difficulty in trying to establish a meaningful interpersonal utility scale, left the expected value of the various alternatives in "raw form," *e.g.*, expected mortality (Emerson, 1975; Pauker and Kassirer, 1975). One analysis left all the value measurements (expected hospital days, expected days off work, expected number of amputations, and expected number of deaths) with the corresponding therapeutic options in table form, allowing decision makers the luxury of choosing the option that matched their value systems (Bunch and Andrew, 1971). This is only possible for a small number of options with a relatively few *tangible* attributes.

Another approach is to use the therapeutic threshold, which indicates that a certain action should be taken if the *a priori* probability of an alternative exceeds a certain threshold level (Pauker and Kassirer, 1975; Safran, Desforges, Tsichlis et al., 1977). These methods and other novel solutions to the utility problem may overcome its present limitations. A thorough investigation of the utility problem including the solution of a problem using a multi-attribute utility function from several perspectives (the doctor, the patient, and the public health official) may be found in Giauque (1972) or Ginsberg (1971). Because of this complexity, the analyses are quite lengthy.

When these analyses are field-tested either against the judgment of experts or against the results of a trial (*e.g.*, looking for the expected mortality under certain circumstances), they seem to perform well (Emerson, 1975; Gorry, Kassirer, Essig et al., 1973).

Experimental Decision Theory

Experimental decision theory attempts to examine, among other things, whether decision theory is a good descriptive model of behavior. Experiments seeking to predict an individual's assessments of probabilities or their ability to generate a linear utility scale for a series of alternatives have been performed since the 1950s. In general, they demonstrate that individuals have limitations in both areas (Edwards, 1954; 1961; Suppes, Davidson and Siegal, 1957). The subjective probability estimates of experts are reliable when the utilities are fixed (as in a pure diagnostic problem) and when the probability estimates are being constantly modified by feedback information in the form of the patients' eventual outcomes (de Dombal, 1975; Salamon, Bernadet, Samson et al., 1976). When the probabilities are fixed but the utilities are variable, the experts seem less reliable. Their actions seem to betray an inconsistency in the values that they attached to the determinants of therapeutic outcomes (Aitchison, Moore, West et al., 1973; Taylor, Aitchison, Parker et al., 1975). In spite of this variability, if there are no mitigating circumstances, one can reason backward from the physician's choice in a decision problem to his approximate utility assignments for the alternatives (Aitchison, 1970). Thus, although decision theory is not a very accurate descriptive model, it can be a valuable tool with which to empirically analyze medical decisions (Taylor, Aitchison, and McGirr, 1971).

Proving Decision Theory

Decision analyses reflect the opinions of experts. However, this only lends a certain amount of credence to the claim that the results of decision analyses are the "best" decisions (Betaque and Gorry, 1971; Gorry, Kassirer, Essig et al., 1973). If one requires actual proof that investing time, effort, and expertise in decision analysis will pay off in terms of better outcomes of one's decisions, then there is a great paucity of data to support this contention. In fact, athough most observers agree that a poor analysis can be misleading, there is no methodology, save common sense, that allows one to distinguish accurate from inaccurate analyses. In spite of the lack of proof, many observers (especially in industry) feel that any analysis is superior to the alternative—*ad hoc* decision making—because at least it makes the presumptions of the decision explicit.

A more limited goal of showing whether the use of decision analysis improved decisions, in terms of previously stated utilities rather than patient outcomes, was attempted by Fryback (1974). The decision making clearly improved, but it was not shown that patient outcomes also improved.

The problem of what to do with counterintuitive results has not been satisfactorily solved. Giauque's conclusion (1972) that all members of a family of a patient with a sore throat should have their throats cultured even though the patient's culture is negative seems unwarranted, just as screening programs for asymptomatic strep carriers outside of military populations seem unwarranted. Is the analysis wrong or are our intuitions incorrect?

Game Theory

Game theory is a subdivision of decision theory which attempts to optimize an action or strategy for two or more decison makers who may either cooperate with each other or conflict with each other (Luce and Raiffa, 1957). An interesting approach to the problem of developing a meaningful interpersonal utility measure was investigated by Hockstra and Miller (1976). They viewed sequential medical diagnostic decisions in terms of a two-person, doctor-and-patient, cooperative game. The patient's utility was to limit cost expenditures, and the doctor's utility was to achieve certainty of diagnosis. The optimal strategy was defined as the Nash equilibrium, which essentially states that, of all the tests that generate information (*i.e.,* increase the probability of diagnosis), the test that minimizes cost should be chosen. The "ceasing testing" rule is that, once a point is reached where any further test yields no increased information, the diagnostic process is terminated.

Unfortunately, this procedure does not solve the fundamental problem of interpersonal utility comparisons. In reality, the doctor and patient have virtually the same set of objectives with regard to a given medical problem. In general, they both want to become more certain of the diagnosis with the least expense, time, and pain. It is not the doctor's need for certainty versus the patient's desire to minimize expenses, but more accurately a situation that demands bargaining between the doctor and patient to arrive at a joint, agreed upon, value for each of the aspects of the various alternatives. Furthermore, one does not ordinarily utilize every test that yields information, but ceases testing at some arbitrary point when the diagnosis is considered "proven."

Decision-Making Under Ignorance and Statistical Inference

If one can specify all the alternative actions as well as all of their respective outcomes for a given decision problem, but is totally ignorant of the probabilities of these outcomes, the decision problem is *under ignorance.* Examples given earlier included the decision whether to take an umbrella or not without a forecast, and the decision to initiate therapy or not without any idea of the likelihood of the alternative diagnoses. These may be visualized in a matrix:

	States of Nature			
	S_1	S_2	$S_3 \ldots$	S_n
Actions A_1	$O_{1,1}/U_{1,1}$			
A_2	$O_{2,1}/U_{2,1}$			
A_3	$O_{3,1}/U_{3,1}$			
.				
.				
.				
A_m	$O_{m,1}/U_{m,1}$			$O_{m,n}/U_{m,n}$

Thus, the decision must be made on the basis of the relative values of the outcomes, not on their chance of occurring. Part of the problem with this notion lies in the concept of complete ignorance. That is, it appears as though simply being able to enumerate the possible outcomes is an admission of some information concerning the relative probabilities of these outcomes as the most probable subset of all possible outcomes (Luce and Raiffa, 1957). A thorough discussion of "complete ignorance" is beyond the scope of this paper.

To handle decisions of complete ignorance, several decision rules have been proposed. Of these, the best known is the Maximin Criterion of Von Neumann (Von Neumann and Morgenstern, 1947). An index is created for each act. The index is the numerical value of its smallest utility assignment for any of its possible outcomes. The maximin decision rule is to choose the action whose index is the largest. It is clear that this decision rule is extremely conservative. Witness the following decision problem:

	S_1	S_2
A_1	0	100
A_2	1	1

The maximin criterion dictates the counterintuitive choice of A_2. Other proposed decision criteria include maximizing the average utility assignment, the Minimax Risk criterion, and the Pessimism-Optimism Index of Hurwicz. None of these criteria is wholly satisfactory, in that one can create examples for which any of these criteria generate counterintuitive results.

Classical statistical inference hypothesis testing is an example of decision making under ignorance, as follows:

	no. with disease	no. without disease
positive test	true positive	false positive type II or β error
negative test	false negative type I or α error	true negative

This is a two-state (disease versus normal), two-action (positive test versus negative test) decision problem. The problem is that, if one of these errors is decreased by a decision rule, the other is invariably increased.

The classical Neymann-Pearson decision solution is to accept a hypothesis if α is ≤ 0.05 or 0.01. This arbitrary rule, like the maximin criterion, clearly fails under certain circumstances. When the cost or disutility of missing a diseased patient is high, such as in phenylketonuria screening, a false-negative rate of less than 0.01 is necessitated even at the expense of a 95% false-positive rate (Galen and Gambino, 1975). The high false-positive rate may be expressed as low predictive accuracy of a positive test TP/TP + FP (Vecchio, 1966). Conversely, if a false-positive is associated with a significant disutility or "cost" and a marginal utility, one would like to minimize type II or β errors. For example, the diagnoses of schizophrenia and alcoholism have significant social detriments and may have little therapeutic importance. If the disutility of an FP or FN is substantial, the Neymann-Pearson decision may be inapplicable. A thorough discussion of the relationship between classical statistical inference and Bayesian decision theory is available in many statistics textbooks (Chernoff and Moses, 1959; Hamburg, 1970).

Many of the unwritten laws in medicine (decision rules of ignorance) have analogies in the legal system (Scheff, 1963). Just as the law uses the rule, "Better a thousand guilty men go free than one innocent man convicted," medicine uses the maxim, "When in doubt, diagnose illness." Both rules are guides to minimizing type I errors.

		Reality	
		innocent	guilty
Verdict	innocent	correct judgment	guilty man freed type II or β error
	guilty	innocent man convicted type I or α error	correct judgment

		Reality	
		ill	not ill
Diagnosis	ill	correct	false diagnosis
	not ill	false bill of health	correct

Both society and medical training promote the minimization of type I errors, in the belief that it is worse to ignore a disease that is present than to diagnose a well person as ill. Scheff points out that certain conditions might require minimizing the false diagnoses, even at the expense of increasing false bills of health. There is a need to establish exactly what conditions benefit by early intervention such that we do minimize false-negative ($\propto$) in these groups. Other diseases, such as benign self-limited conditions or chronic untreatable disorders, would require less hurried investigation.

Another way of visualizing the α and β error relationship is in terms of a Receiver Operating Characteristic (ROC) curve (Lusted, 1968; 1971; Lusted, 1968), as shown in Fig. 4. Once a ROC curve is obtained for a signal detector (*i.e.*, radiologist), several things can be done with it. One can look at the actual outcomes of false-positive and false-negative readings and see if the degree of "conservatism" is warranted. If the patient outcomes do not appear to warrant this degree of conservatism, it has been shown that one can change the ROC characteristics willfully by reading the X-rays with a liberal or conservative attitude (Garland, 1959).

Another use of ROC curves is in training personnel. A specialist's ROC characteristics could be measured and used as a goal for a trainee. A third use is in evaluating new technology in terms of decreased false-positive or false-negative interpretations.

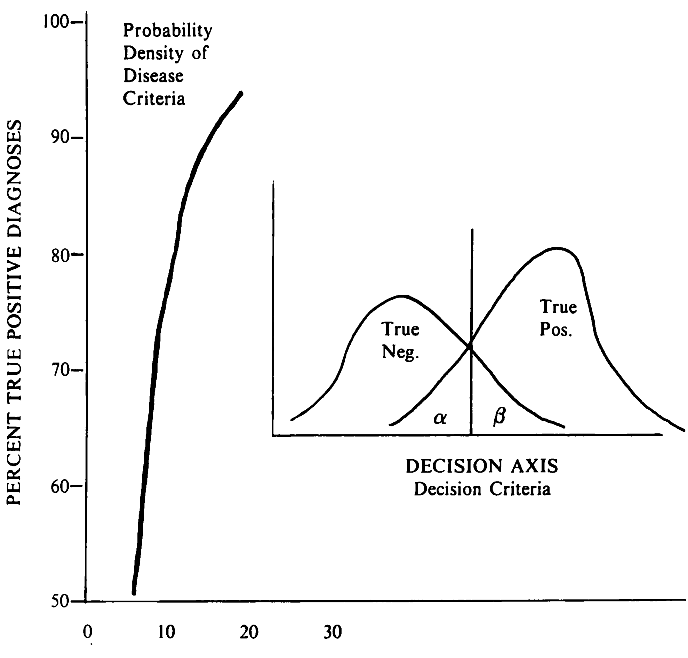


FIG. 4. Receiver Operating Characteristic (ROC) curve. A reciprocal relation is demonstrated between the percentage of true-positive and false-positive diagnoses. Hypothetical population density curves that generated the ROC curve are shown in the upper right diagram. The α and β areas represent false-negative and false-positive diagnoses. (Adapted from Lusted, L. B. 1971. Decision-Making Studies in Patient Management. *The New England Journal of Medicine* 284(8): 416.)

Cost-Benefit Analysis

Cost-benefit analysis is part of the broad field of welfare economics. In spite of its development within the framework of public finance, it may be considered a subset of decision theory (Prest and Turvey, 1965). Within the domain of decision theory it forms one of several normative-group decision theories, and usually constitutes decision making under certainty. It generally enumerates in monetary terms costs and benefits of a willful action, with the decision rule that the alternative with the highest benefit to cost ratio should be adopted.

This technique has been applied to the health-related fields in increasing frequency since the early 1960s and has resulted in studies of heart disease, cancer, stroke, arthritis, venereal disease control, mental illness, and, more recently, to the stool guaiac examination (Fein, 1958; Klarman, 1964: 693; Klarman, 1965: 367; Neuhauser and Lewicki, 1976; Neuhauser and Lewicki, 1975; Department of Health, Education, and Welfare, 1966; Rice, 1962).

Most of the work in cost-benefit analysis has been performed on a public policy level because money is an acceptable, well-understood value measure, and it is the basis on which government agencies often make policy decisions. Even in public policy, however, it is difficult to assess the monetary value of public goods such as clean air (Pliskin and Taylor, 1977). To avoid this trade-off of intangibles such as clean air or lives for dollars, implied by the analysis, a simple method may be used to compare different strategies with similar outcomes in terms of their cost-*effectiveness*. This may be expressed in a monetary or a non-monetary standard, *e.g.*, quality-adjusted years of life (Zeckhauser, 1975).

If one chooses monetary value for the utility measure, there are obvious moral and ethical problems. For example, although a stool guaiac examination costs about $1, when applied as a public policy to six sequential stool guaiac examinations as a screen for silent colorectal cancer, the marginal cost of the sixth guaiac per cancer found is $47,107,214.00 (Neuhauser and Lewicki, 1975). Clearly, this is unlikely to be a cost-effective public policy, but should it for that reason be denied to an individual who requests six guaiac exams? The transition from a prescription for a public policy program to an individual decision is difficult. Cost-benefit analysis has been performed in the individual decision analysis of a space-occupying lesion in the kidney on intravenous pyelogram and in the prevention of acute rheumatic fever (Thornbury and Fryback, 1976; Tompkins, Burnes and Cable, 1977). Cost-benefit analyses of individual decision problems are legitimate when the major therapeutic alternatives differ primarily in costs rather than the more highly subjective qualities of pain, disability, etc.

The absence of probabilities arises from the nature of the analyses. That is, they are generally of existant problems or policies, or future policies with assumed guaranteed outcomes. Of course, this need not be so, and a recent analysis of the colorectal cancer screening program did use probabilities extensively (Neuhauser and Lewicki, 1975: 226).

The choice of money as an interpersonal utility or value standard carries certain advantages and disadvantages. Michalos points out (1970: 67) that maximized expected utility, at least in the Von Neumann-Morgenstern conception of generating numerical utility functions by means of multiple gambles, is thoroughly impractical. He argues for monetary value over maximized expected utilities in situations such as hypothesis acceptance by groups where other standards of utility are lacking. But whose valuation do we use for intangibles such as a year of life—the person's expected income? his loss to society? insurance policy standards? or what an individual would be willing to spend to buy a year of life? (Acton, 1975; Schelling 1968: 217; Tompkins, Burnes and Cable, 1977).

Other cardinal scales of utility instead of money are possible and include disease impact analysis (*i.e.*, theoretic preventable impairment versus actual prevented impairment) and expected length of useful life (*i.e.*, quality-adjusted life years) (Weinstein and Stason, 1976; Williamson, Alexander, and Miller, 1968). However, it is unclear that an individual can readily convert pain or disfigurement into years of life or vice versa.

In conclusion, cost-benefit analysis, as it stands, appears appropriate in evaluating existing public policy programs and problems. With the modification of adding probabilities, it appears valid to evaluate proposed programs with outcomes of less than certain probability, and to evaluate the relatively rare individual decisions where the monetary outcomes are not severely distorted by unmeasurable intangibles. Where disability is a major outcome, the addition of a measure of useful years of life may prove to be an important utility measure. Lastly, in establishing priorities for program innovation, disease impact analysis may be helpful.

Flow Charts

Flow charts are explicit strategies for solving a problem. The problem can be as mundane as the acquisition of raw data in a routine history or as complex as the schematic representation of the evaluation of a clinical finding. The branches are created by additional data. For example, a flow chart evaluating a pleural effusion might have branches representing exudative fluid versus transudative fluid or perhaps unilateral versus bilateral. In most circumstances, the branches are created by the response to a "Yes/No" question or an "Either/Or" condition, but this is not necessary. For example, responses to the initial question in a flow chart on hives include "No," "Yes," "Don't know," and "Don't understand." Alternatively, a flow chart algorithm may be presented as a decision table (Holland, 1975). In spite of the name "decision table" or its representation as a tree diagram, flow charts are only distant relatives of decision analysis.

In terms of a decision theory classification, flow charts are, in general, sequential decision problems where neither utilities nor probabilities are specified. The only aspects of the problem that are specified are the various decision points and the outcomes. The outcomes are often not wholly specified but simply that the choice (which may or may not be a chance phenomenon) of the next decision problem rests with the outcome of the previous one. For example, in evaluating chest pain the physician may be instructed to ask whether it increases with inspiration.

	State	
	Yes	No
Action increases with inspiration	Go to Question 2	Go to Question 3

Implicit in this format is the claim that any chest pain that increases with inspiration increases the probability of certain diagnostic entities such as pleurisy or pulmonary embolus. The Question 2 branch would pursue these diagnostic entities.

In a decision analysis, we would like to know both the value and probability of the outcomes of our actions whether it is performing a lab test, buying stock, or stretching out on the sofa. In the flow chart, we often want a format that will show us how to assess the state some person or object is already in. What we want to know is what label to attach to it or what to do about it. By a series of answers to questions, we will eventually land up with a diagnosis or a therapeutic instruction.

Flow charts are generated by careful clinical observation and introspection. They are useful descriptive and teaching devices, but they do not address the exact same issues as decision analysis (Feinstein, 1974). For example, a flow chart concerning suspected pulmonary embolism may require the acquisition of certain clinical data such as whether the patient had chest pain, shortness of breath, cough, hemoptysis, etc. On the basis of these and other data, the chart might suggest a lung scan. On the other hand, a decision analysis of the same situation might assume that the data has been acquired and might seek to evaluate the relative benefit of proceeding with a lung scan or instituting therapy without this additional datum. Similarly, a flow chart may illustrate the clinical evaluation of a sore throat whereas a decision analysis might investigate the role of the throat culture in the management of patients with a sore throat. Clearly, flow charts are a *codification* of accepted medical practice, whereas decision analyses *investigate* or *evaluate* these practices.

Conclusions

Decision theory is a broad field with numerous facets, even when restricted to medicine. It has accepted application to public health policy in the form of cost-benefit analysis, and in research endeavors in the form of statistical inference. As an empirical tool it has helped sort out some of the logic of our clinical behavior (Moore, 1973; Taylor, 1971). In the form of linear programming, it is used in administration problem solving.

It remains to be seen what impact decision analysis can have on clinical medicine. Of the applications to clinical medicine, automated diagnoses have had the most extensive experience. They are clearly feasible, albeit difficult, time-consuming, and marginally superior to current practices. A most problematic area is management decision analysis under risk. Certain of these analyses are notable successes, but they remain relatively isolated. The reasons are many. There are pragmatic limitations of time, effort, cost, and expertise necessary to perform an analysis that may generate little tangible benefit. On the theoretic side, most of the problems with probability estimates have been satisfactorily resolved; however, the utility problems still plague most analyses (Ransohoff and Feinstein, 1976). When the decision is between alternatives that differ in tangible attributes, such as mortality or significant morbidity, the decision often can be made using these indices as cardinal utility scales. When the alternatives differ in intangible or multiple attributes, there does not seem to be any wholly satisfactory solution currently. It is clear that the utility measurements should thoroughly incorporate the patient's values; but this is often quite difficult, especially with decisions involving life, death, and pain.

In our search to achieve quantitative guides for clinical behavior, we must keep in mind that every formalism is a model. The primary tasks of models are to describe or prescribe accurately (Vaupel, 1973). Although decision theory is suitable for both endeavors, its limitations make it unlikely that it will become a calculus of clinical medicine with or without the aid of a computer (Pauker, Gorry, Kassirer et al., 1976). In spite of these limitations, it will surely play a role in investigating current clinical practice, teaching medical students from a task-oriented viewpoint, and shaping health care policy (Taylor, 1976). Decision theory's true role in medicine will lie between the naive optimism that characterized it as a "new Rosetta stone" (Hall, 1967) and the unmitigated pessimism that characterized it as a "computerized Ouija board" (Ransohoff and Feinstein, 1976).

References

Acton, J. P. 1975. *Measuring the Social Impact of Heart and Circulatory Disease Programs: Preliminary Framework and Estimates.* Rand Corporation Document R1697, NHLI:10, (April). Santa Monica, Calif.: The Rand Corporation.

Aitchison, J. 1970. Decision-Making in Clinical Medicine. *Journal of the Royal College of Physicians, London* 4(3): 195 ff.

———, Moore, M. E., West, S. A. et al. 1973. Consistency of Treatment Allocation in Thryotoxicosis. *Quarterly Journal of Medicine* 42(167): 575 ff.

Anderson, G., Lierena, C., Davidson, D. et al. 1976. Practical Application of Computer-Assisted Decision-Making in an Antenatal Clinic—A Feasibility Study. *Methods of Information in Medicine* 15(4): 224 ff.

Anderson, J. A., and Boyle, J. A. 1968. Computer Diagnosis: Statistical Aspects. *British Medical Bulletin* 24(3): 233 ff.

Barnoon, S., and Wolfe, H. 1972. *Measuring the Effectiveness of Medical Decisions. Springfield, Ill.:* Charles C Thomas.

Bay, K. S., Flatham, D., and Nestman, L. 1976. The Worth of a Screening Program: An Application of a Statistical Decision Model for the Benefit Evaluation of Screening Projects. *American Journal of Public Health* 66(2): 145 ff.

Betaque, N., and Gorry, G. A. 1971. Automating Judgmental Decision Making for a Serious Medical Problem. *Management Science* 17(8): B421 ff.

Bunch, W. H., and Andrew, G. M. 1971. Use of Decision Theory in Treatment Selection. *Clinical Orthopedics* 80(1): 39 ff.

Chernoff, H., and Moses, L. E. 1959. *Elementary Decision Theory.* New York: John Wiley and Sons, Inc.

Cornfield, J. 1964. Bayes Theorem. In *Proceedings of the Sixth IBM Medical Symposium,* p. 163. New York: International Business Machines Corporation.

Croft, D. J. 1972. Is Computerized Diagnosis Possible? *Computers and Biomedical Research* 5(4): 351 ff.

———, and Machol, R. E. 1974. Mathematical Methods in Medical Diagnosis. *Annals of Biomedical Engineering* 2(1): 69 ff.

deDombal, F. T. 1975. Computer-Aided Diagnosis and Decision-Making in the Acute Abdomen. *Journal of the Royal College of Physicians, London* 9(3): 211 ff.

Edwards, W. 1954. The Theory of Decision Making. *Psychological Bulletin* 51(4): 380 ff.

———. 1961. Behavioral Decision Theory. *Annual Review of Psychology* 12: 473 ff.

Elstein, A. S. 1976. Clinical Judgment: Psychological Research and Medical Practice. *Science* 194(4266): 696 ff.

Emerson, P. A. 1975. Decision Theory in the Prevention of Thromboembolism After Myocardial Infarction. *Journal of the Royal College of Physicians, London* 9(3): 238 ff.

———, Teather, D., and Handley, A. J. 1974. The Application of Decision Theory to the Prevention of Deep Vein Thrombosis Following Myocardial Infarction. *Quarterly Journal of Medicine, London* 43(171): 389 ff.

Fein, R. 1958. *The Economics of Mental Health.* New York: Basic Books.

Feinstein, A. R. 1974. An Analysis of Diagnostic Reasoning III—The Construction of Clinical Algorithms. *Yale Journal of Biology and Medicine* 47(1): 5 ff.

———. 1977. Clinical Biostatistics XXXIV—The Haze of Bayes, The Aerial Palaces of Decision Analysis, and the Computer Ouija Board. *Clinical Pharmacology and Therapeutics* 21(4): 482 ff.

Forst, B. E. 1971. *A Doctor's Introduction to Decision Analysis.* Center for Naval Analyses Professional Paper No. 82. Arlington, Va.: Center for Naval Analyses.

Fryback, D. 1974. Use of Radiologist Subjective Probability Estimates in a Medical Decision Making Problem. Ph.D. Dissertation. Ann Arbor, Mich.: University of Michigan.

Galen, R. S., and Gambino, S. R. 1975. *Beyond Normality.* New York: John Wiley and Sons, Inc.

Gardiner, P. E., and Edwards, W. 1975. Public Values: Multiattribute Utility Measurement for Social Decision Making. In *Human Judgment and Decision Processes,* p. 2. Kaplan, M. F., and Schwartz, S., eds., New York: Academic Press, Inc.

Garland, H. 1959. Studies on the Accuracy of Diagnosis. *American Journal of Roentgenology* 82(1): 25 ff.

Giauque, W. 1972. Prevention and Treatment of Streptococcal Sore Throat and Rheumatic Fever. Ph.D. Dissertation. Cambridge, Mass.: Harvard University.

Ginsberg, A. J. 1971. *Decision Analysis in Clinical Patient Management with an Application to the Pleural Effusion Syndrome.* Rand Corporation Document R-751-Rc/NLM: 42. Santa Monica, Calif.: The Rand Corporation.

Ginsberg, A. S., and Offendsend, F. L. 1968. An Application of Decision Theory to a Medical Diagnosis-Treatment Problem. *IEEE Transactions on Systems Science and Cybernetics* SSL-4(3): 355 ff.

Good, I. J. 1959. Kinds of Probability. *Science* 129(3347): 443 ff.

Gorry, G. A., and Barnett, G. O. 1968. Experience With a Model of Sequential Diagnosis. *Computers and Biomedical Research* 1(5): 490 ff.

———. 1970. Modeling the Diagnostic Process. *The Journal of Medical Education* 45(5): 293 ff.

———, Kassirer, J. P., Essig, A. et al. 1973. Decision Analysis as the Basis for Computer-Aided Management of Acute Renal Failure. *American Journal of Medicine* 55(4): 473 ff.

Gustafson, D. H., Kestly, J. J., Griest, J. H. et al. 1976. Initial Evaluation of a Subjective Bayesian Diagnostic System. *Health Services Research* 6(3): 204 ff.

Hall, G. H. 1967. The Clinical Application of Bayes' Theorem. *Lancet* 2: 555 ff.

Hamburg, Morris. 1970. *Statistical Analysis for Decision Making.* New York: Harcourt, Brace and World. p. 741.

Hausner, M. 1954. Multidimensional Utilities. In *Decision Processes,* p. 167, Thrall, R. M. and Coombs, C. H., eds., New York: John Wiley and Sons, Inc.

Hensche, V. K., and Flehinger, B. J. 1967. Decision Theory in Cancer Therapy. *Cancer* 20(11): 1819 ff.

Hockstra, D. J., and Miller, S. D. 1976. Sequential Games and Medical Diagnosis. *Computers and Biomedical Research* 9(3): 205 ff.

Holland, R. R. 1975. Decision Tables: Their Use For The Presentation of Clinical Algorithms. *Journal of the American Medical Association* 233(5): 455 ff.

Huber, O. 1974. An Axiomatic System for Multidimensional Preferences. *Theory and Decision* 5(2): 161 ff.

Hurwitz, L. 1968. In Messick, D. M., ed., *Mathematical Thinking in the Behavioral Sciences,* p. 89. San Francisco: W. H. Freeman and Co.

Kassirer, J. P. 1976. The Principles of Clinical Decision Making: An Introduction to Decision Analysis. *Yale Journal of Biology and Medicine* 49(2): 149 ff.

Kayser, K. 1975. Logic and Diagnosis. *Methods of Information in Medicine* 14(2): 76 ff.

Keeney, R. L. 1972 Utility Functions for Multiattributed Consequences. *Management Science* 18(5): 276 ff.

———, and Raiffa, H. 1976. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* New York: John Wiley and Sons, Inc.

Keynes, J. M. 1956. The Application of Probability to Conduct. Reprinted in Newman, J., ed., *The World of Mathematics,* p. 1360. New York: Basic Books.

King, D. J., and Manegold, R. F. 1965. The Consistency of Cardiologists' Prognostic Judgments in Cases of Myocardial Infarction. *American Journal of Cardiology* 15(1): 27 ff.

Klarman, H. 1964. Socioeconomic Impact of Heart Disease. In *Second National Conference on Cardiovascular Disease* 2: 693. U.S. Public Health Services, Publication No. 1350. Washington, D.C.: U.S. Government Printing Office.

———. 1965. Syphilis Control Programs. In Dorfman, R., ed., *Measuring*

Benefits of Governmental Investments, p. 367. Washington, D.C.: The Brookings Institute.

Kyburg, H., and Smokler, H. 1964. *Studies in Subjective Probability*. New York: John Wiley and Sons, Inc.

Ledley, R. S., and Lusted, L. B. 1959. Reasoning Foundations of Medical Diagnosis. *Science* 130(3366): 9 ff.

Lindley, D. V. 1975. The Role of Utility in Decision-Making. *Journal of the Royal College of Physicians, London* 9(3): 225 ff.

Luce, D., and Raiffa, H. 1957. *Games and Decisions: A Critical Survey*. New York: John Wiley and Sons, Inc.

Lusted, L. B. 1968. *Introduction to Medical Decision Making*. Springfield, Ill.: Charles C Thomas.

______. 1971. Decision-Making Studies in Patient Management. *The New England Journal of Medicine* 284(8): 416 ff.

MacRae, D. 1976. Decision Analysis and Medical Screening Tests. Working Paper. Department of Political Science, University of North Carolina, Chapel Hill, North Carolina.

Mainland, D. 1967. Statistical Ward Rounds — 5. *Clinical Pharmacology and Therapeutics* 8(5): 738 ff.

McKusick, V. A. 1964. *Human Genetics*. p. 151. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

McNeil, B. J., Keeler, E., and Adelstein, S. J. 1975. Primer on Certain Elements of Medical Decision Making. *The New England Journal of Medicine* 293(5): 211 ff.

______, and Adelstein, S. J. 1975. Measures of Clinical Efficacy: The Value of Case Finding in Hypertensive Renovascular Disease. *The New England Journal of Medicine* 293(5): 221 ff.

Menges, G. 1970. On Subjective Probability and Related Problems. *Theory and Decision* 1(1): 40 ff.

Michalos, A. C. 1970. Cost-Benefit Versus Expected Utility Acceptance Rules. *Theory and Decision* 1(1): 61 ff.

Moore, M. F. 1973. First Contact Decisions in General Practice. *Lancet* 1: 817 ff.

Murphy, E. A., and Mutalik, G. S. 1969. The Application of Bayesian Methods in Genetic Counselling. *Human Heredity* 19(2): 126 ff.

Neuhauser, D. 1975. What do we gain from the sixth stool guaiac? *The New England Journal of Medicine* 293(5): 226 ff.

______, and Lewicki, A. 1976. National Health Insurance and the Sixth Stool Guaiac. *Policy Analysis* 2(2): 175 ff.

Overall, J. E., and Williams, C. 1963. Conditional Probability Program for Diagnosis of Thyroid Function. *Journal of the American Medical Association* 183(5): 307 ff.

Pauker, S. G. 1976. Coronary Artery Surgery: The Use of Decision Analysis. *Annals of Internal Medicine* 85: 8 ff.

———, and Kassirer, J. P. 1975. Therapeutic Decision Making: A Cost-Benefit Analysis. *The New England Journal of Medicine* 293(5): 229 ff.

———, Gorry, G. A., Kassirer, J. P. et al. 1976. Towards the Simulation of Clinical Cognition. *American Journal of Medicine* 60(7): 981 ff.

Pauker, S. P., and Pauker, S. G. 1977. Prenatal Diagnosis: A Directive Approach to Genetic Counseling Using Decision Analysis. *Yale Journal of Biology and Medicine* 50(3): 275 ff.

Pliskin, J. S. 1975. The Management of Patients with End Stage Renal Failure: A Decision Theoretic Approach. Ph.D. Dissertation. Cambridge, Mass.: Harvard University.

———, and Beck, C. H. 1976. Decision Analysis in Individual Clinical Decision Making: A Real-World Application in Treatment of Renal Disease. *Methods of Information in Medicine* 15(1): 43 ff.

Pliskin, N., and Taylor, A. K. 1977. General Principles: Cost-Benefit and Decision Analysis. In Bunker, J., ed., *Costs, Risks, and Benefits of Surgery*. Fairlawn, New Jersey: Oxford University Press.

Prest, A. R., and Turvey, R. 1965. Cost-Benefit Analysis: A Survey. *Economic Journal* 75(300): 683 ff.

Public Health Programs for Arthritis. Program Analysis Group on Selected Disease Control Programs (September). Washington, D.C.: Department of Health, Education, and Welfare, 1966.

Raiffa, H. 1968. *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*. Reading, Massachusetts: Addison-Wesley.

Ransohoff, D. F., and Feinstein, A. R. 1976. Is Decision Analysis Useful in Clinical Medicine? *Yale Journal of Biology and Medicine* 49(2): 165 ff.

Rice, D. 1962. Economic Costs of Cardiovascular Diseases and Cancer. Reprinted from the *Report of the President's Commission on Heart Disease, Cancer and Stroke* 2: 439 ff.

Ross, P. 1972. Computers in Medical Diagnosis. *CRC Critical Review in Radiologic Sciences* 3: 197 ff.

Rubel, R. A. 1967. Decision Analysis and Medical Diagnosis and Treatment. Ph.D. Dissertation. Cambridge, Mass.: Harvard University Graduate School of Business Administration.

Safran, C., Desforges, J. F., Tsichlis, P. N. et al. 1977. Decision Analysis to

Evaluate Lymphangiography in the Management of Patients With Hodgkin's Disease. *The New England Journal of Medicine* 296(14): 1088 ff.

———, Tsichlis, P. N., Bluming, A. Z. et al. 1977. Diagnostic Planning Using Computer Assisted Decision-Making for Patients with Hodgkin's Disease. *Cancer* 39: 2426 ff.

Salamon, R., Bernadet, M., Samson, M. et al. 1976. Bayesian Method Applied to Decision Making in Neurology—Methodological Considerations. *Methods of Information in Medicine* 15(3): 174 ff.

Savage, L. J. 1972. Historical and Critical Comments on Utility. In *The Foundations of Statistics,* p 91. New York: Dover Publications.

Scheff, T. 1963. Decision Rules, Types of Error and Their Consequences in Medical Diagnosis. *Behavioral Science* 8: 97 ff.

Schelling, T. C. 1968. The Life You Save May Be Your Own. In Chase, S. B., Jr., ed., *Problems in Public Expenditure Analysis,* p. 217. Washington, D.C.: The Brookings Institute.

Schoenbaum, S. C., McNeil, B. J., and Kavet, J. 1976. The Swine Influenza Decision. *The New England Journal of Medicine* 295(14): 759 ff.

Schwartz, W. B., Gorry, G. A., Kassirer, J. P. et al. 1973. Decision Analysis and Clinical Judgment. *American Journal of Medicine* 55(4): 459 ff.

Shapiro, A. 1977. The Evaluation of Clinical Prediction. *The New England Journal of Medicine* 296(26): 1509 ff.

Sisson, J. C., Schoomacker, E. B., and Ross, J. C. 1976. Clinical Decision Analysis. *Journal of the American Medical Association* 236(11): 1259 ff.

Suppes, P. 1967. Decision Theory. In Edwards, P., ed., *The Encyclopedia of Philosophy.* p. 310. New York: Collier, MacMillan.

———, Davidson, D. and Siegal, A. 1957. *Decision Making: An Experimental Approach.* Palo Alto, Calif.: Stanford University Press.

Swalm, R. O. 1966. Utility Theory—Insights Into Risk Taking. *Harvard Business Review* 44(6): 123 ff.

Taylor, T. R. 1976. Clinical Decision Analysis. *Methods of Information in Medicine* 15(1): 216 ff.

———, Aitchison, J., and McGirr, E. M. 1971. Doctors as Decision-Makers: A Computer-Assisted Study of Diagnosis as a Cognitive Skill. *British Medical Journal* 3(5765): 35 ff.

———, Aitchison, J., Parker, L. S., and Moore, M. F. 1975. Individual Differences in Selecting Patients for Regular Hemodialysis. *British Medical Journal* 2(5967): 380 ff.

Teather, D., Emerson, P. A., and Wandley, A. J. 1974. Decision Theory Applied to the Treatment of Deep Venous Thrombosis. *Methods of Information in Medicine* 13(2): 92 ff.

Thornbury, J. R., and Fryback, D. G. 1976. Evaluation of a Computerized Bayesian Model for Diagnosis of Renal Cyst Versus Tumor Versus Normal Variant from the Excretory Urogram Information. *Investigative Radiology* 11(2): 102.

———. 1977. Cost-Benefit Analysis, Medical Decision Making and the Individual Radiologist. *Current Problems in Diagnostic Radiology* 7(2): 3 ff.

Tompkins, R. K., Burnes, D. C., and Cable, W. E. 1977. An Analysis of the Cost-Effectiveness of Pharyngitis Management and Acute Rheumatic Fever Prevention. *Annals of Internal Medicine* 86: 481 ff.

Tversky, A., and Kahneman, D. 1974. Judgment Under Uncertainty: Heuristics and Biases. *Science* 185(4157): 1124 ff.

Vaupel, J. W. 1973. Muddling Through Analytically. Institute of Policy Analysis and Public Affairs, Working Paper No. 573 (May). Durham, North Carolina: Duke University.

Vecchio, T. J. 1966. Predictive Value of a Single Diagnostic Test in an Unselected Population. *The New England Journal of Medicine* 274(21): 1171 ff.

Von Neumann, J., and Morgenstern, O. 1947. *Theory of Games and Economic Behavior.* Second edition. Princeton, New Jersey: Princeton University Press.

Warner, H., Toronto, A. F., Veasey, L. G. et al. 1961. A Mathematical Approach to Medical Diagnosis-Application to Congenital Heart Disease. *Journal of the American Medical Association* 177(3): 177 ff.

Weinstein, M. C., and Stason, W. G. 1976. Foundations of Cost-Effectiveness Analysis for Health and Medical Practices. *The New England Journal of Medicine* 296(13): 716 ff.

Williamson, J., Alexander, M., and Miller, G. E. 1968. Priorities in Patient-Care Research and Continuing Medical Education. *Journal of the American Medical Association* 204(4): 303 ff.

Winkler, R. L. 1972. *Introduction to Bayesian Inference and Decision.* New York: Holt, Rinehart and Winston.

Woodbury, M. A. 1963. Inapplicabilities of Bayes' Theorem to Diagnosis. *Proceeding from the Fifth International Conference in Medical Electronics in Liege, Belgium.* Springfield, Ill.: Charles C Thomas.

Zeckhauser, R. J. 1975. Procedures for Valuing Lives. *Public Policy* 23(4): 419 ff.

Acknowledgments: The author is indebted to Michael Resnik, Ph.D., of the Department of Philosophy, The University of North Carolina at Chapel Hill, and David McKay, M.D., of the Department of Internal Medicine, and The Robert Wood Johnson Program at The University of North Carolina School of Medicine, for their critical review of the manuscript. The author also wishes to acknowledge the editorial assistance of Karen Strain.

Address correspondence to: Daniel A. Albert, M.D., Division of Rheumatology and Immunology, Department of Medicine, University of California, San Diego, School of Medicine, San Diego, California 92104.