General Problems in Medical Decision Making
With Comments on ROC Analysis

Lee B. Lusted

Medical decision-making studies continue to focus on two questions: How do physicians make decisions? How should physicians make decisions? Researchers pursuing the first question emphasize human cognitive processes and the programming of symbol systems to model observed human behavior. Those researchers concentrating on the second question assume that there is a standard of performance against which the physician's decisions can be judged, and to help the physician improve his performance, an array of tools is proposed. These tools include decision trees, Bayesian analysis, decision matrices, receiver operating characteristics (ROC) analysis, and cost–benefit considerations including utility measures. Medical decision-making questions must be answered in an ethical context where ethics and decision analysis are intertwined.

TWO STREAMS of research on medical decision making can be identified. The streams are parallel at certain places and at some points they merge. One stream consists of studies of how physicians make decisions. These studies are related to artificial intelligence, and they emphasize computer augmentation of clinical judgment.

The second stream consists of a development of mathematical tools with which the physician may sharpen his decisions. An underlying thought is that language is a weak and inadequate vehicle for the usual understanding of things. The language of probabilities and utilities (deeply rooted in probability) provides a more sensitive method for transmitting diagnostic information and diagnoses among physicians and patients than the present crude language of "frequently" and "rarely."

Although this article is a survey and a short summary, two subjects are discussed at somewhat greater length: (1) relationships of ethics and medical decision making, and (2) ROC analysis.

ETHICS AND MEDICAL DECISION MAKING

In the past several years the focus has shifted in discussions of medical decision analysis from questions of diagnostic probabilities to questions of cost–benefit relationships, risks, and values. The "what's at stake" issues of decision analysis tend to predominate.

Economists and ethicists writing on medical care problems emphasize that if people want more medical care they are going to have to pay for it, and although economics can help to make the choices more rational, economics cannot provide the ethics and value judgments that guide our decisions. Ethics is not independent of economics. We need to adopt an economic point of view guided by values. But we must choose.

Values are a reflection of ethical beliefs. Bioethics is concerned with the broadest implications of biologic knowledge, and the principal task of medical ethics is to reconcile the welfare of the individual with the welfare of mankind—both must be served.

However, in the absence of a revealed and generally accepted moral philosophy, physicians and patients alike are confused about answers to moral, ethical, and value questions. The answers are to questions about the purpose of the field of medicine in all its related manifestations in our society. Physicians alone should not be expected to determine the priorities for use of medical resources. The physicians must join with other professionals and representatives of the public in setting priorities. Group decisions, however, will depend heavily on the quantity and quality of information provided by physicians. Medical decision analysis studies have for the most part been directed at optimizing the quantity and quality of this medical information even though the goals of optimization haven't always been clear.

To return to the relationship of medical decision analysis and ethics, if the policy that it is better to choose than not to choose can be agreed upon, then rational choices can be developed from the point of view of decision analysis guided by ethical and moral values.

From the Department of Radiology, University of California, San Diego, Calif.
Reprint requests should be addressed to Lee B. Lusted, M.D., Box 2209, Rancho Santa Fe, Calif. 92067.
© 1978 by Grune & Stratton, Inc.
0001-2998/78/0804-0008$01.00/0

Seminars in Nuclear Medicine, Vol. VIII, No. 4 (October), 1978
There are two themes shared by medical ethics and decision analysis. The first is that definite decision criteria should be constructed. The second theme is that decision guidelines should be consistent. These tasks are not easy but the interested reader will find help in three books published recently. The books on decision analysis are Lindley,1 Making Decisions, and Wulff,2 Rational Diagnosis and Treatment. A very interesting book on medical ethics with decision analysis methods applied to ethical problems is Brody,3 Ethical Decisions in Medicine.

Since we have as yet no universally accepted answers to health and medical care questions, individual preferences tend to prevail and our action has often seemed paralyzed in the presence of such diversity of opinion. Medicine, through considerations of decision analysis and ethics, may increase our understanding of a system of values based on a new sense of an interdependence of persons that could gain the assent and allegiance of a majority of us. Some of the underlying value questions may be apparent as we proceed to consider medical decision-making studies.

Medical Decision Analysis

We should continually ask ourselves why we should work on medical decision analysis and how we are going to demonstrate the value of decision analysis to medical students and physicians. In a recent editorial, Bear and Schneiderman4 provide some answers. They review two general areas in which decision analysis is useful in medical decision making: first, in medical decisions that can be clarified by straightforward “cost-benefit analysis” and second, in complex decisions of a recurring nature that involve the advisability, in discrete populations, of tests or procedures with definite risks. They provide further support for the decision analysis concept by demonstrating, for the first time, that decisions made by such analyses frequently differ from those made by clinicians using clinical judgment in a tertiary referral center.

As Lindley1 emphasizes, we should do decision analysis using the ideas of probability, utility, and maximization of expected utility because we believe that better medical decisions will result from the use of a consistent choice strategy. This should apply to individual patients as well as to patients in the aggregate.

You might say that decision analysis operates on a “divide and conquer” philosophy. Problems are broken up into manageable components. We are going to have to demonstrate that decision analysis offers substantial advantages in practical medical diagnosis and treatment situations, in screening situations, and in public health problems, and we are going to have to use language and examples that medical students and clinicians can understand.

Probably 95%-99% of medical decisions, including problems of screening for disease, can use decision analysis. Some problems don’t need a formal decision structure, because analysis wouldn’t be worth the effort. Some problems need a simple Bayes’ analysis while some need only a 2 × 2 decision matrix; but some problems will need the entire array of analysis tools including sensitivity analysis and assessment of patient utilities.

I have found it helpful to discuss medical decision analysis as a sequence of five procedures or tools beginning with decision trees and ending with cost-benefit relationships and utilities. The sequence and relationship are as follows:

Decision Trees
- Decision tree is constructed to show all important courses of action.
- Decision tree is revised or simplified using Bayes’ theorem.
- Utilities are assigned at ends of branches.

Bayes’ Theorem
- Bayes’ theorem is considered as a Bayesian technique when medical decision making is viewed from the Bayesian standpoint.
- Bayes’ theorem is used with diagnostic test outcome to calculate probability of disease.
- 2 × 2 decision matrices are introduced.

Decision Matrices
- 2 × 2 decision matrix is related to a decision node in the decision tree.
- Matrix is related to sensitivity and specificity of a diagnostic test and to prevalence of the disease in question.
- Matrix is related to a decision operating point on a receiver operating (ROC) curve, which is to be used with prevalence of disease in question.
ROC Analysis
Analysis demonstrates the possible trade-offs for various decision outcomes.
Diagnostic usefulness of a test or sequence of tests may be evaluated in terms of error rates, information content, or average cost and average net benefit.

Cost–benefit Relationships; Utilities and Maximization of Expected Utility
We consider the use of multiattribute utility analysis. These considerations lead to a discussion of ethical decisions in medicine.

Decision Trees
A decision tree is a graphic device to display all possible courses of action (acts) with all possible consequences (states). The acts and states are drawn in the order in which they occur in time. The acts branch from square decision nodules, which may be imagined as small $2 \times 2$ decision matrices. At the end of each act the possible states branch from a round chance node.

Utilities, either the physician's or the patient's, are assigned at the ends of the decision-tree branches to be used by the decision maker.

Sometimes it is convenient to reorder the probability decision tree in a manner called "flipping" because the original tree structure presents data in the form given while the restructured or flipped tree shows processed data in a form appropriate for further analysis.

Construction and use of decision trees for clinical problems are discussed by McNeil and Pauker.

Bayes' Theorem
The use of Bayes' theorem in computer-aided diagnosis is now well known as a result of several hundred papers published on different medical applications of the theorem. It is the Bayesian standpoint applied to medical decision making that I wish to emphasize, a standpoint well illuminated and exemplified in the works of de Finetti, Lindley, Savage, and Cornfield.

The Bayesian standpoint supports the use of subjective probabilities as a degree of belief, the use of Bayes' theorem to modify prior probabilities by likelihood ratios to obtain posterior probabilities, and a consideration of the consequences of incorrect decisions.

The Bayesian standpoint, drawing a contrast between simplified but honest methods of comparing decisions and sophisticated but perplexing methods, is well illuminated by Lindley:

In our everyday decision-making we have developed, because there are no basic rules to guide us, some bad habits. One of them is the tendency to shy away from the simple and take refuge in the complex, where it is not so easy to have one's incompetence exposed. As has been said: 'Practical decision-makers instinctively want to avoid the rather awful clarity that surrounds a really simple decision.' The reply to the accusation of guessing at probabilities and utilities is simply that if you can't do simple problems how can you do complicated ones?

Bayes' theorem may be used with a diagnostic test, or a sequence of diagnostic tests, to calculate the probability that a patient does or does not have a certain diagnosis.

Bayes' theorem and the diagnostic test. If we know the sensitivity and specificity of a diagnostic test and the prior probability of disease, then we can use Bayes' theorem to calculate the probability of disease in a patient with a positive test. Bayes' theorem can be expressed by the following formula,* using the notation of McNeil:

\[
P(D+|T+) = \frac{P(T+|D+)P(D+)}{P(T+|D+)P(D+) + P(T+|D-)P(D-)}
\]

where \( P(D+/T+) \) = conditional probability that a patient has disease, \( D+ \), given a positive test, \( T+ \); \( P(T+/D+) \) = true positive (TP) ratio; fraction of diseased patients who have a positive test result; \( P(D+) \) = prior probability of disease, \( D \); \( P(T+/D-) \) = false positive (FP) ratio; fraction of nondiseased patients who have a positive test result; \( P(D-) \) = prior probability of no disease (not \( D \)).

Bayes' theorem has been used to assess the impact of diagnostic information on opinion in a study of the efficacy of diagnostic radiologic procedures.

Written in odds-likelihood ratio form, Bayes' theorem says that

\[
FO = IO \times LR
\]

*See derivation in the first article in this issue.
where \( I_0 \) is the initial odds of the diagnosis, \( F_0 \) is the final odds of the diagnosis, and \( LR \) is the likelihood ratio or diagnostic usefulness of the data. Technically, \( LR \) is the ratio of the probability of a given test result if the diagnosis is true to the probability of the same test result if the diagnosis is not true.

In this multiplicative form, Bayes' theorem offers an inconvenient definition of diagnostic information impact. A diagnostically useless datum has an \( LR \) of 1; a minimum is zero and maximum is infinity. The asymmetry can be removed by using the additive relationships of Bayes' theorem in logarithmic form. Using \( L \) for logarithm to base 10,

\[
L_{F_0} = L_{I_0} + LLR
\]

or

\[
LLR = L_{F_0} - L_{I_0}
\]

\( LLR \) is the direct measure of the diagnostic efficacy of information from radiographic or nuclear medicine procedures.

The relations among the possible outcomes of a diagnostic test, namely TP, false negative (FN), true negative (TN), and FP may be displayed and studied as a nomogram, a \( 2 \times 2 \) decision matrix, or as an ROC plot.

**Decision Matrices**

More than 4.5 billion medical laboratory tests were performed in the United States in 1975 at an annual cost exceeding $12 billion. Yet physicians usually have very little knowledge of the sensitivity, specificity, and predictive values of the tests they use. Moreover, they have little understanding of how prevalence of the disease for which they are testing influences the predictive value of the laboratory test result. Since these misconceptions influence the physician's diagnostic thinking and therapeutic decisions, it seems useful to review the subject of decision matrices and diagnostic tests in the clinical decision-making context.

With a decision matrix, we can logically relate the results of a diagnostic test to clinical or pathological outcome. This type of analysis is most readily applied to the simple decision of whether disease is present, \( D+ \), or absent, \( D- \), as in screening for disease, when the test is abnormal (positive), \( T+ \), or normal (negative), \( T- \). The presence of disease is always defined with respect to "some other test," and how we define "the truth" of this other test can in itself offer serious problems that we will not discuss here.

A group of matrices are developed in Table 1 to show various relationships of diagnostic outcomes.

The \( 2 \times 2 \) decision matrix was adapted from material in an interesting book entitled *Beyond Normality*, by Galen and Gambino.\(^{17}\) In this example, we have assumed that the prevalence of this disease being tested is 10,000/100,000 = 10\%, the FP rate is 5\%, and the FN rate is 5\%.

The sensitivity of the diagnostic test

\[
(\text{TP}/D+) = P(T+ | D+)
\]

(positivity in disease) =

\[
\frac{A}{A + B} = \frac{9,500}{10,000} = 95\%
\]

The specificity of the diagnostic test

\[
(\text{TN}/D-) = P(T- | D-)
\]

(negativity in health) =

\[
\frac{D}{C + D} = \frac{85,500}{90,000} = 95\%
\]

<table>
<thead>
<tr>
<th>Presence of Disease</th>
<th>Test Result</th>
<th>Positive (( T+ ))</th>
<th>Negative (( T- ))</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(Number with + test)</td>
<td>(Number with - test)</td>
<td></td>
</tr>
<tr>
<td>Present (( D+ ))</td>
<td></td>
<td>( 9,500 )</td>
<td>( 500 )</td>
<td>( 10,000 )</td>
</tr>
<tr>
<td>(number sick)</td>
<td></td>
<td>( 500 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent (( D- ))</td>
<td></td>
<td>( 4,500 )</td>
<td>( 85,500 )</td>
<td>( 90,000 )</td>
</tr>
<tr>
<td>(number not sick)</td>
<td></td>
<td>( 85,500 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>( 14,000 )</td>
<td>( 86,000 )</td>
<td>( 100,000 )</td>
</tr>
</tbody>
</table>
Table 2. Effect of Prevalence on Predictive Value When Sensitivity and Specificity Equal 95%\(^\text{17}\)

<table>
<thead>
<tr>
<th>Prevalence (%)</th>
<th>Predictive Value of a Positive Result (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>2</td>
</tr>
<tr>
<td>1.0</td>
<td>16</td>
</tr>
<tr>
<td>2.0</td>
<td>28</td>
</tr>
<tr>
<td>5.0</td>
<td>50</td>
</tr>
<tr>
<td>10.0</td>
<td>68</td>
</tr>
<tr>
<td>50.0</td>
<td>95</td>
</tr>
</tbody>
</table>

The predictive value of a positive test

\[
P(\text{D+} \mid \text{T+}) = \frac{A}{A + C} = \frac{9,500}{14,000} = 68\%.
\]

This means that 68% of all patients with a positive test result have the disease.

Notice in Table 2 the important effect of prevalence of disease on the predictive value of a positive test. The prevalence rate of a disease equals the number of patients per 100,000 population who have disease at the time of the study. The incidence rate is the number of new cases per 100,000 population per year. (Prevalence = incidence \times duration of disease.)

Since much medical testing is carried out in low disease prevalence situations, it is not surprising that predictive values of positive tests are low—usually unknown to the clinicians.

Few diagnostic tests have simple positive or negative outcomes. Most tests, whether medical laboratory tests, x-ray examinations, or pathology specimens, yield a continuous scale of values, of which one of several can be selected as the decision line to separate the normal subjects from the patients with disease. There is a trade-off depending on relative costs associated with diagnostic errors of FP and FN patient classifications.

Decision matrices alone do not give much insight concerning the cost of these errors of misclassification. However, by relating the matrices to ROC analysis we can analyze outcome trade-offs in cost–benefit terms.

**ROC Analysis and Medical Decision Making**

In the late 1950s, I was studying problems of observer variability in radiology, and I presented an analysis of film reader accuracy as an operating characteristic curve.\(^\text{18}\) At that time, I was not considering a strategy for medical decision making, and it was not until 1967\(^\text{19}\) that I could show the importance of including ROC analysis as a procedure in the medical decision-making sequence that bridges decision matrices and cost–benefit relationships. ROC analysis, in addition, offers a viewpoint from which image analysis appears in greatest simplicity. It is helpful to remember that a node in a decision tree can be represented by a 2 \times 2 decision matrix and has an ROC representation.

I wish to acknowledge the substantial contributions of my colleagues Metz,\(^\text{20,21}\) Goodenough,\(^\text{22,23}\) and Starr\(^\text{24}\) to the theory and application of ROC analysis to visual detection performance in medicine. Recent work by Metz\(^\text{21}\) has shown that ROC analysis can be generalized to cover a range of problems in medical decision making. This follows from the demonstration that the ROC curve describes possible relationships among the probabilities of the various types of correct and incorrect decisions. Therefore, it plays a central role in optimizing diagnostic strategies using the general techniques of decision analysis.

I would like to interest pathologists in ROC analysis. In almost all studies of the efficacy of diagnostic tests or treatment, the tissue pathologist plays an important role as a source of final “truth.” Only a few studies have been done on observer error among pathologists, and these studies show quite clearly that observer disagreement does exist.\(^\text{25,26}\) Pathologists provide answers that are used in assessing the usefulness of certain screening and diagnostic tests. One would hope that ROC analysis would be adopted by pathologists for all tissue diagnosis studies and especially those for cancer screening of breast, lung, cervix, etc. Difficult value judgments concerning the cost–benefit relationships of screening and diagnostic tests might be made with less uncertainty.

To conclude the section on ROC analysis, an operating characteristic curve is used to interpret findings of Loop and Bell\(^\text{27}\) on the efficacy of skull examinations in the evaluation of patients with head injury for the presence of skull fracture. The object of the exercise is to obtain a better understanding of the marginal cost of detecting skull fractures, and to apply similar evaluation methods to nuclear medicine and radiographic procedures.

Loop and Bell\(^\text{27}\) found five high-yield criteria
Table 3. Fracture Yield for Five High-Yield Findings

<table>
<thead>
<tr>
<th>Odds for fracture 50:50 or 9:1 certain</th>
<th>1 Fracture per no. of Examinations</th>
<th>Percent of Fractures Associated with Findings</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/3</td>
<td>67</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Discharge from ear</td>
<td>1/3</td>
<td>30</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Patient stuporous, semiconscious or comatose</td>
<td>1/6</td>
<td>43</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>More than five minutes of unconsciousness</td>
<td>1/7</td>
<td>44</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Eardrum discoloration</td>
<td>1/4</td>
<td>23</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

From Table 3, high-yield findings have selected a group of patients in which the price of fracture detection is about $70 (based on a skull examination cost of $30) because only about 100 patients need to be examined—but in this group more than half the fractures are missed. A selection policy that accepts up to 10 FP examinations for each FN examination will detect 23 fractures in a new group of 133 patients at a cost of $160 for each of these fractures. Similarly a policy that accepts 100 FP for each FN will pick up another 25 fractures in an added group of 350 patients. Finally, for 1000 unnecessary radiographic examinations for each additional fracture found, we detect 6 more fractures in a new group of 600 patients at a cost of $3000 for each fracture. More positive cases are found at an increasingly greater cost per case. At what point on the operating characteristic curve do we prefer to operate—or at what point can we afford to operate?

Cost-Benefit Relationships: Utilities and Maximization of Expected Utility

Explicit formulation of benefit-risk relationships and assignment of utilities to medical action outcomes have become the subject of medical investigations only recently. This may account, in part, for the controversies over the use of utilities in medical decisions. I will point out some of the questions. We must continue to develop this important aspect of medical decision analysis.

There are more than a few critics who believe that values of different types and the relationship of values to costs are incommensurable. Opposed are those who share the view of Hardin29 that "in real life incommensurables are commensurable. All that is needed is a criterion of judgment and a system of weighting. . . . It is when hidden decisions are made explicit that arguments begin. The problem in the
years ahead is to work out an acceptable theory of weighting.”

Obvious difficulties arise when patient utilities and physician utilities must be related. A statement of utilities is to judge every action on the basis of its expected utility. It is the supreme tool for the delegation of authority: give your assistant your utility and tell him to maximize its expected value. By using this approach the patient may delegate authority to the physician in a more meaningful way.

Pauker recently published a study of the choice between coronary bypass surgery and medical therapy in patients with angiographically documented coronary artery disease. A decision tree was constructed and probabilities assigned to all outcomes. Patients were debriefed on their preferences for three possible outcomes: immediate death, 5 yr of life with severe pain, and 2.5 yr of life without pain. With some further questioning a point of indifference can be found and a consistent utility assigned to each outcome. Probability of outcome and the patient’s utilities are then combined to predict the choice of treatment that has maximum expected utility for the patient.

Pauker emphasizes that using decision analysis in clinical problems combines medical expertise and patient preferences in an explicit method of clinical decision making. The use of a decision tree and an examination of patient preferences facilitates the education of a patient as to his options and their consequences. But medical decisions that include patient utilities may raise questions of quality of care and cost containment in which society has a stake—decisions in which ethics and decision analysis are intertwined. Thus we return to a consideration of topics which opened the article.

REFERENCES

and localization of radiographic images. Radiology 116:533, 1975


