

Introduction to Clinical Decision Making

Dennis D. Patton

In the last few years there has been a remarkable increase in the amount of clinical data in the average hospital chart, and more and more problem-solving algorithms have been developed. We need better "thinking tools" to help us handle the flow of information. The term "clinical decision making" is used to describe a systematic way to handle data and algorithms to decide on a best course of action. This in-

CLINICAL DECISION MAKING is certainly no new idea—it's been around since the dawn of medicine. Physicians have always reasoned, either verbally or nonverbally: "If I do this, such and such will probably happen, but if I do that so and so will probably happen." They then would choose the course that seems best to them.

The term "clinical decision making" as used today connotes something more concrete: a systematic way to handle data and algorithms to decide on a best course of action. A systematic approach has become more and more necessary as we gather more data about patients and develop better problem-solving algorithms. (An algorithm is merely a step-by-step procedure for solving a problem.) The concepts used in clinical decision making do not require the use of a computer, though computers not only save time but greatly expand the range of applications and the power of the technique. Computers are in fairly widespread use in nuclear medicine and are used mainly for handling data. It is likely that we will see computers used more and more to handle problem-solving algorithms as well.

A systematic approach to deciding on the best course of action might go something like this: (1) List all possible actions, (2) list all possible outcomes, (3) predict the probability of each outcome from each action, and then (4) select the best action based on outcome likelihood and outcome utility.

The term "actions" could refer to deciding whether to do a test or not to do it, deciding which test to do, deciding how to interpret the results, and deciding on patient management. Actions are not limited to the physician—the patient can and should participate in the decision-making process. Chance may also play a role. The patient may turn out to have a certain disease or not have it; a diagnostic test may

cause adverse results or it may not. (Decision trees, which help analyze the effect of various actions and chance events, are discussed in the article by Pauker and Kassirer in this issue.¹)

troductory article discusses some of the problems in establishing a decision criterion, both for a population and for an individual patient. Comparing the probabilities and utilities of various diagnostic outcomes (true positive, false positive, etc.) leads to a diagnostic strategy. The article also discusses conditional probability, Bayes' theorem, and likelihood ratios.

The term "outcome" refers to the results of a course of action, and the interpretation of the term depends on the situation. In deciding whether to do a test or not do it, or in deciding which test to do, possible outcomes are measured in terms of cost, time, patient discomfort, risk, morbidity, and mortality. In deciding how to interpret a test, the outcome is measured in terms of whether or not the interpretation was correct, and the probability and penalty of being wrong.

In the past few years there have been several books and articles on general principles in clinical decision making. These include books by Lusted,² Feinstein,³ and Murphy.⁴ Articles have been written by McNeil et al,⁵ Kassirer,⁶ Pauker and Kassirer,⁷ Burke,⁸ and Patton.⁹ Specific problems in clinical decision making are discussed in articles on the value of screening tests, by McNeil,¹⁰ the value of a normal finding, by Gorry,¹¹ and one on potential misuse of additional data, by Sisson.¹² In a book on the subject, Galen¹³ discusses the importance of disease prevalence on clinical decision making.

Diagnostic algorithms have been discussed in a wide variety of clinical applications. To reference only a few: electrocardiograms,¹⁴ cystourethrograms,¹⁵ the neurologic work up,¹⁶ FUD, lymphangiography,¹⁸ and bone tumors.¹⁹ Diagnostic algorithms in nuclear medicine studies have been described for lung

From the Health Science Center of the University of Arizona, Tucson, Ariz.

Reprint requests should be addressed to Dennis D. Patton, M.D., University of Arizona, Health Science Center, Division of Nuclear Medicine, Tucson, Ariz. 85724.

© 1978 by Grune & Stratton, Inc.

0001-2998/78/0804-0009\$02.00/0

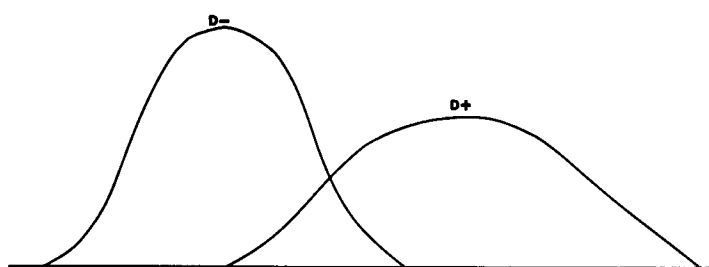


Fig. 1. Results of a hypothetical test (T) in a population with disease ($D+$) and in a normal population ($D-$). Test results overlap as the test is not perfect. Horizontal axis: test result. Vertical axis: number of individuals with given test result.

scans by McNeil,²⁰ for the evaluation of thyroid nodules by Sisson,²¹ and for liver scans by Drum.²² These lists are by no means complete, and many other references are given in the following articles.

THE DECISION CRITERION

In using a test to distinguish normal from abnormal there are two general situations: establishing a diagnostic strategy for a population, and making a decision about an individual patient. The strategies may be quite different.

Decision Strategy for a Population

If a test is to be used to distinguish normals from patients with disease (D), some definite criterion will be needed to classify test results. In the simplest case there are only two possible test outcomes, normal and abnormal, and there are two populations, normal and patients with disease (D). The distributions of the populations may or may not be Gaussian but they should be known. Figure 1 shows a hypothetical situation in which $D+$ represents people with disease, and $D-$ represents people without it. The two populations almost inevitably overlap with respect to test results; hardly any test is perfect.

Where should the line be drawn to separate what one is going to call "normal" from what

one is going to call "abnormal"? A logical first step might be to put the line through the point where the two curves cross (Fig. 2). Then every test result to the right of this criterion line, X_c , will be called "positive" and every result to the left of it "negative." All people with disease, D , whose test results are higher than X_c will be correctly classified as having D and will be true positives (TP). All people without D whose test results are lower than X_c will be true negatives (TN; Fig. 3). Those patients with D whose test results happen to be lower than X_c will be misclassified and will be false negatives (FN), while those people without D whose test results happen to be higher than X_c will be false positives (FP; Fig. 4).

This location of X_c leads to approximately as many FNs as FPs. This is acceptable if the penalty for being wrong is the same in either case, i.e., if it is just as bad to overcall a noncase as to miss a case. Whether the penalties are symmetric depends on what happens as a result of the test. If missing a case is infinitely worse than overcalling a noncase, the strategist will move his criterion line X farther to the left, so as to miss fewer cases. The problem of FN can be completely eliminated easily enough—just call all patients "positive" (or even just those patients whose test results are higher than those of the case having the lowest test result). Figure

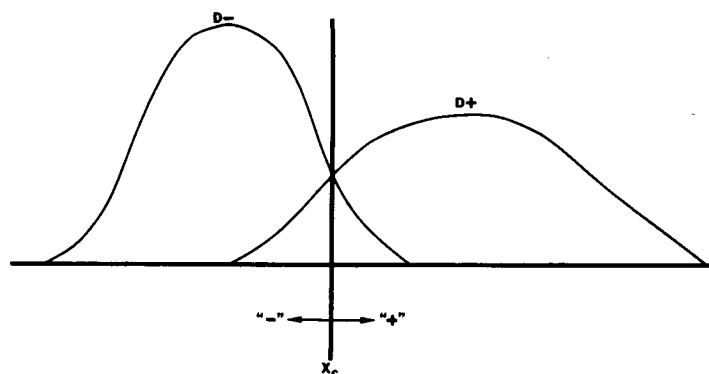


Fig. 2. Criterion X_c is drawn through the point where the two curves overlap. Test results to the right of X_c will be called positive ($T+$), and to the left negative ($T-$).

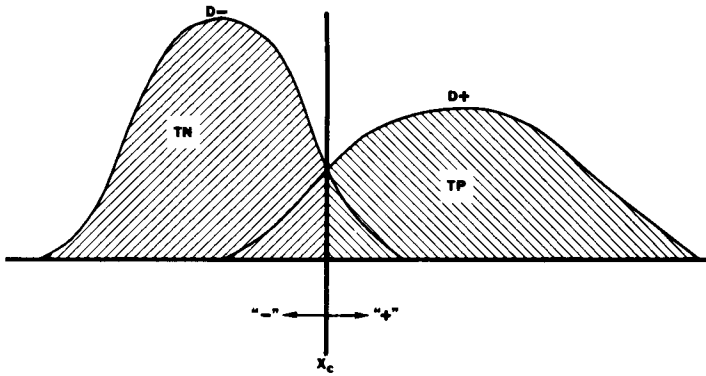


Fig. 3. Persons with disease ($D+$) whose results are to the right of X_c will be TP, and persons without disease whose results are to the left of X_c will be TN.

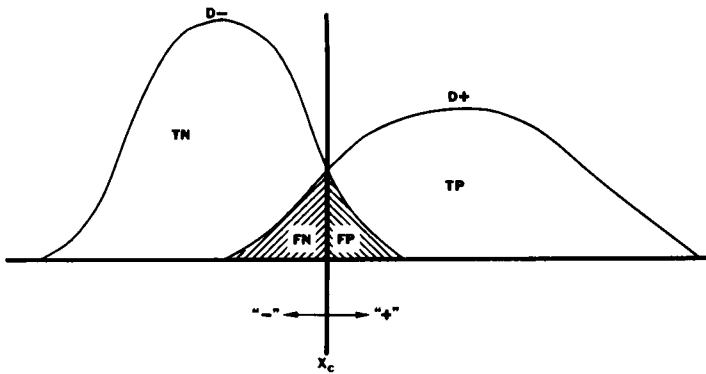


Fig. 4. Persons with disease ($D+$) whose test results happen to be to the left of X_c will be FN, and normals whose results are to the right of X_c will be FP.

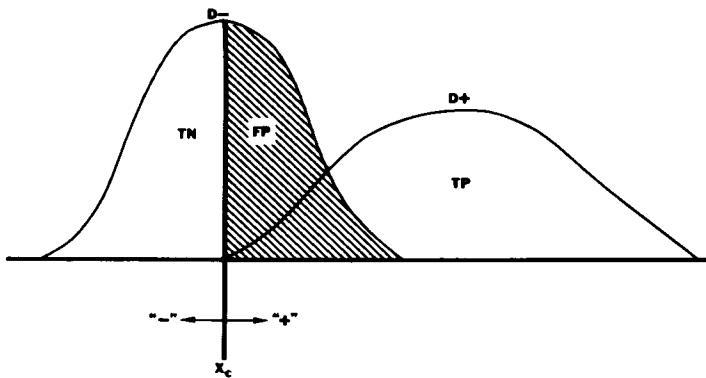


Fig. 5. FN have been eliminated by moving X_c to the lower end of the $D+$ curve. As a result of eliminating FN, FP have been greatly increased.

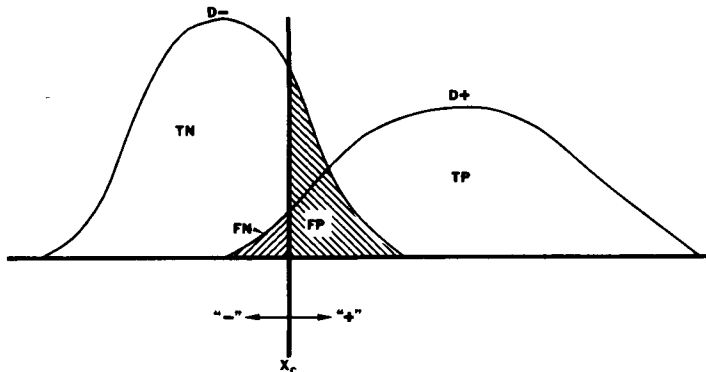


Fig. 6. A compromise. Criterion line X_c has been placed so as to accept a small number of FN in exchange for a manageable number of FP.

5 shows this latter strategy. The criterion line has eliminated all FN results. But look what has happened to the FP—half of the normal population will be called positive under this strategy. This would be unacceptable in almost any real situation, in which a positive test result leads to some further test or action incurring cost or risk. The penalty for being wrong with this strategy is too many FPs. To balance the penalties for FP and FN the strategist will move his criterion line back to the right (Fig. 6) until the penalty for FP times the frequency of this outcome balances the penalty for FN times the frequency of this outcome.

The penalties depend on a number of tangible and intangible factors, including cost, risk, morbidity, mortality, radiation dose, discomfort, etc. Comparing the penalty of the two types of incorrect outcomes is at least partly subjective, and different strategists will assign different penalties; as a result, different physicians will assign different values of X_c to determine what they as individual doctors can live with. The question of utility of outcomes is discussed at greater length in the articles by Bell²³ and by Pauker and Kassirer¹ in this issue.

Note that in Fig. 6 if the prevalence of D were higher, TP and FN would both increase if X_c were left constant.

Decision Strategy for the Individual Patient

It should be stressed that the strategy by which one assigns a criterion line (Figs. 1-6) is one that applies to a population, not to an individual patient. The physician who has to make clinical decisions will find the criterion line that best fits his own personal preferences. What does he do then with a patient who has test result X_1 (Fig. 7)? This test result is higher than the criterion line, so the patient is classified as being positive. What are the chances that he actually is? By examining the two curves at the point X_1 the physician can see that there is actually more chance that the patient is normal ($D-$) than there is that he is abnormal ($D+$). Then why call him abnormal? Because of the penalty for being wrong if he is called a noncase but turns out to be a missed case. The patient may be told that there is a 67% probability (or 2:1 odds) that he is normal, but that prudent regard for his welfare implies that he should carry a working diagnosis of D until proved

otherwise. Similarly, the patient with test result X_2 has a 50% probability (1:1 odds) of being normal, and the patient with test result X_3 has a 33% probability (1:2 odds) of being normal.

Thus, the same curves can be used to determine a strategy for classifying populations and to find the probability (or odds) for an individual. It should be stressed that this analysis applies to any test-disease combination for which the normal and abnormal test distributions are known. Actually, one could discard the criterion line concept and work with probabilities alone, but modern medicine demands classification, and few clinicians would be content to work just with probabilities.

Outcomes can be compared with test results in a 2×2 table (simplest case):

		Final Diagnosis	
		D+	D-
Test result	T+	TP	FP
	T-	FN	TN

Sensitivity is the fraction of cases that the test calls positive:

		Final Diagnosis	
		D+	D-
Test result	T+	TP	FP
	T-	FN	TN
		$\frac{TP}{TP + FN}$	

$$\text{Sensitivity} = \frac{\text{TP}}{\text{actual cases}} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Specificity is the fraction of noncases that the test calls noncases (i.e., normal):

		Final Diagnosis	
		D+	D-
Test result	T+	TP	FP
	T-	FN	TN
		$\frac{\text{TN}}{\text{FP} + \text{TN}}$	

$$\text{Specificity} = \frac{\text{TN}}{\text{actual noncases}} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

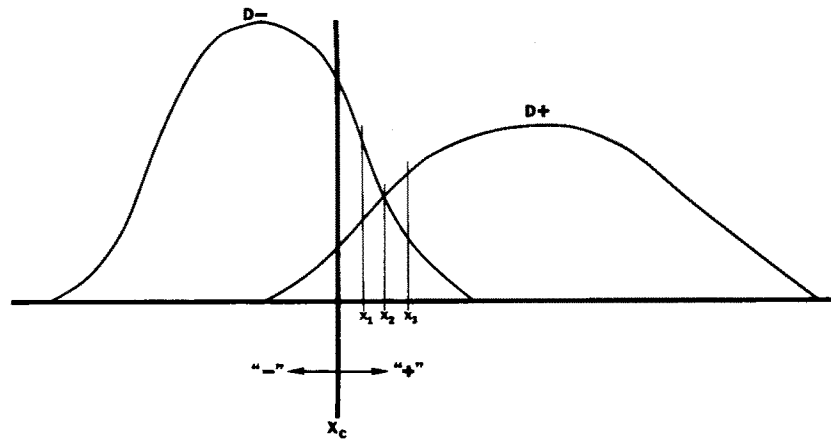


Fig. 7. An individual patient with test result X_1 will be called positive because his result is to the right of X_c , even though the odds are 2:1 that he is actually normal. Patient with X_2 has 1:1 odds of being normal. Patient with test result X_3 has 2:1 odds of having the disease.

The predictive value of a “positive” test result is the fraction of all “positive” reports that turn out to be cases:

		Final Diagnosis		
		D+	D-	Sum
Test result	T+	TP	FP	TP + FP
	T-	FN	TN	

Predictive Value

$$= \frac{TP}{\text{all positive reports}} = \frac{TP}{TP + FP}$$

Similarly, the predictive value of a negative report is $TN/(TN + FN)$.

Accuracy is defined as the sum of all correct outcomes divided by the total number of tests done:

Accuracy

$$= \frac{TP + TN}{\text{all tests}} = \frac{TP + TN}{TP + TN + FP + FN}$$

CONDITIONAL PROBABILITY

Before going on to Bayes’ theorem we need to say a word about conditional probability. The probability of a given situation may depend on some condition; for example, the probability of drawing a spade from a deck of cards depends on whether cards have already been drawn from it. The condition is indicated after a short vertical line: $P(\text{spade} | \text{full deck}) = 13/52$, but $P(\text{spade} | 1 \text{ club already removed}) = 13/51$. In clinical decision making there are many such

conditions. Say a patient has fever and abdominal pain; the probability that he has appendicitis | fever, pain, appendix = 0.1 but $P(\text{appendicitis} | \text{fever, pain, no appendix}) = 0$. Conditional probability is written in the form $P(X | Y)$ and translated to “the probability that X is true, given the condition that Y is true.”

There are two important relationships in conditional probability that the reader should understand before going further. Patients with disease D either have a positive test ($T+$) or a negative test ($T-$) (assuming that only these test outcomes are possible). Then, given $D+$,

$$P(T+ | D+) + P(T- | D+) = 1 \tag{1}$$

which is another way of saying the preceding sentence (using the notation of McNeil⁶). In other words, if we list all the possibilities for a given condition (in this case $D+$), they should add up to 1, or certainty. This hypothetical test has only two allowable (and mutually exclusive) outcomes, $T+$ and $T-$; other situations may have a greater number of possible outcomes, but the probabilities must all add up to 1.

Similarly, $P(T+ | D+)$ and $P(T+ | D-)$ represent all patients with a positive test result; they are either TP or FP, respectively. There are no other possibilities. So,

$$P(T+) = \frac{P(D+)P(T+ | D+) + P(D-)P(T+ | D-)}{P(T+ | D-)} \tag{2}$$

the probability of getting a positive test result regardless of whether there is disease or not. The two terms in equation 2 show that the probability of getting a positive test result equals the probability of having disease, $P(D+)$, times the probability of being TP, $P(T+ | D+)$,

plus the probability of not having disease, $P(D-)$, times the probability of being FP, $P(T+ | D-)$.

BAYES' THEOREM

One of the most important concepts in modern clinical decision making was developed by Thomas Bayes (1702–1761), an English mathematician/theologian. His work was published posthumously in 1763. Bayes' theorem gives a method for recalculating a probability based on new evidence. The old (prior) probability must be known, as well as some other data regarding the new evidence. In its simplest medical form Bayes' theorem is given as the expression

$$P(D|S) = \frac{P(D)P(S|D)}{P(S)} \quad (3)$$

where $P(D|S)$ = the probability of disease D given symptom S ; $P(D)$ is the prior probability of disease D (i.e., the probability of D before the evidence regarding S was added); $P(S|D)$ is the probability of symptom S given disease D (i.e., the probability that a patient will have S if he has D); and $P(S)$ is the probability of finding symptom S in a patient regardless of whether he has disease D or not. For example, suppose hyperthyroidism (D) occurs in 5% of the population ($P(D) = 0.05$), and that insomnia (S) is found in 10% of the general population ($P(S) = 0.1$) but in 30% of the hyperthyroids ($P(S|D) = 0.3$). Then $P(D|S)$, the probability that a patient has hyperthyroidism given that he has insomnia, is $(0.05 \times 0.3)/0.1 = 0.15$, or 15%. In other words, the new evidence (S) makes it three times as likely that he has hyperthyroidism than it was before we knew S . Note that S can represent not only a symptom but a sign, laboratory test result, or any piece of information about the patient.

Another way of looking at Bayes' theorem is to consider $P(S|D)/P(S)$ as a factor that modifies $P(D)$ so as to take into account the effect of new information (S):

$$P(D|S) = P(D) \times \frac{P(S|D)}{P(S)} \quad (4)$$

If $P(S|D)/P(S)$ is greater than 1, the new

probability $P(D|S)$ will be greater than the old probability $P(D)$. This will happen if S is found in more patients with D than it is in the general population, in other words if there is an association between S and D . On the other hand if $P(S|D)/P(S)$ is less than 1, the new probability $P(D|S)$ will be less than the old probability $P(D)$, which is hardly surprising because if $P(S|D)/P(S)$ is less than 1, it means that S is found less often in patients with D than it is in the general population, and there is a disassociation between S and D ; finding S would then make D less likely, which Bayes' theorem bears out quantitatively.

Using the standard notation in this issue, Bayes' theorem would be written,

$$P(D+ | T+) = \frac{P(D+)P(T+|D+)}{P(T+)} \quad (5)$$

$$= P(D+) \times \frac{P(T+|D+)}{P(T+)} \quad (6)$$

This notation will be used throughout this issue.

It should be stressed that Bayes' theorem merely recalculates an old probability based on new evidence; it does not magically create a probability out of raw data. The new probability obviously depends completely on the old probability, and if the old is wrong, the new will be wrong too.*

Bayes' theorem can be written in a number of forms. Note that the denominator in equation 5, $P(T+)$, can be written $P(D+)P(T+|D+) + P(D-)P(T+|D-)$, i.e., all patients with $T+$ either have the disease, $P(D+)$, and are TP, $P(T+|D+)$, or they do not have the disease, $P(D-)$, and are FP, $P(T+|D-)$. Then Bayes' theorem can be written.

$$P(D+ | T+) = \frac{P(D+)P(T+|D+)}{P(D+)P(T+|D+) + P(D-)P(T+|D-)} \quad (7)$$

Note also that $P(D-)$ can be written $1 - P(D+)$.

Bayes' theorem can be written in simple form if the numbers of TP and FP are known:

*Probably.

$$P(D+|T+) = \frac{TP}{TP + FP} \quad (8)$$

Note that this expression is the same as predictive value. Similarly, if the TP fraction [TPF = TP/(TP+FN)] and FP fraction [FPF = FP/(FP+TN)] are known, Bayes' theorem can be written.

$$P(D+|T+) = \frac{P(D+) TPF}{P(D+) TPF + (1 - P(D+)) FPF} \quad (9)$$

Be sure not to confuse TP (units: number of patients) with TPF (no units). The reader can show that equations 8 and 9 are equivalent by substituting the definitions of TPF and FPF in equation 9 and noting that TP + FN = $N \times P(D+)$, where N is the total number of patients ($N = TP + FP + TN + FN$), and that $TN + FP = N \times (1 - P(D+))$.

With reference to Fig. 6, Bayes' theorem can be written in integral form if the curves $D-$ and $D+$ are known analytically:

$$P(D+|T+) = \frac{\int_{x_c}^{\infty} [D-] dx}{\int_{x_c}^{\infty} [D+] dx + \int_{x_c}^{\infty} [D-] dx} \quad (10)$$

where $[D+]$ and $[D-]$ are the analytical forms of the curves $D+$ and $D-$.

Another form uses the likelihood ratio, LR , which is the ratio of the probability of a given test result in the presence of disease to the probability of the same test result in the absence of disease:

$$LR = \frac{P(T+|D+)}{P(T+|D-)} \quad (11)$$

The likelihood ratio is the ratio of the TPF to the FPF. (It assumes the FPF is not zero.) If we divide each term in equation 7 by $P(T+ | D-)$, we get

$$P(D+|T+) = \frac{P(D+) LR}{P(D+) LR + 1 - P(D+)} \quad (12)$$

or,

$$P(D+|T+) = \frac{1}{1 + \frac{1 - P(D+)}{P(D+) LR}} \quad (13)$$

The reader should satisfy himself that all these forms of Bayes' theorem are equivalent.

Bayes' theorem has found its way into the medical literature; a 1977 review by Mai and Hachman²⁴ described 73 clinical applications.

ASSESSING PHYSICIANS' PERFORMANCE

A number of studies have been done to find out whether physicians use sound clinical decision-making principles in actual practice. As it turns out, many physicians do, though there is much room for improvement. The American College of Radiology began an efficacy study in 1971, collecting nearly 9000 diagnostic studies (the 7 most frequently used x-ray procedures in hospital emergency services).²⁵ Some of the findings of this study are discussed in the article by Bell in this issue. Several articles on assessing physicians' performance in decision-making tasks have appeared in the recent literature. Physicians' use of laboratory tests was described by Skendzel,²⁶ a barium enema efficacy study was discussed by MacEwan,²⁷ and Berlin²⁸ discussed the missed radiographic diagnosis.

As our clinical decision-making tools become more refined and as knowledge of them becomes more widespread, we can look forward to commonplace—and hopefully standard—use of these techniques in the practice of medicine.

GLOSSARY

Accuracy. The fraction of test results that are correct; $(TP+TN)/\text{All results} = (TP+TN)/(TP+FP+TN+FN)$.

Actual diagnosis. The diagnosis that is most likely according to the reference test (usually some pathologic test). See *reference test*.

Algorithm. A step-by-step procedure for solving a problem. The algorithm for finding the mean of a group of numbers might be: find the sum of all the numbers; call it Q ; find the number of entries, call it N ; divide Q by N .

A posteriori. Retrospective reasoning; classically, from effect to cause (inductive reasoning), but as used in clinical decision making, the best explanation of all findings based on analysis of all known data. An a posteriori diagnosis is the diagnosis that is finally reached when all clinical and laboratory data have been evaluated.

A priori. Prospective reasoning; classically, from cause to effect (deductive reasoning), but as used in clinical decision making, the best estimate of the likelihood of disease prior to testing. An a priori diagnosis is the diagnosis based on prevalence of disease in the population or on clinical evaluation prior to a diagnostic test. See *prior probability*.

Bandwidth. Range of spatial frequencies (q.v.) included in an image.

Bayes' theorem. A technique for recalculating an old

(prior) probability, based on new evidence. See discussion in this article.

Conditional probability (CP). The probability that something is true given that something else is true. The CP of drawing a spade from a deck of cards given that the 4 of spades has already been removed is 12/51. (Without the condition it would have been 13/52.) See discussion in this article.

Contrast. Difference in density (counts, blackness, concentration, etc.) between an area of interest and its immediate surrounding. In nuclear medicine, object contrast refers to differences in amounts of radioisotope in the object being imaged, while image contrast refers to differences in count density in the image.

Contrast gradient. Change in density (i.e., contrast) per unit length. At the edges of an image, a high contrast gradient gives the effect of edge sharpness.

Contrast threshold. The level of contrast necessary for an observer to just barely detect some feature of an image.

Cost-benefit ratio. A hypothetical (at present) ratio of the negative features of a course of action to the positive features. Negative factors are financial cost, patient discomfort, morbidity, mortality, time, effort, etc. Positive factors are reduction of morbidity or mortality, establishing a firm diagnosis or course of management, etc. Improving the cost-benefit ratio means making it lower.

Count-limited image. An image in which the number of counts is so low that the statistical fluctuation in the count rate overshadows any count rate changes that may be due to important features of the image. In terms of contrast, a count-limited image is one in which statistical fluctuation due to low counts has reduced the image contrast to the contrast threshold or below.

Data. Facts or figures from which conclusions can be inferred. The term data is sometimes used to mean facts not necessarily in readily intelligible form (such as the ones and zeros in binary computer language), in contrast to information, which suggests facts in ordinary language. This distinction is used irregularly.

Data base. The set of data used to form a conclusion; the set of facts known about a patient.

Decision matrix. A table of all possible outcomes of a test. See discussion in this article.

Decision tree. A figure showing all possible outcomes of all possible courses of action and chance events, used to evaluate the utility of various outcomes so as to plan the best course of action. See articles in this issue by Pauker and Kassirer,¹ and by Lusted.²⁹

Density. Concentration per unit area. In radiographs, film density is the logarithm (to the base 10) of the ratio of the incident intensity to the transmitted intensity. In scintigraphs, count density is the number of counts per unit area. In a transparency of a scintigram one can speak of both film density and count density.

Diagnosis. A term representing a recognizable disease entity; the process of arriving at an understanding of a disease process. A few diagnoses are mere translations: "erythematous dermatitis" carries no more information than "red, inflamed skin."

Diagnostic certainty. That level of confidence at which the referring physician is willing to stop further testing. See article in this issue by Bell.²³

Diagnostic efficacy. Usefulness of a diagnostic

procedure in terms of the impact it has on the diagnostic thinking of the referring physician. See article in this issue by Bell.²³

Efficacy. Power to produce intended results. In a diagnostic test, the intended result is to establish a diagnosis or to rule it out, and a test is efficacious to the extent that it does this. See article by Bell, in this issue.²³

False negative. A test result which is negative despite the fact that the patient actually has disease.

False negative fraction. The fraction of all patients having disease in whom test results are negative.

False positive. A test result that is positive despite the fact that the patient actually does not have disease.

False positive fraction. The fraction of all patients not having disease in whom test results are positive. $FPF = [FP/(FP + TN)] = 1 - \text{specificity (q.v.)}$.

Gold standard. See *reference test*.

Image contrast. See *contrast*.

Incidence. Number of new cases of a given disease per 100,000 population per year. See *prevalence*.

Information. See *data*.

Interobserver disagreement. The degree to which two qualified observers, examining the same data, come to disparate conclusions. An observer may disagree with himself upon reviewing previously analyzed data without knowing he has seen it before.

Intransitive. In analyzing preferences, describes a system in which a person could prefer A to B, could prefer B to C, and could prefer C to A. Since many subjective preferences are intransitive, it is difficult to deduce a preferred course of action by considering only a small number of such trade-off preferences.

Just noticeable density difference. See *contrast threshold*.

Likelihood ratio. Probability of a given test result in patients with disease divided by the probability of the same test result in patients without disease. $LR = P(T+ / D+) / P(T+ / D-)$, where $P(T+ / D+)$ = probability that test is positive in the presence of disease (true positive), and $P(T+ / D-)$ = probability that test is positive in the absence of disease (false positive). See *log likelihood ratio*.

Line spread function. The image of an infinitesimally thin line by an imperfect imaging system; the activity appears to be spread around the line in a quasi-Gaussian distribution.

Log likelihood ratio. The logarithm (to the base 10) of the likelihood ratio. The likelihood ratio can vary from zero to infinity, the extremes being most useful in terms of diagnostic efficacy. An LR of 1 is diagnostically useless (see article by Lusted in this issue).²⁹ The log likelihood ratio varies between small negative numbers and small positive numbers, and is symmetric about $LLR = 0$, a value that suggests uselessness.

Management efficacy. The extent to which a diagnostic procedure changes patient management. See article in this issue by Bell.²³

Modulation transfer function (MTF). Description of the ability of an imaging system to reproduce different spatial frequencies (q.v.). The MTF is usually shown as a graph in which system response is plotted against spatial frequency.

Noise. In imaging, features of an image that do not arise from signal features in the object, but rather from ex-

traneous sources (background activity in the object, background activity in the room, spurious counts arising in the electronics, etc.).

Normal. In most numerical diagnostic tests, that range of values lying within two standard deviations of the mean of the undiseased population. Normal range is usually between mean -2 SD and mean $+2$ SD.

Object contrast. See *contrast*.

Odds. Ratio of the chances of one event to the chances of some other event. The events should be mutually exclusive. The odds of drawing a spade from a complete deck are 13:39 or 1:3 (i.e., ratio of spades to nonspades). The odds of *not* drawing a spade are 39:13, or 3:1. Odds and probability are related: if the odds of X versus Y are $X:Y$, the probability of X is $X/(X + Y)$.

Outcome efficacy. Usefulness of a course of action in optimizing the final outcome. See article by Bell in this issue.²³

Pixel. The smallest resolvable element of an image. For most gamma camera scintigrams, a pixel is roughly 1 sq cm of object space.

Posterior probability (posttest probability). The probability of disease after a given test result is known. See *a posteriori*.

Predictive value. Probability that a patient has disease, given that the test result is positive: $PV = P(D+ | T+)$. The predictive value of a negative test is the probability that the patient is free from disease, given that the test result is negative: $PV (\text{neg.}) = P(D- | T-)$. See discussion in this article.

Pretest probability. The probability of disease estimated just before a given diagnostic test is done. See *a priori*.

Prevalence. The number of cases of a given disease per 100,000 population. Prevalence = incidence \times duration of disease (in years). See *incidence*.

Prior probability. See *pretest probability*. Prior probability is often written $P0$ or $P(D)$ or $P(D+)$. It is the prevalence (q.v.) of the disease in the population being tested.

Probability. The chance that a given event will occur. Probability is written either as a decimal number between 0 (impossible) and 1 (certain), or as a percent between 0% and 100%. See *odds*.

Receiver operating characteristic curve (ROC curve). The relationship between true positive and false positive responses from an observer interpreting a number of images. Note from Figs. 1-6 in this article that there is a direct relationship between TP and FP, and that more TP can be gained only at the cost of more FP. See articles by Metz²⁰ and by Lusted²⁹ in this issue.

Reference test. A diagnostic test that is agreed upon to establish the ultimate and final diagnosis, and against which all other tests are to be validated. The reference test in many cases is autopsy or surgical pathology. The fallibility

of reference tests in some situations makes validation of other tests difficult.

Resolution. In an image, the closest that two point sources can be and yet be discerned as two sources.

Screening test. A test performed to pick up as many cases as possible out of the population being tested; it should have high sensitivity (q.v.).

Sensitivity. Ratio of true positives to all patients with disease; fraction of diseased patients diagnosed as such: $TP/(TP + FN)$. See discussion in this article.

Sensitivity analysis. Analysis of a decision tree to see to what extent small changes in assigned values of probability or utility will significantly alter a decision. See article by Pauker and Kassirer in this issue.¹

Sharpness. See *contrast gradient*.

Signal-to-noise ratio (SNR). A measure of the quality of an image in terms of its ability to convey significant data to the observer. If noise in a scintigram consists of random background activity, the $SNR = S/\sqrt{B}$, where S = signal counts/pixel and B = background counts/pixel.

Spatial frequency. A means of representing the size of an object in terms of a frequency in cycles per unit length. The object is understood to occupy one-half of a full cycle, so the spatial frequency of an object X cm in diameter is $1/2 X$ cycles/cm. Smaller objects have higher spatial frequencies. This concept is essential to calculation of modulation transfer function (q.v.).

Specificity. Ratio of true negatives to all patients without disease; fraction of nondiseased patients diagnosed correctly: $TN/(TN + FP)$. See discussion in this article.

Survival. Length of time from diagnosis of a disease to death. Early diagnosis may appear to lengthen survival even if nothing is done to alter the course of the disease, unless correction is made for the increased lead time.

Threshold probability. A probability level at which a decision maker should consider two courses of action equivalent with respect to expected benefit for the patient. See article by Pauker and Kassirer in this issue.¹

Transitive. In analyzing preferences, describes a system in which if a person prefers A to B and prefers B to C, he would necessarily prefer A to C. See *intransitive*.

True negative. A test result that is negative in a patient who does not have disease.

True negative fraction (TNF). Fraction of all patients not having disease in whom test result is negative. $TNF = TN/(TN + FP) = \text{specificity (q.v.)}$.

True positive. A test result which is positive in a patient with disease.

True positive fraction (TPF). Fraction of all patients having disease in whom test result is positive. $TPF = TP/(TP + FN) = \text{sensitivity (q.v.)}$.

Utility. The relative desirability of a given outcome; a subjective assessment of worth that may be made on the basis of morbidity and mortality and may also include input from the patient or his family. See articles in this issue by Pauker and Kassirer,¹ Lusted,²⁹ McNeil,³¹ and Drum.³²

REFERENCES

1. Pauker SG, Kassirer JP: Clinical application of decision analysis: A detailed illustration. *Semin Nucl Med* 8:324-335, 1978
2. Lusted LB: Introduction to medical decision making. Springfield, Charles C Thomas, 1968
3. Feinstein AR: Clinical judgment. Baltimore, Williams & Wilkins, 1967
4. Murphy EA: The logic of medicine. Baltimore, Johns Hopkins Univ Pr, 1976
5. McNeil BJ, Keeler E, Adelstein SJ: Primer on certain

- elements of medical decision making. *N Engl J Med* 293:211-215, 1975
6. Kassirer JP: The principles of clinical decision making: An introduction to decision analysis. *Yale J Biol Med* 49:149-164, 1976
 7. Pauker SG, Kassirer JP: Therapeutic decision making: A cost benefit analysis. *N Engl J Med* 293:229-234, 1975
 8. Burke MD: Diagnosis: Art or science? *Diagn Med* 1:25-32, 1978
 9. Patton DD: A radiologist looks at clinical decision making. *Appl Radiol* 5:119-123, 1976
 10. McNeil BJ, Adelstein SJ: Determining the value of diagnostic and screening tests. *J Nucl Med* 17:439-448, 1976
 11. Gorry GA, Pauker SG, Schwartz WB: The diagnostic importance of the normal finding. *N Engl J Med* 298:486-489, 1978
 12. Sisson JC, Schoemaker EB, Ross JC: Clinical decision analysis: The hazard of using additional data. *JAMA* 236:1259-1263, 1976
 13. Galen RS, Gambino SR: *Beyond Normality: The Predictive Value and Efficiency of Medical Diagnosis*. New York, Wiley, 1975
 14. Rifkin RD, Hood WB: Bayesian analysis of electrocardiographic exercise stress testing. *N Engl J Med* 297:681-686, 1977
 15. Saenger EL, Kereiakes JG, Cavanagh DJ, et al: Cystourethrography procedures in children. *Radiology* 118:123-128, 1976
 16. Salamon R, Bernadet M, Samson M, et al: Bayesian method applied to decision making in neurology—Methodological considerations. *Methods Inform Med* 15:174-179, 1976
 17. Vickery DM, Quinell RK: Fever of unknown origin: An algorithmic approach. *JAMA* 238:2183-2186, 1977
 18. Safran C, Desforges JF, Tschlis PN, et al: Decision analysis to evaluate lymphangiography in the management of patients with Hodgkin's disease. *N Engl J Med* 296:1088-1092, 1977
 19. Lodwick GS: A probabilistic approach to the diagnosis of bone tumors. *Radiol Clin North Am* 3:487-497, 1965
 20. McNeil BJ: A diagnostic strategy using ventilation-perfusion studies in patients suspect for pulmonary embolism. *J Nucl Med* 17:613-616, 1976
 21. Sisson JC, Bartold SP, Bartold SL: The dilemma of the solitary thyroid nodule: Resolution through decision analysis. *Semin Nucl Med* 8:59-71, 1978
 22. Drum DE, Christacopoulos JS: Hepatic scintigraphy in clinical decision making. *J Nucl Med* 13:908-915, 1972
 23. Bell: Efficacy . . . What's That?? *Semin Nucl Med* 8:316-323, 1978
 24. Mai N, Hachman E: Anwendung des Bayes-Theorems in der medizinischen Diagnostik—Ein Literaturübersicht. *Metamed* 1:161-205, 1977
 25. Lusted LB: Efficacy of diagnostic x-ray examination (Letter to the editor). *JAMA* 239:109, 1978
 26. Skendzel LP: How physicians use laboratory tests. *JAMA* 239:1077-1080, 1978
 27. MacEwan DW, Kavanagh S, Chow BA, et al: Manitoba barium enema efficacy study. *Radiology* 126:39-44, 1978
 28. Berlin L: Does the "missed" radiographic diagnosis constitute malpractice? *Radiology* 123:523-527, 1977
 29. Lusted LB: General problems in medical decision making with comments on ROC analysis. *Semin Nucl Med* 8:299-306, 1978
 30. Metz CE: Basic principles of ROC analysis. *Semin Nucl Med* 8:283-298, 1978
 31. McNeil BJ: Rationale for the use of bone scans in selected metastatic and primary bone tumors. *Semin Nucl Med* 8:336-345, 1978
 32. Drum DE: Optimizing the clinical value of hepatic scintiphography. *Semin Nucl Med* 8:346-357, 1978