Application of Information Theory to Clinical Diagnostic Testing

The Electrocardiographic Stress Test

GEORGE A. DIAMOND, M.D., MICHAEL HIRSCH, M.S., JAMES S. FORRESTER, M.D., HOWARD M. STANIOFF, M.D., RAN VAS, PH.D., STEPHEN W. HALPERN, M.D., AND H.J.C. SWAN, M.D., PH.D.

SUMMARY The inherent imperfection of clinical diagnostic tests introduces uncertainty into their interpretation. The magnitude of diagnostic uncertainty after any test result may be quantified by information theory. The information content of the electrocardiographic ST-segment response to exercise, relative to the diagnosis of angiographic coronary artery disease, was determined using literature-based pooled estimates of the true- and false-positive rates for various magnitudes of ST depression from \( < 0.5 \text{ mm} \) to \( \geq 2.5 \text{ mm} \). This analysis allows three conclusions of clinical relevance. First, the diagnostic information content of exercise-induced ST-segment depression, interpreted by the standard 1.0-mm criterion, averages only 15% of that of coronary angiography. Second, there is a 41% increase in information content when the specific magnitude of ST-segment depression is analyzed, as opposed to the single, categorical 1-mm criterion. Third, the information obtained from ECG stress testing is markedly influenced by the prevalence of disease in the population tested, being low in the asymptomatic and typical angina groups and substantially greater in groups with nonanginal chest pain and atypical angina.

The quantitation of information has broad relevance to selection and use of diagnostic tests, because one can analyze objectively the value of different interpretation criteria, compare one test with another and evaluate the cost-effectiveness of both a single test and potential testing combinations.

INFORMATION THEORY was developed as a means for quantifying the information content and the noise content in the audio signal transmitted over telephone channels. Because the general theory provides a formal mathematical basis for quantifying information and uncertainty, it has been widely applied in other sciences as diverse as economics and engineering, but has been little used in clinical medicine presumably because its relevance was not apparent.

The application of information analysis to diagnostic testing in medicine derives from the recognition that all clinical tests are imperfect. This imperfection introduces uncertainty (or "noise") into the interpretation of the test. The information (and, conversely, the uncertainty existing before and after a diagnostic test) can be quantified if one knows its three determinants. These are the pretest probability of disease in the individual being tested (prevalence), the conditional probability of the test response observed in a diseased population (the true-positive rate), and the conditional probability of this same response in a non-diseased population (the false-positive rate). The change in diagnostic uncertainty that occurs as a result of testing, then, is a measure of the test's effectiveness (fig. 1).

The purpose of this report, therefore, is to derive an equation that defines the effectiveness of any diagnostic test throughout the range of these three determinants and to illustrate the value of this equation by analysis of the ST-segment response to exercise as a diagnostic marker of coronary artery disease. Three fundamental questions relative to this test are analyzed: What is the diagnostic information content of the electrocardiographic ST-segment response to stress in comparison to that of angiography? What criteria for ST-segment interpretation provide the greatest diagnostic information? Does the information derived from the test vary significantly in different test populations?

Methods

The medical literature was reviewed unselectively to obtain pooled, least-biased estimates4 for the prevalence of coronary artery disease in groups of patients defined according to the presence of various chest pain syndromes6 (table 1), and for the true-positive and false-positive rates for various magnitudes of exercise-induced horizontal or downsloping ST-segment depression in relation to the angiographic presence of greater than 50% diameter narrowing of at least one major coronary artery6 (table 2). Using these data, the information content (I) of the ECG stress test, expressed in binary digits, or bits, was evaluated for all prevalence values from 0–1 at 0.001 increments according to the equation

\[
I = \sum_{f} \left[ \log_{2} \left( \frac{ap \cdot \log_{2} (ap) + bq \cdot \log_{2} (bq)}{(ap + bq) \cdot \log_{2} (ap + bq)} \right) - \sum_{i} p_{i} \cdot \log_{2} p_{i} \right]
\]
diagnosis of coronary artery disease with certainty. In practical terms, the value of a bit may be thought of as a percentage that expresses the relationship between existing information and certainty.

The information content of two formats for ST-segment analysis were compared. The categorical format considered the test as "positive" on the basis of each 0.5-mm threshold level of ST depression, from ≥ 0.5 mm to ≥ 2.5 mm. In contrast, the compartmental format considered the actual magnitude of observed ST depression as a component of the sum of all closed ST intervals, m ≤ ST < n, where m and n are adjacent 0.5-mm increments.

The variances associated with I and its integrated average relative to prevalence, I, were calculated according to methods described in the Appendix. Statistical comparisons were by analysis of variance (F test).

Results

The calculated indexes of information content for exercise induced ST-segment depression are summarized in Table 3.

Categorical Analysis

The information content for a categorical ("positive" vs "negative") interpretation of the ECG stress test was determined for each threshold level of ST depression from 0.5-2.5 mm. Figure 2 illustrates the information content for each threshold value over the entire range of disease prevalence from 0-100%. These data are compared to the hypothetical maximum amount of information obtainable with a "perfect" test defined by a true-positive rate of 1 and a false-

![Figure 1](http://circ.ahajournals.org/)

**Figure 1.** This schematic provides conceptualization of the terms used in this manuscript. A patient to be tested has a pretest likelihood of coronary artery disease of 10%. If the sensitivity and specificity of the test are known, the likelihood of the disease after testing can be calculated by Bayes' theorem. In this example, it is assumed to be 60%. Information theory allows calculation of the diagnostic uncertainty before and after testing (and conversely, the information gained from performance of the test). The information content of the test, which is a measure of its effectiveness, is defined mathematically in the Methods and Appendix.

**Table 1.** Prevalence of Coronary Artery Disease

<table>
<thead>
<tr>
<th>Symptom class</th>
<th>Patients</th>
<th>p(D+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymptomatic</td>
<td>23,996</td>
<td>0.045 ± 0.001</td>
</tr>
<tr>
<td>Nonanginal chest pain</td>
<td>913</td>
<td>0.160 ± 0.012</td>
</tr>
<tr>
<td>Atypical angina</td>
<td>1,931</td>
<td>0.499 ± 0.011</td>
</tr>
<tr>
<td>Typical angina</td>
<td>2,108</td>
<td>0.889 ± 0.007</td>
</tr>
</tbody>
</table>

Each value in tables 1 and 2 represents a percent proportion (P), expressed as a decimal, ± standard error of the percent (equivalent to standard deviation),

\[
SE_p = \sqrt{\frac{P(1-P)}{N}}
\]

where N is the number of patients from which the value of P is derived. The complete bibliography for the literature review from which these data were obtained is given in reference 5.

where \( a = p(T_+ | D+) \) = true positive rate, \( b = p(T_+ | D-) \) = false positive rate, \( p = p(D+) \) = prevalence of disease, \( q = p(D-) = 1-p \) = prevalence of non-disease.

The derivation of this equation is outlined in the Appendix. In concept, information content of a test is proportional to the magnitude of change in probability of having disease that occurs secondary to a test result. Because some test results are correct and others are false, the information content is the magnitude of valid reduction in diagnostic uncertainty because of performance of the test. Shannon and Weaver quantified information in a unit called the bit. If one had all the information necessary to answer a question (i.e., complete certainty) he would have 1 bit of information. Information is therefore expressed as a fraction of a bit; for example, 0.50 bit is equal to half the information necessary to establish or exclude
positive rate of 0. The figure shows that information obtained from the ST-segment response varied considerably as a function of prevalence. It reached 0 at the two prevalence extremes (0 and 100%), because no diagnostic information could be derived where no prior uncertainty existed. Conversely, the information derived from the ST-segment response tended to peak near 50% prevalence, where maximum uncertainty ("50–50 chance") existed before testing. From the graph, two conclusions may be drawn. The first is that the average information content (I), represented by the integrated area under the curve, was very low, averaging only 0.11 × 0.03 bit, compared with 0.72 bit for a perfect test (p < 0.001). The greatest information content was obtained using the criterion ≥ 1.0 mm (I = 0.14), but the criterion of ≥ 0.05 mm was almost as powerful (I = 0.13) and, above a prevalence of 70%, was slightly superior (p = NS). The second conclusion is that use of different threshold criteria produces little difference in information content (fig. 3). The amount of information falls as more strict categorical criteria are used (p < 0.001), but the

---

**Table 2. True- and False-positive Rates for ST Depression**

<table>
<thead>
<tr>
<th>Observation (j)</th>
<th>True-positive rate p (Tj[D+]</th>
<th>False-positive rate p (Tj[D−])</th>
<th>Likelihood ratio p (Tj[D+]) p (Tj[D−])</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST ≥ 0.5</td>
<td>0.857 ± 0.033</td>
<td>0.375 ± 0.057</td>
<td>2.3 ± 0.4</td>
</tr>
<tr>
<td>ST ≥ 1.0</td>
<td>0.649 ± 0.009</td>
<td>0.148 ± 0.008</td>
<td>4.4 ± 0.2</td>
</tr>
<tr>
<td>ST ≥ 1.5</td>
<td>0.416 ± 0.023</td>
<td>0.038 ± 0.011</td>
<td>11.0 ± 3.2</td>
</tr>
<tr>
<td>ST ≥ 2.0</td>
<td>0.328 ± 0.017</td>
<td>0.017 ± 0.006</td>
<td>19.3 ± 3.9</td>
</tr>
<tr>
<td>ST ≥ 2.5</td>
<td>0.195 ± 0.016</td>
<td>0.005 ± 0.005</td>
<td>39.9 ± 39.1</td>
</tr>
</tbody>
</table>

**Compartmental criteria**

| 0 ≤ ST < 0.5   | 0.143 ± 0.033               | 0.625 ± 0.057                 | 0.2 ± 0.1                              |
| 0.5 ≤ ST < 1.0 | 0.208 ± 0.034               | 0.227 ± 0.058                 | 0.9 ± 0.3                              |
| 1.0 ≤ ST < 1.5 | 0.233 ± 0.025               | 0.110 ± 0.014                 | 2.1 ± 0.4                              |
| 1.5 ≤ ST < 2.0 | 0.088 ± 0.029               | 0.021 ± 0.013                 | 4.3 ± 2.9                              |
| 2.0 ≤ ST < 2.5 | 0.133 ± 0.023               | 0.012 ± 0.008                 | 11.1 ± 7.6†                            |
| 2.5 ≤ ST < ∞   | 0.193 ± 0.016               | 0.005 ± 0.005                 | 39.0 ± 39.1                            |

Abbreviations: p (Tj[D+]) = true-positive rate for the observation j; p (Tj[D−]) = false-positive rate for the observation j.

**Table 3. Information Content of the Electrocardiographic Stress Test**

<table>
<thead>
<tr>
<th>Observation (j) (mm)</th>
<th>$\bar{I} = \sigma_I$ (bits)</th>
<th>$I_{max} = \sigma_{I_{max}}$ (bits)</th>
<th>$%I_{max}$ (%)</th>
<th>$\bar{I}/I_{max}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categorical analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST ≥ 0.5</td>
<td>0.127 ± 0.035</td>
<td>0.188 ± 0.049</td>
<td>52.5</td>
<td>17.6</td>
</tr>
<tr>
<td>ST ≥ 1.0</td>
<td>0.135 ± 0.007</td>
<td>0.201 ± 0.010</td>
<td>48.0</td>
<td>18.7</td>
</tr>
<tr>
<td>ST ≥ 1.5</td>
<td>0.114 ± 0.013</td>
<td>0.169 ± 0.018</td>
<td>43.7</td>
<td>15.8</td>
</tr>
<tr>
<td>ST ≥ 2.0</td>
<td>0.100 ± 0.008</td>
<td>0.148 ± 0.012</td>
<td>41.9</td>
<td>13.9</td>
</tr>
<tr>
<td>ST ≥ 2.5</td>
<td>0.063 ± 0.007</td>
<td>0.094 ± 0.010</td>
<td>40.0</td>
<td>8.8</td>
</tr>
<tr>
<td>Compartmental analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Sigma m \leq ST &lt; \eta$</td>
<td>0.190 ± 0.084</td>
<td>0.277 ± 0.115</td>
<td>47.5</td>
<td>26.3</td>
</tr>
</tbody>
</table>

**Figure 3. Relationship of average information content ($\bar{I}$) to the categorical magnitude of ST depression.**
magnitudes of this reduction remains small until the 2.5-mm criterion is reached.

Compartmental Analysis

Figure 4 is a comparison of the information content of the ECG stress test obtained by compartmental analysis to that of ≥ 1.0 mm, the "best" categorical criterion. The graph allows two conclusions. The first is that compartmental information content is only 26% of that obtained by angiography (I = 0.19, p < 0.001). The second is that compartmental analysis provides a substantial increase (41%, p < 0.02) in information content compared to the use of categorical criteria for interpretation of the same test response.

Information Content of the ECG Stress Test
As a Function of Pretest Likelihood

Figure 5 illustrates the information content for compartmental analysis of the ECG stress test in four prevalence groups defined by type of chest pain (table 1). There is a fivefold difference in information content derived from testing in these four groups, from a low of only 0.05 bit when ECG stress testing is used in asymptomatic patients (pretest likelihood 4.5%) to 0.28 bit in patients with atypical angina (pretest likelihood 49.9%). These data illustrate that the most information from an ECG stress test is obtained in those patients with an intermediate level of pretest likelihood, the groups with nonanginal chest pain and atypical angina. The least information is obtained by using the test in an asymptomatic population, and only slightly more diagnostic information is obtained when the test is used in patients with typical angina.

Discussion

Since its classic description by Shannon in 1948, information theory has been validated by both rigorous mathematical proof6 and by practical application to fields as diverse as engineering8 and economics.10 This broad applicability derives from the fundamental statistical nature of information: What we "know" invariably contains some degree of uncertainty. The clinical diagnostic process also concerns itself with the gathering of information under conditions of uncertainty. It is not surprising, therefore, that it might also lend itself to a similar approach.

Our selection of the electrocardiographic ST-segment response for informational analysis was based on the mass of literature available for estimation of the necessary conditional probabilities, and on the wide clinical experience to which the derived inferences might be compared. Analysis of the exercise-induced ST-segment response relative to the diagnosis of angiographic coronary artery disease using information theory provided several insights with direct clinical relevance.

First, the information content of exercise-induced ST depression was low, averaging only 15% and, at best, 26% of its reference standard, angiographic presence or absence of coronary disease. This observation supports the view of several investigators6, 7, 11-14 that diagnostic reliance on the ST response alone is overly simplistic and at times may be seriously misleading. Moreover, it has been suggested that the stress ECG may be even less reliable in the face of resting ST-segment abnormalities, in women, and in those with less than typical symptoms of ischemia. Weiner et al.,11 for instance, recently reported the results of stress testing in 2045 patients who underwent coronary angiography. Applying our methods to their data, the information content of ≥ 1.0 mm ST depression in the 1528 patients with a normal resting ECG was identical.
to that reported here (\( \bar{I} = 0.14 \)). In the 517 patients with an abnormal resting ECG, however, the same diagnostic criterion had 57% less \( (p < 0.05) \) information content (\( \bar{I} = 0.06 \)). Of additional interest, although the information content in males averaged three times that of females (\( \bar{I} = 0.13 \) vs 0.04), this difference was not statistically significant. Similarly, no significant differences were observed when the population was grouped by symptom class. These observations are in accord with the authors' own conclusions, which were based upon a more traditional analysis of sensitivity and specificity. These data, therefore, support the use of additional independent procedures in diagnosis to improve the information content of the testing process. This theoretical justification of multiple testing is consistent with the improvement in diagnostic accuracy when the ECG stress test is analyzed in conjunction with thallium scintigraphy, cardiac fluoroscopy or cardiokymography.\textsuperscript{15-18}

A second important observation in this study was that the information content of the ECG stress test may be substantially influenced by the prevalence of disease. Thus, there was a 500% difference in information content between the intermediate prevalence group ("atypical angina") and the extreme prevalence groups (typical angina or asymptomatic). This finding also has direct relevance to the selection of patients for diagnostic testing and the criteria used for test interpretation. The information to be derived from a test varies greatly with the pretest likelihood, while the cost of the test remains fixed, so important and quantifiable differences exist in the cost-effectiveness of a test.\textsuperscript{19} The low information content of the ECG stress test when used for asymptomatic patients, for example, is the basis for using more strict diagnostic criteria in low-prevalence patients to minimize false-positive diagnoses. The overlap of the individual information content curves for each categorical ST-segment criterion (fig. 2) indicates, however, that such modification would have little impact upon the effectiveness of the ECG stress test. This occurs because the information gain achieved by reduction in false diagnoses is offset by a counterbalancing reduction in the number of true diagnoses.\textsuperscript{18, 20}

A third conclusion of relevance to clinical ECG stress testing is that compartmental analysis substantially increases information content. This finding is predicted by information theory,\textsuperscript{1} which states that the greater the number of mutually exclusive test probabilities, the greater the information. This observation is also consistent with studies in which the predictive accuracy of the ECG stress test was found to increase by application of multivariate analysis to the magnitude of ST depression.\textsuperscript{21} To illustrate the value of compartmental analysis in a practical clinical setting, consider a patient with a pretest likelihood of 0.5. If this patient has a "positive" ST-segment response to exercise (defined as \( \geq 1.0 \) mm), then the categorical post-test likelihood increases to 0.81 (calculated from the data in table 2 according to equations 4 and 5 in the Appendix). Depending on the actual observed magnitude of ST depression, however, compartmental analysis results in a spread of this single pooled probability from 0.68 (for 1.0 mm of depression) to 0.98 (for 2.5 mm of depression). Likewise, if the patient instead had a "negative" response (\(< 1.0 \) mm), then his categorical post-test likelihood falls to 0.29. Again, compartmental analysis reveals a range of this value from 0.19 (for 0 mm of depression) to 0.48 (for 0.5 mm of depression). Although the physician learns little by a change in probability from 0.5 to 0.48, patients can and do show such equivocal responses. Categorical analysis of probability at least allows one to know more precisely when one has learned something and when one has not. Such knowledge is indispensable for rational and cost-effective clinical decision-making. An intermediate probability, for example, might indicate the need for additional diagnostic testing. If the decision is made to perform an additional test, information theory then allows one to predict the comparative effectiveness and cost-effectiveness of the available tests before their actual performance.\textsuperscript{18}

One might wonder why compartmental analysis is not often used when it results in a 40% increase in diagnostic information content at no increase in cost. The answer probably relates to two factors: First, the value of compartmental analysis has been largely unrecognized. The second factor relates to the complexity of its practical application. Tables that provide the likelihood (and associated statistical variance) for the presence of coronary artery disease in relation to age, sex, symptoms and magnitude of ST-segment depression, however, have recently been published\textsuperscript{4} and computer programs incorporating these data are also available. We believe, therefore, that application of these concepts is both practical and relevant, and can routinely provide an assessment of probability and information content of that probability as part of the report from our stress and nuclear scintigraphy laboratories.

The validity of the conclusions drawn from this analysis depends on the accuracy of the input probabilities. Literature-based pooled estimates, although subject to many limitations,\textsuperscript{4} appear to provide a reasonable basis for such preliminary determinations.\textsuperscript{4} As more and better data become available, however, these conclusions may be modified or refined. Important responses such as exercise-induced hypotension\textsuperscript{22} or ST-segment slope\textsuperscript{23} might be incorporated into this analysis, if they are stochastically independent of ST depression, or if the analytical equations are modified by appropriate corrective weighting functions.\textsuperscript{3} Informational analysis, therefore, may provide a basis for optimizing the selection and interpretation of diagnostic tests and, thereby, may allow for more rational and cost-effective use of these procedures.

Acknowledgment

We are grateful to Melinda Unsworth and Tina Orvis for secretarial assistance, to Lance Laforteza for help with the illustrations, and to Patricia Edwards for photography.
Calculation of Information Content

Information theory defines the average uncertainty for a set, $i$, of mutually exclusive probabilistic events (such as the presence ($D^+$) and absence ($D^-$) of coronary artery disease) in terms of the probability associated with each event. If the events are denoted as $D_i$ and the prior probability, or frequency of occurrence, by $p(D_i)$, then a priori uncertainty, expressed in binary digits, or bits, is given by

$$U_{\text{pre}} = - \sum_i p(D_i) \log_2 p(D_i)$$

(1)

Similarly, the a posteriori uncertainty (resulting from performance of a test) is defined in relation to the set of posterior probabilities, $p(D_i | T_j)$, as the average uncertainty associated with all mutually exclusive test responses ($T_j$), weighted by their respective frequencies of occurrence, $p(T_j)$:

$$U_{\text{post}} = - \sum_j p(T_j) \sum_i p(D_i | T_j) \log_2 p(D_i | T_j)$$

(2)

The information content of the test is defined as the average reduction in uncertainty, obtained by subtracting equation 2 from equation 1:

$$I = \sum_j p(T_j) \sum_i p(D_i | T_j) \log_2 p(D_i | T_j) - \sum_i \sum_j p(D_i) \log_2 p(D_i)$$

(3)

The conditional probabilities in equation 3 are interrelated through Bayes' theorem:

$$p(D_i | T_j) = p(D_i)p(T_j | D_i)/p(T_j)$$

(4)

$$p(T_j) = p(D+)p(T_j | D+) + p(D-)p(T_j | D-)$$

(5)

The term $p(T_j | D^+)$ is the conditional probability for the test observation $j$ in a population of diseased patients (true-positive rate), and $p(T_j | D^-)$ is the conditional probability for $j$ in the complementary population, $D^-$, of nondiseased patients (false-positive rate). By substitution of equations 4 and 5 into equation 3, therefore, one obtains the information content, $I$, expressed, as a function of the three variables $p(D^+)$, $p(T_j | D^+)$, and $p(T_j | D^-)$:

$$I = \sum_j \{ \log_2 (p(D^+)p(T_j | D^+)+p(D^-)p(T_j | D^-)) - \sum_i p_i \log_2 p_i \}$$

(6)

where $a = p(T_j | D^+)$, $b = p(T_j | D^-)$, $p = p(D^+)$ and $q = p(D^-)$. Inspection of equation 6 reveals that $I=0$ for any set of $a$ and $b$ when $p=1$ or 0, and monotonically reaches a single peak, $I_{\text{max}}$, at some prevalence, $p_{\text{max}}$, between these limits.

The average information content of the test is represented by the area under the information curve, given by integration of equation 3 with respect to $p(D^+)$ over the interval 0 to 1 (for simplicity of notation, we will substitute $p$ for $p(D^+)$):
The Statistical Variance of Information Content

Each of the probabilities incorporated into the calculation of information content is associated with a variance, \( \sigma^2 = PQ/N \), where \( P \) is the probability, \( Q \) is \( 1 - P \) and \( N \) is the population sample from which \( P \) was estimated. The probabilities are independently determined, so their variances are independent and uncorrelated. The variance associated with the calculated information content, \( \sigma^2_I \), is, therefore, given by the partial differential equation:

\[
\sigma^2_I = \left( \frac{\partial I}{\partial a} \right)^2 \sigma^2_a + \left( \frac{\partial I}{\partial b} \right)^2 \sigma^2_b + \left( \frac{\partial I}{\partial p} \right)^2 \sigma^2_p \tag{8}
\]

We may evaluate these partials by substituting equation 4 into equation 3 and expanding relative to \( p(D) \):

\[\frac{\partial I}{\partial a} = \sum_j p(D+) \log_2 p(D+ | T_j) \tag{9}\]

\[\frac{\partial I}{\partial b} = \sum_j p(D-) \log_2 p(D- | T_j) \tag{10}\]

\[
\frac{\partial I}{\partial p} = \log_2 [p(D-) p(D+)] + \sum_j p(T_j | D+) \log_2 p(D+ | T_j) - \sum_j p(T_j | D-) \log_2 p(D- | T_j) \tag{11}\]

Similarly, the variance of \( I \) is given as the integrated average variance of \( I \) over the entire range of \( p \):

\[
\sigma^2_I = \int_0^1 \sigma^2_I dp \tag{12}\]

These calculated variances were used in the statistical comparisons reported in this manuscript.

Information Content vs Likelihood Ratio for Assessing Test Effectiveness

The likelihood ratio, which is the ratio of the true-positive rate to the false-positive rate, has been widely used as a means of assessing a test’s diagnostic effectiveness.\(^\text{7-20}\) Table 2 shows that the likelihood ratio for the ECG stress test increases progressively, from 2.3 for the categorical criterion \( \geq 0.5 \) mm to 39.0 for \( \geq 2.5 \) mm. As discussed earlier, however, information content fell slightly as the categorical threshold criterion was increased from 0.5 mm to 2.5 mm, so that a negative correlation existed between the two variables. The likelihood ratio, however, does not consider the prevalence of disease in the subject being tested, nor is it weighted for the frequency of occurrence of the observation. The term effectiveness therefore may have two meanings. As determined by information content, effectiveness refers to the weighted value of the test relative to all possible observations. As determined by likelihood ratio, effectiveness refers only to the relative value of a specific observation in normal and abnormal subjects in isolation from the other possible observations. These theoretical limitations, coupled with its failure to correlate positively with information content suggest that the likelihood ratio is a misleading index of test effectiveness.

Information Content of Diagnostic Likelihood

There are numerous potential sources of error in application of compartmental probability analysis. Indecision and misclassification can arise, for instance, in determination of symptoms and in measurement of the magnitude of ST depression. These errors influence the uncertainty associated with the resultant assessment of diagnostic likelihood, and this uncertainty also can be quantified by information theory if one knows the variance associated with the posterior likelihood. If we denote \( p(D+ | T_j) \) as \( \bar{p} \) and \( q = 1 - \bar{p} \), then this variance is given by the following equation, analogous to equation 8:

\[
\sigma^2_p = \left( \frac{\bar{p} q}{\bar{p}^2} \right)^2 \left[ \left( \frac{\sigma_p}{\bar{p}} \right)^2 + \left( \frac{\sigma_q}{q} \right)^2 \right] \tag{13}\]

The posterior likelihood and its variance define a specific density function, termed the \( \beta \) frequency distribution:\superscript{24}

\[
B(n,r) = \frac{\Gamma(n)}{\Gamma(r) \Gamma(n-r)} \cdot \bar{p}^r \cdot q^{n-r} \tag{14}\]

where \( p \) is probability, \( q = 1 - p \), \( n = \bar{p} q / \sigma_p^2 \), \( r = n \bar{p} \) and the symbol \( \Gamma \) denotes the gamma function. The information content of this continuous distribution is defined in a manner identical to that for the case of discrete probabilities, given in equations 1-3, except that the summation of the discrete probabilities is replaced by integration of the probability distribution (logarithms have been converted to base \( e \)):

\[
\B(n,r) = 1 + 2 \int_0^1 B(n,r)(p \ln p + q \ln q) dp \tag{15}\]

This equation, therefore, expresses the total information content of all the findings in a given subject, relative to a final categorical diagnosis for or against disease.
Application of information theory to clinical diagnostic testing. The electrocardiographic stress test.
G A Diamond, M Hirsch, J S Forrester, H M Staniloff, R Vas, S W Halpern and H J Swan

doi: 10.1161/01.CIR.63.4.915

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1981 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/63/4/915

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/