

## Interval Likelihood Ratios: Another Advantage for the Evidence-Based Diagnostician

**Michael D. Brown, MD**  
**Mathew J. Reeves, PhD**

*From the Grand Rapids MERC/  
Michigan State University  
Program in Emergency  
Medicine, Grand Rapids, MI  
(Brown), and the Department of  
Epidemiology, Michigan State  
University, East Lansing, MI  
(Reeves).*

**See editorial, p. 298.**

Emergency physicians are often confronted with making diagnostic decisions on the basis of a test result represented on a continuous scale. When the results of continuous data are expressed as binary outcomes using a single cutoff, loss of information and distortion may occur. In this setting, interval likelihood ratios provide a distinct advantage in interpretation over those based on a dichotomized sensitivity and specificity. Dividing the data into intervals uses more of the information contained in the data and allows the clinician to more appropriately interpret the test results and to make valid clinical decisions. This article illustrates the advantages of interval likelihood ratios with examples and demonstrates how to calculate them on the basis of different data formats. Authors and journals need to be encouraged to report the results of studies of performance of diagnostic tests using interval ranges rather than simple dichotomization when the tests involve continuous variables.

[*Ann Emerg Med.* 2003;42:292-297.]

Copyright © 2003 by the American  
College of Emergency Physicians

0196-0644/2003/\$30.00 + 0  
doi:10.1067/mem.2003.274

## INTRODUCTION

Clinicians look to the results of diagnostic tests such as peripheral WBC counts and cardiac biomarkers as the basis for modifying their estimates of how likely it is that a particular patient has a clinically important disease, condition, or injury. Studies of the performance of such tests commonly simplify their results by calculating sensitivity and specificity in comparison to a criterion standard for the presence or absence of the disease entity, where all values above a single threshold level are considered “positive,” and all those below it are considered “negative.” This implies that all test results above the threshold increase the likelihood that the disease is present to the exact same degree.

However, using appendicitis as an example, clinicians instinctively recognize that a WBC count of  $18 \times 10^3/\mu\text{L}$  renders a patient with abdominal pain and a consistent clinical presentation more likely to have appendicitis than if the WBC count were only  $12 \times 10^3/\mu\text{L}$  (even though both values are elevated). Similarly, a clinician’s suspicion that a patient with chest pain is having a myocardial infarction is likely to be much greater if the troponin I level were  $5.0 \mu\text{g/L}$  than if it were  $0.5 \mu\text{g/L}$ , even though both are above a typical standard cutoff of  $0.4 \mu\text{g/L}$ . In this article, we will clarify the nature of the problem presented by diagnostic tests having a continuous set of possible results. We will show how likelihood ratios based on interval ranges help to quantify the differences in diagnostic effect that clinicians instinctively recognize in these settings, and how their use can help to avoid erroneous interpretations of such a test when the results lie near a dichotomous cutoff value.

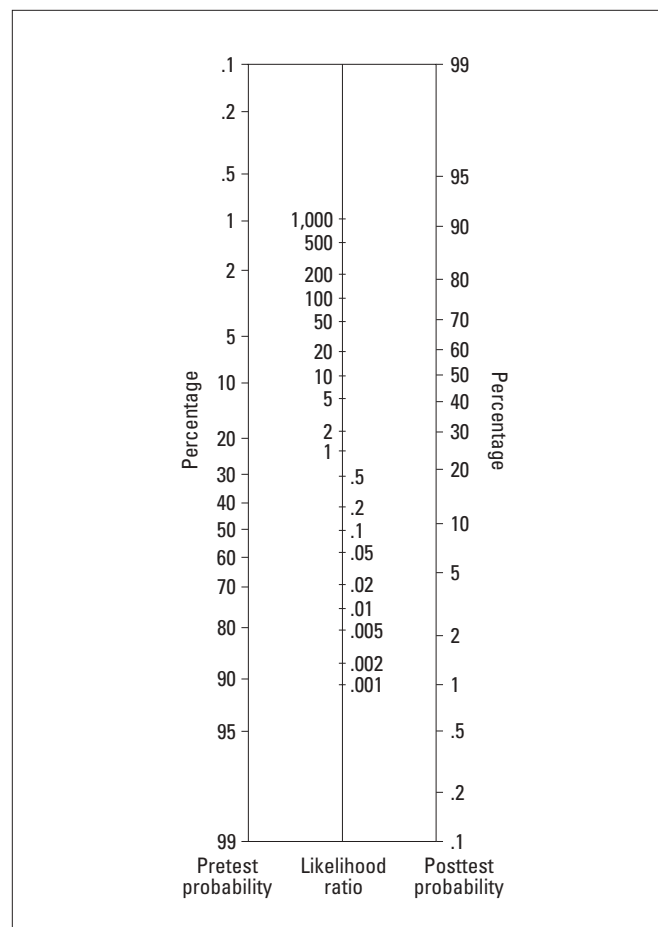
## HOW DO LIKELIHOOD RATIOS HELP CLINICIANS MAKE DECISIONS?

A previous installment of *Annals’* “Skills for Evidence-Based Emergency Care” series explained the unique value of likelihood ratios in the interpretation of diagnostic test results.<sup>1</sup> The likelihood ratio is the ratio of the probability of a given test result in patients with disease to the probability of the same test result in patients

without disease. This ratio represents the magnitude of change from a clinician’s initial suspicion for disease (pretest probability) to the likelihood of disease after the test result (posttest probability). Using the likelihood ratio for a particular test result, the Fagan nomogram (Figure 1) provides a graphic representation of

**Figure 1.**

Nomogram for interpreting diagnostic test results. Place a straight edge on the left side of the figure at a point corresponding to an estimate of the probability of a disease before the performance of a test. Connect that point to the likelihood ratio corresponding to the test result in the middle of the figure. The point of intersection of the straight edge with the right-hand side of the figure corresponds to the probability of the disease implied by the test result. Adapted from Fagan TJ. Nomogram for Bayes theorem [letter]. *N Engl J Med.* 1975;293:257. Copyright © 1975 Massachusetts Medical Society. All rights reserved.<sup>11</sup>



how the test result changes the probability of disease.<sup>1</sup> Test results with likelihood ratios greater than 1.0 or less than 0.1 effect a substantial change on disease probability over a broad range of pretest estimates. A test result with a likelihood ratio of 1 leaves the probability of disease unchanged.

Diagnostic test characteristics are typically reported in terms of sensitivity and specificity, parameters that do not directly allow clinicians to understand the effect of a test on the likelihood of disease.<sup>1</sup> To calculate sensitivity and specificity for a continuous variable, the test results must be forced into a binary or dichotomous classification.<sup>2</sup> The optimal cutoff value is often derived using a receiver operating characteristic (ROC) curve.<sup>3</sup> The cutoff value, or test threshold, is often chosen as the point on the curve where the sensitivity and specificity of the test are both maximized. However, simple dichotomization of continuous data may result in a loss of valuable information and may cause distortions in the interpretation of test results when used in clinical practice.<sup>3,4</sup> The distortions are especially exaggerated when the patient's test result is close to the established cutoff value.<sup>4</sup> The advantages of likelihood ratios particularly come to the fore in this context. Compared with the traditional approach of calculating sensitivity and specificity using just 2 intervals of data, "positive" and "negative," dividing the data into several intervals of test results and calculating a likelihood ratio for each interval uses more of the information.

## WHY DO INTERVAL LIKELIHOOD RATIOS USE MORE OF THE DATA?

To demonstrate the advantage of using interval likelihood ratios rather than a simple dichotomization of a continuous variable, we will use data from a study by Andersson et al.<sup>5</sup> The authors evaluated the utility of a broad range of criteria, including the WBC count, in the diagnosis of acute appendicitis in patients admitted to 2 hospitals in Sweden. When the results are displayed in a simple 2×2 table format using a WBC count of  $10.0 \times 10^3/\mu\text{L}$  as the cutoff, likelihood ratios for positive ( $>10.0 \times 10^3/\mu\text{L}$ ) and negative ( $<10.0 \times 10^3/\mu\text{L}$ ) test

results of 2.4 and 0.3, respectively, can be calculated (Table 1). However, the data as reported by Andersson et al allow the test results to be stratified into more than 2 levels or intervals, as shown in Table 2. The calculation of the interval likelihood ratios for each of the 3 ranges of WBC count results is simply the percentage of patients with appendicitis having WBC counts in each range, or interval, divided by the percentage of patients without appendicitis having WBC counts in the same range.<sup>1</sup> When this method is applied to the data in Table 2, the likelihood ratios range from 0.2 to 7.0. The wider range of likelihood ratios suggest that the data are pro-

**Table 1.**

*Deriving likelihood ratios from a standard 2×2 table for WBC count in the diagnosis of acute appendicitis.<sup>5</sup>*

Group	Appendicitis+	Appendicitis–	Total
Test result positive (WBC $\geq 10 \times 10^3/\mu\text{L}$ )*	148	95	243
Test result negative (WBC $< 10 \times 10^3/\mu\text{L}$ )†	42	202	244
Total	190	297	487

\*Positive likelihood ratio is the probability of a positive test result when appendicitis is present divided by the probability of a positive test result when appendicitis is absent (Positive likelihood ratio =  $(148/190)/(95/297) = 2.4$ ).

†Negative likelihood ratio is the probability of a negative test result when appendicitis is present divided by the probability of a negative test result when appendicitis is absent (Negative likelihood ratio =  $(42/190)/(202/297) = 0.3$ ).

**Table 2.**

*Deriving interval likelihood ratios\* from a 2×3 table for WBC count in the diagnosis of acute appendicitis.<sup>5</sup>*

WBC Count ( $\times 10^3/\mu\text{L}$ )	Appendicitis+	Appendicitis–	Total
$\geq 15^\dagger$	63	14	77
8 to $<15^\ddagger$	111	130	241
$<8^\S$	16	153	169
Total	190	297	487

\*Interval likelihood ratio is the probability of the test result when appendicitis is present divided by the probability of the test result when appendicitis is absent.

†Likelihood ratio for WBC count  $\geq 15 = (63/190)/(14/297) = 7.0$ .

‡Likelihood ratio for WBC count 8 to  $<15 = (111/190)/(130/297) = 1.3$ .

§Likelihood ratio for WBC count  $<8 = (16/190)/(153/297) = 0.2$ .

viding “more” clinically useful information than when the data are presented in a dichotomized format (where the likelihood ratio ranged from 0.3 to 2.4).

The same clinical example also demonstrates how distortion can occur when quantitative data are expressed as a binary outcome. Using the data shown in Table 1, a patient being evaluated for possible appendicitis having a WBC count of  $9.0 \times 10^3/\mu\text{L}$  would be regarded as having a “negative” test result associated with a likelihood ratio of 0.3. The Fagan nomogram<sup>1</sup> (Figure 1) allows the user to quickly ascertain that this likelihood ratio would lower a clinician’s estimate of the probability of acute appendicitis from 50% to just a little above 20%. However, when the data are tabulated using the 3 interval ranges shown in Table 2, the same WBC count of  $9.0 \times 10^3/\mu\text{L}$  corresponds to a likelihood ratio of 1.3. The estimate of the probability of acute appendicitis would now be increased from 50% to almost 60% by this likelihood ratio. In fact, Andersson et al<sup>5</sup> chose to report their data using an even larger number of WBC count intervals. In their own table, a WBC count of  $9.0 \times 10^3/\mu\text{L}$  corresponds to a likelihood ratio of 0.83. Both 1.3 and 0.83 are very close to 1, suggesting that a WBC count of  $9.0 \times 10^3/\mu\text{L}$ , when assessed using interval likelihood ratios, has negligible effect on the likelihood of appendicitis.

This example illustrates how the clinical significance of a test result may be exaggerated when the result is close to the established dichotomous cutoff value. The WBC count data could, of course, be divided into smaller and smaller intervals. Snyder and Hayden<sup>6</sup> derived likelihood ratios for multiple intervals for the WBC count in the diagnosis of acute appendicitis from published data and concluded that the WBC count was only useful at the extremes (ie,  $<7$  or  $>17 \times 10^3/\mu\text{L}$ ). Only at these levels do the likelihood ratios produce meaningful changes in the estimate of the probability of disease (Table 3). To be sure, unless the data set is very large, a significant loss of precision occurs when multiple intervals are created.<sup>7</sup> The 95% confidence interval (CI) can be calculated for likelihood ratios<sup>8</sup> and the loss of precision can be seen by examination of the wide 95% CIs in this example (Table 3).

## HOW DO INTERVAL LIKELIHOOD RATIOS RELATE TO ROC CURVES?

The ROC curve provides a visual display of the relationship between the choice of possible cutoff values and the corresponding sensitivities and specificities for a continuous diagnostic test variable (Figure 2).<sup>2</sup> The ROC curve is generated by plotting the true-positive rate along the y-axis and the false-positive rate along the x-axis for a series of different potential dichotomous cutoffs.<sup>3</sup> Although reporting results of a diagnostic test in the form of an ROC curve is informative, application of the results to clinical practice is often difficult.<sup>3</sup> However, by deriving interval likelihood ratio estimates from the ROC curve, this information can be more readily applied to the clinical setting.

As already noted, the likelihood ratio for a dichotomous test result is equal to the true-positive rate divided by the false-positive rate. As a result, it can be shown mathematically that the slope of the tangent to any point on an ROC curve is equal to the likelihood ratio of the test result corresponding to that point on the curve.<sup>2</sup> Alternatively, the slope of a line connecting any 2 points on the ROC curve is equal to the likelihood ratio for that interval.<sup>2</sup> The slope of a line is calculated by dividing the difference of the true-positive rates for the 2 points

**Table 3.**

*Likelihood ratios for WBC counts in the diagnosis of acute appendicitis based on 8 defined intervals. Adapted from Snyder BK, Hayden SR. Accuracy of leukocyte count in the diagnosis of acute appendicitis. Ann Emerg Med. 1999;33:565-574. Reprinted with permission.<sup>6</sup>*

WBC Count ( $\times 10^3/\mu\text{L}$ )	Likelihood Ratio (95% CI)
4–7	0.10 (0–0.39)
7–9	0.52 (0–1.57)
9–11	0.29 (0–0.62)
11–13	2.8 (1.2–4.4)
13–15	1.7 (0–3.6)
15–17	2.8 (0–6.0)
17–19	3.5 (0–10)
19–22	NA

NA, Not applicable.

by the difference of the false-positive rates for the same 2 points. Readers who remember their mathematics lessons from high school may recognize this formula for the slope as “the rise over the run.”

As an example, Harrison et al<sup>9</sup> evaluated the use of B-type natriuretic peptide to predict the likelihood of cardiac events in patients with congestive heart failure. They summarized their results using the typical binary approach. The positive likelihood ratio and negative likelihood ratio values calculated using the cutoff suggested by the authors of 480 pg/mL were 5.7 and 0.4, respectively. However, by using 2 cut-points of 230 pg/mL and 480 pg/mL identified by the authors, 3 interval likelihood ratios may be estimated from the ROC

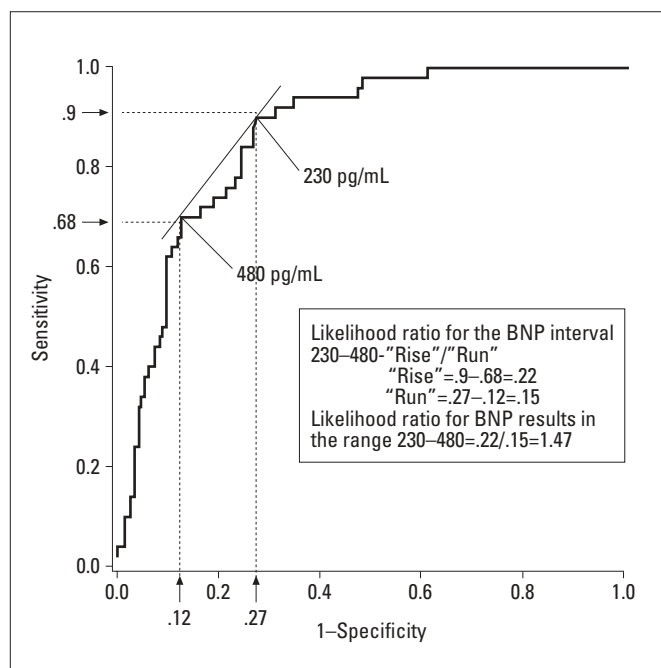
curve using the previously described approach (Figure 2 and Table 4).

Again, the interval likelihood ratios provide the clinician attempting to interpret a test result in the clinical setting with more useful information. Using the dichotomous cutoff of 480 pg/mL, a patient with a B-type natriuretic peptide of 400 pg/mL, corresponding to a likelihood ratio of 0.4, would seem to have a moderate decrease in the probability of having a cardiac event. Using the interval likelihood ratios derived from the ROC curve, the probability of having such an event would be minimally increased by virtue of the likelihood ratio of 1.5 for the 230 to 480 pg/mL interval.

This is a second example of the “paradoxical” distortion that may arise when a test result falls near a dichotomous cutoff value chosen for a continuous diagnostic test variable. Viewed from the perspective of the dichotomous cutoff, the result appears to have the opposite direction of effect on the pretest probability that would be implied when interval likelihood ratios are used. Although similar distortion may occur with a test result near any cutoff, whether treated dichotomously or within an interval range, the degree of distortion will be less when a greater number of intervals are used. Hopefully, as investigators start to report results in terms of interval likelihood ratios, it will become unnecessary for the evidence-based diagnostician to have to perform this type of extrapolation from the ROC curve.<sup>10</sup>

**Figure 2.**

An example of the calculation of interval likelihood ratios from an ROC curve for B-type natriuretic peptide as a predictor of cardiac events in patients with congestive heart failure. **BNP**, B-type natriuretic peptide. Adapted from Harrison A, Morrison LK, Krishnaswamy P, et al. B-type natriuretic peptide predicts future cardiac events in patients presenting to the emergency department with dyspnea. *Ann Emerg Med*. 2002;39:131-138.<sup>9</sup> Reprinted with permission.



**Table 4.**

Likelihood ratios for B-type natriuretic peptide levels estimated from the ROC curve.<sup>9</sup>

BNP, ng/mL	Likelihood Ratio
>480*	5.7
230-480†	1.5
<230‡	0.1

\*Likelihood ratio >480=(change in true positive rate/change in false positive rate)=(.68-0/.12-0)=5.7.  
†Likelihood ratio 230 to 480=1.5 (Figure 2).  
‡Likelihood ratio <230=(change in true positive rate/change in false positive rate)=(1-.9/1-.27)=0.1.

In summary, emergency physicians make clinical decisions on the basis of the results of diagnostic tests of continuous variables on a daily basis. To interpret these results appropriately, the clinician must understand the limitations and distortions that may occur when a continuous variable is converted to a dichotomous variable by using a single cutoff for defining positive and negative test results. In these situations, interval likelihood ratios use more of the information contained in the data and are less likely to exaggerate or to distort the clinical significance of test results than are likelihood ratios derived from a dichotomous "sensitivity" and "specificity." Consequently, the clinician is able to more appropriately interpret the results of a diagnostic test for an individual patient.

We acknowledge the assistance of Peter C. Wyer, MD, in preparing the manuscript.

Received for publication October 2, 2002. Revisions received January 12, 2003, and February 6, 2003. Accepted for publication February 11, 2003.

The authors report this study did not receive any outside funding or support.

**Reprints not available from the authors.**

**Address for correspondence:** Michael D. Brown, MD, Emergency Medicine, Spectrum Health-Butterworth, 100 Michigan Avenue, NE, Grand Rapids, MI 49503; 616-391-3588, fax 616-391-3674; E-mail brownm@msu.edu.

## REFERENCES

1. Hayden S, Brown M. Likelihood ratio: a powerful tool for incorporating the results of a diagnostic test into clinical decisionmaking. *Ann Emerg Med.* 1999;33:575-580.
2. Choi BC. Slopes of a receiver operating characteristic curve and likelihood ratios for a diagnostic test. *Am J Epidemiol.* 1998;148:1127-1132.
3. Gallagher E. Clinical utility of likelihood ratios. *Ann Emerg Med.* 1998;31:391-397.
4. Tandberg D, Deely JJ, O'Malley AJ. Generalized likelihood ratios for quantitative diagnostic test scores: slopes of a receiver operating characteristic curve and likelihood ratios for a diagnostic test. *Am J Emerg Med.* 1997;15:694-699.
5. Andersson RE, Hugander AP, Ghazi SH, et al. Diagnostic value of disease history, clinical presentation, and inflammatory parameters of appendicitis. *World J Surg.* 1999;23:133-140.
6. Snyder BK, Hayden SR. Accuracy of leukocyte count in the diagnosis of acute appendicitis. *Ann Emerg Med.* 1999;33:565-574.
7. Dujardin B, Van den Ende J, Van Gompel A, et al. Likelihood ratios: a real improvement for clinical decision making? *Eur J Epidemiol.* 1994;10:29-36.
8. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol.* 1991;44:763-770.

9. Harrison A, Morrison LK, Krishnaswamy P, et al. B-type natriuretic peptide predicts future cardiac events in patients presenting to the emergency department with dyspnea. *Ann Emerg Med.* 2002;39:131-138.
10. Pulliam PN, Attia MW, Cronan KM. C-reactive protein in febrile children 1 to 36 months of age with clinically undetectable serious bacterial infection. *Pediatrics.* 2001;108:1275-1279.
11. Fagan TJ. Nomogram for Bayes theorem [letter]. *N Engl J Med.* 1975;293:257.