Enhancing Clinical Decision-Making

Likelihood ratios: an intuitive tool for incorporating diagnostic test results into decision-making

Edouard Timsit DVM, PhD
Renaud Léguillette DVM, PhD
Brad J. White DVM, MSc
Robert L. Larson DVM, PhD
Sébastien Buczinski DrVet, MSc

Many clinicians use positive and negative predictive values to interpret diagnostic test results; however, these values have limitations that restrict their usefulness for this purpose. Likelihood ratios can overcome these limitations and provide a more intuitive way to incorporate results of a diagnostic test into clinical decision-making. Indeed, by incorporating the suspected probability that the patient has the disease before a test is even performed (ie, pretest probability of disease) as suggested by clinical experience (or previous test results), likelihood ratios can be used to easily estimate the likelihood that a patient truly has the disease given the test results (ie, posttest probability of disease) using the Bayes theorem. The purpose of the present article is to describe how likelihood ratios and Bayes theorem can be routinely used in a clinical setting.

Sensitivity, Specificity, and Predictive Values

Sensitivity and specificity have traditionally been used to evaluate the diagnostic accuracy of a test of interest. In this process, results of the evaluated test are compared against those of a reference or gold standard (ie, the test with an accuracy closest to 100%), generally by creation of a 2 X 2 table for a dichotomous test (Table 1). Sensitivity is the proportion of truly diseased individuals with a positive test result, and this value reflects how well the evaluated test detects a disease (or condition) when that disease is truly present. Specificity is the proportion of truly nondiseased individuals with a negative test result, and this value represents how well the evaluated test rules out a disease when that disease is truly absent. A sensitive test yields few false negative results in disease-positive populations, whereas a specific test yields few false positive results in disease-negative populations.

However, sensitivity and specificity have limitations in clinical use. Indeed, sensitivity answers the question “If an animal has the disease, what is the probability that the test result will be positive?” whereas specificity answers the question “If an animal does not have the disease, what is the probability of the test result being negative?” Yet when performing a diagnostic test, clinicians typically seek to answer the question “If the test result is positive, what is the probability this animal truly has the disease?” or “If the test result is negative, what is the probability this animal truly does not have the disease?” These questions can be answered by positive and negative predictive values, respectively.

Positive and negative predictive values can be also calculated from a 2 X 2 table (Table 1). These values correspond to the probability that an individual truly has or does not have a disease given a positive or negative test result, respectively. Unfortunately, predictive values do not easily translate into clinical practice because they are critically dependent on the population of interest and the prevalence of the disease in the setting in which they were derived. Indeed, positive and negative predictive values can decrease tremendously as prevalence of the disease in the tested population decreases or increases, respectively. Therefore, these values are not portable from one population to another (without additional calculations). Furthermore, because they reflect pop-

Table 1—Example of a 2 X 2 table showing how to calculate the diagnostic accuracy of a test with a binary outcome (ie, positive vs negative results).

<table>
<thead>
<tr>
<th>Test result</th>
<th>True disease status</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Disease present</td>
<td>Disease absent</td>
<td>Total</td>
</tr>
<tr>
<td>Positive</td>
<td>a</td>
<td>b</td>
<td>a + b</td>
</tr>
<tr>
<td>Negative</td>
<td>c</td>
<td>d</td>
<td>c + d</td>
</tr>
<tr>
<td>Total</td>
<td>a + c</td>
<td>b + d</td>
<td>a + b + c + d</td>
</tr>
</tbody>
</table>

Sensitivity = a/(a + c).
Specificity = d/(b + d).
Positive predictive value = a/(a + b).
Negative predictive value = d/(c + d).
LR+ = (a/(a + c))/(b/(b + d)) = sensitivity/(1 – specificity).
LR– = (c/(a + c))/(d/(b + d)) = (1 – sensitivity)/(specificity).

This is another installment in a series of articles first introduced in July 2015 to help practitioners enhance their ability to convert research information into clinical application.
ulation characteristics, predictive values cannot be readily applied to individuals. Finally, predictive values do not allow clinicians to rapidly test scenarios in which the pretest probability of disease differs from animal to animal. Fortunately, these limitations can be overcome by using likelihood ratios.2

**Likelihood Ratios and Bayes Theorem**

**Likelihood ratios**

A likelihood ratio is the percentage of diseased individuals with a given test result, divided by the percentage of nondiseased individuals with the same test result.2 For a test with a binary outcome (eg, positive vs negative results), 2 likelihood ratios are reported: one for a positive test result (LR+) and another for a negative test result (LR−; Table 1). The LR+ is the percentage of diseased individuals with a positive test result (ie, sensitivity), divided by the percentage of nondiseased individuals with a positive test result (ie, false positive fraction = 1 − specificity). In other words, the LR+ addresses the question “How much more likely is a diseased animal to have a positive test result, compared with a nondiseased animal?” The LR− is the percentage of diseased individuals with a negative test result (ie, false negative fraction = 1 − sensitivity), divided by the percentage of nondiseased individuals with a negative test result (ie, specificity). Consequently, the LR− addresses the question “How much more likely is a diseased animal to have a negative test result, compared with a healthy animal?”

Interpretation of LRs is intuitive; the larger the LR+, the greater the ability of a diagnostic test to rule in a disease or condition in a patient, whereas the smaller the LR−, the greater the ability of the test to rule out a disease or condition. It is generally accepted that a test with an LR+ > 10 is useful for ruling in a disease, whereas a test with a LR− < 0.1 is useful for ruling out a disease.3

Likelihood ratios have several attractive features, compared with predictive values. First, likelihood ratios do not change with prevalence because they are based on a ratio of sensitivity and specificity. Therefore, they are portable from one population to another or from one patient to another.2 Another interesting feature of likelihood ratios is that they can be estimated for diagnostic test results that are binary (positive or negative), involve > 2 ordinal categories, or are continuous (numeric). Consequently, clinicians can use likelihood ratios to interpret not only positive and negative test results but also strong positive, weak positive, weak negative, and strong negative results. Indeed, for tests that yield numeric data, such as serum biochemical values or serum antibody titers, likelihood ratios can be calculated for each of several cut points used to designate positive or negative results. Finally, if clinicians can estimate a pretest probability of disease for a given animal, then likelihood ratios can be easily applied to that animal with the aid of Bayes theorem.

**Bayes theorem and clinical application of likelihood ratios**

Given an animal’s signalment, history, and clinical examination or previous test results, clinicians can deduce an estimate of the probability of disease or the condition of interest in that animal. Then, they can decide whether to perform another test to increase this probability if it is too low to initiate a treatment (ie, the test bears a high risk of a false positive result) or too high to rule out the disease or condition (ie, the test bears a high risk of a false negative result).

Bayes theorem facilitates the translation of likelihood ratios and pretest probabilities into clinical decision-making in that the posttest probability of disease depends on the pretest probability and the result of the diagnostic test applied. Mathematically, Bayes theorem (using odds) is as follows:4

\[ \text{Posttest odds of disease} = \frac{\text{Pretest odds of disease} \times \text{LR}}{1 - \text{Pretest odds of disease} \times \text{LR}} \]

where LR is the likelihood ratio (LR+ or LR−, depending on the test result).

However, because odds are not intuitive for clinicians, odds are often converted into probabilities by means of the following formula:4

\[ \text{Odds} = \frac{\text{Probability}}{1 - \text{Probability}} \]

Bayes theorem can therefore be converted to the following formula:4

\[ P_{\text{posttest}} = \frac{P_{\text{pretest}} \times \text{LR}}{1 - P_{\text{pretest}} \times \text{LR}} + [P_{\text{pretest}} \times \text{LR}] \]

where \( P_{\text{pretest}} \) is the pretest probability of the disease or condition and \( P_{\text{posttest}} \) is the posttest probability.

In other words, Bayes theorem can be used to incorporate estimates from clinicians (or previous test results; ie, the pretest probability of disease), combined with the test’s characteristics (namely sensitivity and specificity, used to calculate the likelihood ratio), to answer the question “What is the probability of this animal truly having this disease now that I have a positive (or negative) test result?”

Although the posttest probability of disease in a given animal can be difficult to calculate in clinical settings using the preceding formulae, several online or mobile applications are available to make this easier (Figure 1). A Fagan nomogram, a graphical tool based on Bayes theorem, can also be used for this purpose.4 Such tools enable clinicians to go directly from a pretest probability of disease or condition to a posttest probability without having to convert probabilities to odds, or odds to probabilities.2

**Practical Example with Binary Diagnostic Test Outcomes**

As an example of the clinical application of Bayes theorem, consider a 6-year-old Holstein cow evaluated for anorexia and a severe decrease in milk production. Heart rate is 98 beats/min. Given the right-sided...
Tympanic resonance detected during physical examination, this cow is suspected of having right displacement of the abomasum (RDA) or abomasal volvulus (AV), which are difficult to differentiate before surgery. Because AV is a life-threatening condition and could lead to a negative outcome (ie, death or early culling) because of ischemic lesions to the abomasum, the producer wants to know the probability of a negative outcome before deciding on treatment (eg, surgery and supportive treatment). In this situation, clinical data can be useful for predicting (or diagnosing) this negative outcome.

Two types of clinical data have been shown to be useful for predicting outcome in Holstein cows with RDA or AV: blood lactate concentration and heart rate. For this example, data regarding the reported sensitivity and specificity of blood lactate concentration at various cut points have been used to calculate the LR+ (positive result defined as a negative outcome) for these cutpoint categories (Table 2). Data on the sensitivity (76%) and specificity (68%) of a heart rate ≥ 90 beats/min for predicting a negative outcome can also be used to calculate the LR+ for heart rate as a diagnostic test as follows: 0.76/(1 – 0.68) = 2.38.

Before Bayes theorem can be used to calculate the posttest probability of the cow in question having a negative outcome, the pretest probability needs to be estimated. For this estimate, we can turn to a previous study involving 102 Holstein cows with RDA or AV, in which the prevalence of a negative outcome was 21%. When we use 21% as the pretest probability and apply Bayes theorem, we get a posttest probability of a negative outcome for this cow of 38.7% (Figure 2). However, this posttest probability may not be informative enough for the producer to decide whether it would be worth investing in treatment (given the value of the cow and the price of milk), and performing a second test could help improve certainty in this regard.

Practical example with ordinal diagnostic test outcomes

Continuing with our clinical scenario, we can also use blood lactate concentration as an independent test to estimate the likelihood of a negative outcome for our cow. Using the same pretest probability as for heart rate (21%) and the values for LR+ obtained for blood lactate concentration at various cut points (Table 2), we can calculate the posttest probability of a negative outcome for each cutpoint category. In this situation, if the cow had a blood lactate concentration < 2 mmol/L, the posttest probability would be only 6.5%; if the concentration were between 2 and 6 mmol/L, the posttest probability would increase to 44.1%; and if the concentration were > 6 mmol/L, it would increase to 86.0%.

Practical Example with a Sequential Testing Approach

In the previous examples, heart rate and lactate results were considered as independent tests. However, another useful feature of likelihood ratios and Bayesian reasoning is that the posttest probability for a first test can become the pretest probability for a second test, provided both tests are conditionally independent (ie, the 2 tests results are not highly correlated). Consider this time a cow with RDA or AV, a heart rate of 98 beats/min, and a blood lactate concentration of 4 mmol/L. We
can use the posttest probability of a negative outcome from the heart rate test (38.7%) as the pretest probability in our estimation of the posttest probability following the blood l-lactate test, which becomes 65.2% (Figure 2). The combined likelihood ratio (LR+ or LR−) for ≥ 2 conditionally independent tests can also be calculated directly by multiplying the likelihood ratio for each, and that combined value can be used to estimate posttest probabilities (rather than calculating them in series as in the previous example). In the example here, the combined likelihood ratio would be 7.07 (2.38 × 2.97).

Limitations of the Bayes Approach

Despite the attractiveness of likelihood ratios and Bayes theorem for application in clinical practice, limitations exist that should be considered. First, few studies of diagnostic test accuracy have been reported in the veterinary literature. Second, existing studies generally involve small numbers of animals and imprecise accuracy estimates (ie, wide confidence intervals) that may limit the use of diagnostic tests for decision-making. When considering reported estimates of sensitivity, specificity, and therefore likelihood ratios, it is also important to think about potentially associated biases. For example, clinicians should consider the context in which a given diagnostic test was evaluated and whether that setting is similar to theirs, in which a different spectrum of disease or patients might be encountered. Finally, the likelihood ratio and Bayes approach is highly dependent on the pretest probability of disease, which can differ tremendously among clinical settings and is often unclear or unknown.

Clinical Summary

In summary, likelihood ratios have numerous advantages over predictive values. Because they are based on a ratio of sensitivity and specificity, they do not vary with prevalence and are therefore easily portable from one patient population to another. They can also be used with ordinal and continuous diagnostic test outcomes. Furthermore, they can be directly used at the individual patient level, provided clinicians have an idea regarding the pretest probability of disease for that patient. Indeed, with the Bayes theorem and associated tools, clinicians can easily calculate posttest probabilities of disease and test multiple scenarios based on various pretest probabilities.

Table 2—Likelihood ratios for the use of blood l-lactate concentration at various cutpoints for the prediction of a negative outcome (ie, death or early culling) in 102 Holstein cows with RDA or AV.

<table>
<thead>
<tr>
<th>Blood l-lactate (mmol/L)</th>
<th>No. (%) of cows with a positive outcome</th>
<th>No. (%) cows with negative outcome</th>
<th>LR+</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2</td>
<td>67 (66)</td>
<td>5 (5)</td>
<td>0.26</td>
</tr>
<tr>
<td>2–6</td>
<td>13 (13)</td>
<td>10 (10)</td>
<td>2.97</td>
</tr>
<tr>
<td>&gt; 6</td>
<td>1 (1)</td>
<td>6 (6)</td>
<td>23.14</td>
</tr>
<tr>
<td>Total</td>
<td>81 (79)</td>
<td>21 (21)</td>
<td></td>
</tr>
</tbody>
</table>

Likelihood ratios were calculated for this article from previously reported data.

Figure 2—Illustration of the use of the smartphone application in the example in Figure 2 to incorporate the posttest probability of a negative outcome for a cow with a clinical diagnosis of RDA or AV and a heart rate of 98 beats/min (ie, 38.7% in that figure; now the pretest probability of that negative outcome) with information on blood l-lactate concentration (4 mmol/L in this cow) to obtain a final posttest probability of the negative outcome. In this example, the 2 diagnostic tests (blood l-lactate concentration [Table 2] and heart rate) were considered independent, and rather than entering sensitivity and specificity of blood l-lactate concentration for predicting a negative outcome as was done in the previous example, the calculated LR+ for 2 to 6 mmol of l-lactate/L (2 mmol/L < L-lac < 6 mmol/L) was entered instead. See Figure 1 and Table 1 for remainder of key.
Acknowledgments

The manuscript was not supported by any grant nor presented at any meeting.

The authors thank Dr. John Kastelic for editorial assistance.

Footnotes


References