

EPIDEMIOLOGY: DIAGNOSTIC TEST USE

A **diagnostic test** is any procedure used to rule a differential diagnosis in or out, which includes patient history details, physical exam findings, imaging results and specific treatment response as well as conventional diagnostic tests. A test can be a combination, such as a rectal temperature threshold, a clinical appearance score and a cattle source (e.g., sale barn). The goal is to increase diagnostic certainty sufficiently to take the next step, such as making a prognosis or applying a treatment. In addition to direct material and labor costs, test costs include the risks of the test procedure to patient and the costs of improper actions (i.e., not doing something that should be done or doing the wrong thing) due to false test results. Costs of false positive versus false negative tests are often significantly different.

Test selection depends on cost, risk to the patient, turnaround time and opportunity for lab error as well as basic test performance. Highly technical tests sometimes have high batch-to-batch variation, are sensitive to laboratory conditions and operator skill and require concurrent positive and negative controls. Pressure to reduce cost often leads to less than optimum use of replicates and controls, even in commercial kits. Some laboratory certification programs are not blind, resulting in overestimation of laboratory performance.

The clinically useful test outcome measures are the situation-specific **predictive values**, which are presented as probabilities or odds ratios. **Predictive value positive (Pvp)** is, given that the test is positive, the probability that the animal has the condition. **Predictive value negative (Pvn)** is, given that the test is negative, the probability that the animal does not have the condition. Pvn and Pvp depend upon the likelihood that the particular animal has the condition before testing and the intrinsic performance characteristics of the test in that patient population (e.g., acute cases vs. chronic cases, subclinical cases vs. clinical cases, competing conditions).

The **prior probability** or **prevalence (Pr)** is the clinician's estimate, based on other information (e.g., history, exam findings, typical herd prevalence), of the probability the animal has the condition before the test. Without specific prior information, the best estimate is the expected condition **prevalence** in groups of similar animals.

The information gain is the difference between the prior probability and the appropriate predictive value. Because diagnostic tests are imperfect, meaning false negative and false positive results occur, the information gained depends in a non-intuitive way on the prior probability or estimated prevalence of the condition before running the test and on the performance characteristics of the test in such animal populations. **Baye's Theorem** underlies the mathematics.

To remove the effect of prior probability or prevalence, test performance is reported as two probabilities, sensitivity and specificity. **Sensitivity (Se)** is the probability the test is positive in an animal that has the condition. Sensitivity depends on the patient's location in disease natural history (e.g., incubating, subclinical vs. clinical, acute vs. chronic, mild vs. severe, recovered). This is different from **analytic sensitivity**, which is the ability of the test to detect the factor being measured, such as an antibody or antigen, a biochemical or an infectious agent, and is denoted by terms such as detection threshold.

Specificity (Sp) is the probability that the test is negative in an animal that does not have the condition. Specificity depends on the range of competing conditions in animals to which the test is applied.

When a **rule-in test** is positive, the animal very likely has the condition but the test may have a high proportion of false negatives. In a diagnostic situation, high specificity tests are "rule-in" tests. A pathognomonic sign is such a "test." When a **rule-out test** is negative, the animal very likely does not have the condition but the test may have a high proportion of false positives. High sensitivity tests are "rule-out" tests.

For most tests, sensitivity and specificity are inversely related, meaning changing the cutoff to improve one lowers the other. If the test measure is linearly related to the disease, a reasonable assumption is that the likelihood of the animal's status is higher (more likely negative or more likely positive) than the predictive value of the result indicates if the result is some distance from the cutoff. For some test formats, such as competitive ELISA's, this relationship is not linear so distance from the cutoff doesn't supply additional diagnostic information. Because the patients used to develop and validate tests tend to be more severe, longer duration cases (academic referral practices) than those in the typical private practice situation, test performance may be lower than published estimates.

If the test has low sensitivity but high specificity and is measuring something that is highly variable, such as intermittent bacterial shedding, repeated testing over time can increase diagnostic certainty if the interval between tests is longer than the typical peak to valley shedding interval. Prior testing of a herd with treatment or removal of test-positive animals tends to reduce the performance of subsequent testing unless sufficient time has passed for the disease to advance in the remaining animals. When using a test for screening, the disease spectrums are those of cross-sections of subclinical animals rather than those individuals being diagnosed in clinical settings.

"Back-of-the-envelope" calculations using a two-by-two table for a binary (positive or negative) test are the easiest way to understand and to explain testing results to clients. As the relationships are counter intuitive, select several test performance combinations, such as 0.98 Se, 0.98 Sp (a very good test), 0.75 and 0.75, (a so-so test), 0.25 and 0.98, (common ELISA performance for subclinical infection), and then calculate the predictive value pairs for a range of prior probabilities or prevalences, such as 0.001 (screening for a rare condition), 0.01 (infrequent), 0.10, 0.50 (toss up), and 0.75 (likely). In the following table, the two rightmost columns represent the disease states (disease present, disease absent) and the two bottom rows represent the clinician's knowledge after a test is applied (test is positive or negative) and the 4 cells represent the four possible test outcomes.

Outcome of Diagnostic Testing for Target Disease <i>(The Clinician's Axis)</i>	Actual Disease Status - To be estimated by clinician <i>(Mother Nature's Axis)</i>	
	Target Disease Present	Target Disease Absent
Test is Positive	True Positive (Tp)	False Positive (Fp)
Test is Negative	False Negative (Fn)	True Negative (Tn)

The following table contains the formulas for each cell of the above table. If $N = 1$ (testing an individual animal), the numbers are the probabilities that the animal will be in one of the four interior cells above. If the clinician is testing a group of N animals with an estimated disease prevalence Pr in the group, the numbers in the cells are the approximate number of animals that will have each of the four test outcomes.

Test Result:	Target Disease Present	Target Disease Absent	
Positive	$Tp = Se \times (Pr \times N)$	$Fp = (1 - Sp) \times ((1 - Pr) \times N)$	No. of Test Positives (Sum of $Tp + Fp$)
Negative	$Fn = (1 - Se) \times (Pr \times N)$	$Tn = Sp \times ((1 - Pr) \times N)$	No. of Test Negatives (Sum of $Fn + Tn$)
	$Pr \times N$	$(1 - Pr) \times N$	N

$$Pvp = Tp / (Tp + Fp) = Se \times Pr / ((Se \times Pr) + (1 - Sp) \times (1 - Pr))$$

$$Pvn = Tn / (Fn + Tn) = Sp \times (1 - Pr) / (Sp \times (1 - Pr) + (1 - Se) \times Pr)$$

If you don't wish to do these calculations by hand, many Bayes calculators are available on-line.

Find test performance reports by entering "sensitivity", "specificity" and the name of the condition or the infectious agent into PubMed. Including the format, such as ELISA, will reduce the number of hits. Producers of licensed test kits will provide such information on request as they are required to develop it as part of the licensing process. The more reliable test evaluations follow the STARD (STANDards for the Reporting of Diagnostic accuracy) reporting guideline. Above all, look for evaluations performed in animals and under circumstances similar to those in which you intend to use the test.

Suggested Reading

[Journals]

Banoo, S., Bell, B., Bossuyt, P., Herring, A, et al. Evaluation of diagnostic tests for infectious diseases: General principles. *Nature Reviews: Microbiology*, 2010;S16-S28. doi:10.1038/nrmicro1523
http://www.nature.com/nrmicro/journal/v8/n12_suppl/nrmicro1523.pdf

Otter, A., Davies, I. Disease features and diagnostic sampling of cattle and sheep postmortem examinations. *In Practice* 2015;37(6):293-305 doi: 10.1136/inp.h2482

Theurer, M.E., White, B.J., Renter, D.G. Optimizing feedlot diagnostic test strategies using test characteristics, disease prevalence, and relative costs of misdiagnosis. *Veterinary Clinics: Food Animal Practice* 2015;31(2):483-493.

[Books]

Callan, R.J., ed. *Ruminant Diagnostic Medicine*. *Veterinary Clinics: Food Animal Practice* 2007;23(3):377-604.

Fletcher, R.S., Fletcher, S.W., Fletcher, G.S. *Clinical Epidemiology: The essentials*. 5th ed. Lippincott, Williams & Wilkins, 2014.

Hulley, S.B., Cummings, S.R., Browner, W.S., Grady, D.G., Newman, T.B. *Designing Clinical Research*. 4th ed. Lippincott Williams & Wilkins, 2013.

Thrusfield, M.V. Veterinary Epidemiology. 3rd ed. Blackwell Publishing Co, 2007.

[Internet Resources]

The Equator Network - Reporting guidelines for main study types <http://www.equator-network.org/>

NLM Research Reporting Guidelines and Initiatives: By Organization -
https://www.nlm.nih.gov/services/research_report_guide.html

Sergeant, E.S.G. EpiTools epidemiological calculators. AusVet Animal Health Services and Australian Biosecurity Cooperative Research Centre for Emerging Infectious Disease, 2016.
<http://epitools.ausvet.com.au>

Author John M. Gay

Consulting Editor Kaitlyn A. Lutz