Available online at www.sciencedirect.com
ScienceDirect
Theriogenology 70 (2008) 560–564
Theriogenology
www.theriojournal.com

Interpretation of diagnostic testing strategies for production and economic usefulness

R.L. Larson *

College of Veterinary Medicine, Kansas State University, 111B Mosier Hall, Manhattan, KS 66506, United States

Abstract

Veterinarians providing reproductive services use a variety of diagnostic testing methods, including physical examination, laboratory testing, diagnostic imaging, and performance record evaluation. The diagnostic end point may be a physical diagnosis of pregnancy, attainment of puberty, or adequate quality and quantity of sperm; furthermore, it may be a medical diagnosis of reproductive tract pathology, presence of an infectious pathogen, or abnormal hormonal status. Proper interpretation of test results requires an understanding of how sensitivity and specificity (as measures of test accuracy), and prevalence of the condition, affect the interpretation of an individual result. For many diagnostic questions, the proper use of more than one test, either in series or in parallel, allows veterinarians to optimize their diagnostic accuracy and the economic return for the testing strategy.

Keywords: Physical examination; Laboratory test; Sensitivity; Specificity; Prevalence

1. Introduction

Veterinary practitioners help clients with beef cattle meet a number of production and economic goals; these include having a high percentage of females exposed to bulls becoming pregnant (i.e. enhance fertility or minimize infertility), minimizing the effects of infectious, metabolic, and other disease processes that can cause reproductive loss, and minimizing the effects of infectious, metabolic, and other disease processes that can cause disease and death. To aid their evaluation of an operation's production and health, practitioners can use a number of tests for screening, monitoring, and diagnosing populations, as well as individual animals. Commonly utilized diagnostic procedures include the use of diagnostic laboratory tests for infectious disease agents (e.g., serology, immunohistochemistry, PCR, virus isolation, bacterial culture, etc.). These tests can be used to screen for infectious diseases in apparently healthy animals, as well as to investigate disease outbreaks. In addition, veterinarians use diagnostic procedures to examine individual livestock and entire operations to find and change avoidable risks, for breeding soundness examination of bulls and heifers, for body condition scoring, to diagnose pregnancy via transrectal palpation or ultrasonography, and for feed and ration evaluation (nutritional and toxicological).

2. Determining the usefulness of a diagnostic test

A valid question confronting veterinary practitioners is whether to use available diagnostic tests to screen a particular herd for a particular condition [1]. The input that one needs to arrive at a logical conclusion includes: epidemiologic data regarding the condition or disease, diagnostic test sensitivity and specificity data, disease or

* Tel.: +1 785 532 4257; fax: +1 785 532 4989.
E-mail address: Rlarson@vet.ksu.edu.

0093-691X/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.theriogenology.2008.04.010

condition epidemiology, and economic costs of the condition and its treatment [2]. Literature review and mathematical aids (e.g., computer spreadsheets), are the tools used to create the necessary outputs, including post-test predictive values of diagnostic tests, economic value of testing, sensitivity of the decision to the individual inputs, and the importance of individual inputs to the decision. These outputs are used to evaluate alternate diagnostic testing strategies in order to determine the best testing strategy, and to identify control points that will be monitored for change that can trigger a re-evaluation of the decision.

2.1. Sensitivity and specificity of diagnostic tests

Sensitivity and specificity are properties of a diagnostic test that are determined by comparing the test to a ‘‘gold standard’’. The gold standard is considered the true diagnosis and may be made using a variety of information, such as clinical examination, expert opinion, laboratory results, or postmortem results. Sensitivity is the proportion of known positive (gold standard-positive) samples that the test in question identifies as positive. Specificity is the proportion of known negative samples that the test in question identifies as negative. In other words, sensitivity answers the question, ‘‘How effective is the test at identifying animals with the condition?’’ and specificity answers the question, ‘‘How effective is the test at identifying animals without the condition?’’

Diagnostic tests try to separate two populations. One is abnormal, diseased, or has an undesired condition and the other population is normal, or has a desired condition. In most conditions of interest to veterinarians, the affected and unaffected populations overlap, based on available diagnostic tests; therefore, both laboratory and clinical examination tests must use an arbitrary cut-off to separate test-positive and test-negative populations. Where cut-offs for diagnostic tests are placed is very important when deciding into which of the two distributions or outcomes a particular animal or herd lies. Because diagnostic measurements of affected and unaffected populations overlap to some extent, sensitivity and specificity are inversely related; therefore, placing the cut-off is always a trade-off between the impacts of false negative and false positive results (Fig. 1). For test cut-offs not set by the practitioner (i.e. set by a test manufacturer or at a diagnostic laboratory), it is important for the practitioner to be informed, as they will have different diagnostic goals depending on the context (e.g., a screening situation versus a confirmatory diagnostic situation). Where one places a diagnostic cut-off is always a trade-off between false negatives and false positives, due to the overlap between normal and abnormal populations [3].

2.2. Prevalence

Prevalence is the proportion of the number of cases of a particular condition at a given time to the size of the population at that time. The known or suspected prevalence is the probability of the condition being present determined before the investigation begins. Unfortunately, there are limited published, timely prevalence data for most infectious diseases and reproductive conditions of interest to veterinary medicine. Each practitioner’s judgment, based on history and clinical examination of both individuals and the population, aided by any available prevalence

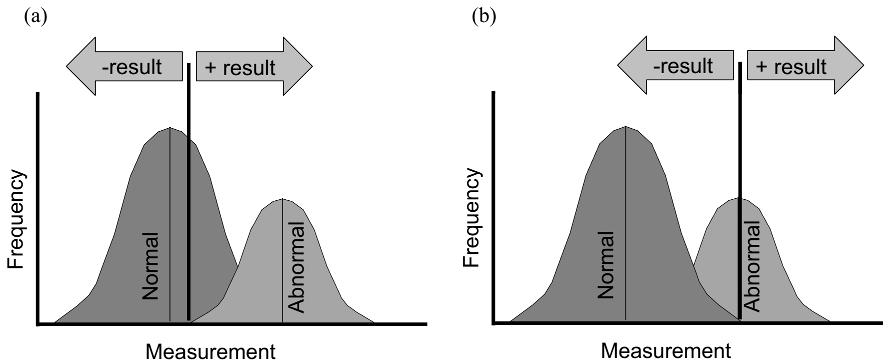

Fig. 1. When diagnostic measurements of affected and unaffected populations overlap to some extent, sensitivity and specificity are inversely related, and placing the cut-off is always a trade-off between the impacts of false negative and false positive results. (a) The cut-off is placed to maximize sensitivity (few false-negative results), but false-positive results occur. (b) The cut-off is placed to maximize specificity (few false-positive results), but false-negative results occur.

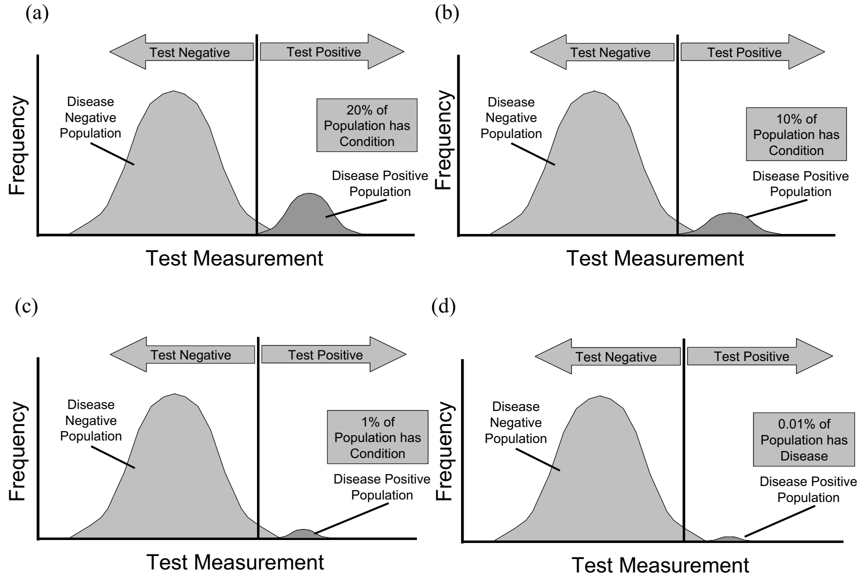

Fig. 2. If a test has imperfect test specificity (i.e. <100%), as the prevalence of the condition decreases, the positive predictive value of a positive test (PPV) decreases. In these examples, the test sensitivity and the test specificity are consistently 99.99% and 95%, respectively, but the prevalence of the condition differs. (a) If 20.0% of the population has the condition, the predictive value of a negative test is 100.0% and the predictive value of a positive test is 88.3%. (b) If 10.0% of the population has the condition, the predictive value of a negative test is 100.0% and the predictive value of a positive test is 67.0%. (c) If 1.0% of the population has the condition, the predictive value of a negative test is 100.0% and the predictive value of a positive test is 16.0%. (d) If 0.01% of the population has the condition, the predictive value of a negative test is 100.0% and the predictive value of a positive test is 0.2%.

data, are the basis to estimate the probability of infectious diseases and reproductive conditions.

2.3. Post-test predictive value

The post-test predictive values of a test are determined in the field (not the laboratory), and they tell a diagnostician if a valid test is useful. The positive predictive value is the proportion of animals with a positive test result that are actually positive and is influenced by test specificity. The negative predictive value is the proportion of animals with a negative test result that are truly negative and is influenced by test sensitivity. Both the positive and negative predictive values of a test are affected by the prevalence of the condition. As the prevalence of the condition is raised, we have more animals with the condition in the population, and we have greater confidence that a positive test result is correct. With increasing prevalence, the positive predictive value of the test is increased and the negative predictive value is decreased, whereas the reverse is true as the prevalence of the condition is decreasing (Fig. 2).

It is often not possible to estimate prevalence with any confidence, but one must still consider predictive value in test interpretation. When screening a herd, one often has no prior data to suggest that an individual animal is in a particularly high prevalence group. In such a mode, a negative test result has a high negative predictive value and is useful in striking a rule-out off the list, but a positive test result (which is most likely a false positive) is useful only in keeping a rule-out on the active list.

2.4. Diagnostic testing strategy

The interpretation of a test result is dependent on whether the test is assisting the decision to rule-out a preliminary diagnosis, or to screen for regulatory or biosecurity reasons; this is true whether one is considering an infectious disease agent such as *Brucella* sp. or Bovine Viral Diarrhea Virus, or a reproductive condition such as a non-pregnant cow or subfertile bull. To rule-in a preliminary diagnosis, many times it is necessary to use more than one test, either in series or in parallel. Running tests in series, where a second test is submitted only after the first test returns a positive result, is used to confirm a positive test with a low positive predictive value (low specificity or low prevalence). A two-test series is interpreted as negative if either test results in a negative response. Running two or more tests in parallel, where two or more tests are submitted at the same time, is used to confirm a positive test with low negative predictive value (low sensitivity or high prevalence). Parallel tests are interpreted as positive if either test results in a positive response.

3. Determining the cost of disease

Biology and economics intersect when veterinarians determine the cost of a negative condition. A number of tools are available to approximate the cost of a negative condition (disease presence, suboptimal body condition, non-pregnant after bull exposure, subfertile bull, etc.), and the biologic characteristics of the condition determine the proper economic analysis. Partial budgets are appropriate for diseases that are horizontally transmitted and immunity or other responses (death, sterility, removal from population, sale of feedlot pen, etc.) confine the negative effect of the disease to a short interval; or for conditions whose negative effects are confined to a short interval (i.e. correction of low body condition score effect on subsequent time interval's feed costs [4]). Multi-year enterprise analyses are more appropriate to estimate the economic cost of diseases that are vertically transmitted, due to an environmental source, have a chronic production-losing component, or to estimate the cost of conditions that have an impact on costs in subsequent years (e.g., non-pregnant cows, subfertile bulls, etc.).

4. Determining the economic benefit of a diagnostic strategy

Once the cost of the disease is determined (either by partial budget or multi-year enterprise analysis), the cost-effectiveness of alternate diagnostic testing strategies can be compared with a partial budget. In this partial budget, the post-test predictive values, test cost, cost of the negative condition, treatment cost, and cost of false positives are used to calculate the return for true positives, true negatives, false positives, and false negatives. The economic benefit is simply the costs for true negatives, false positives, and false negatives subtracted from the return for true positives.

When utilizing a two-test series strategy for screening for a condition in a population with low prevalence where initial positives are re-tested, the positive predictive value of the initial test becomes the prevalence used in the post-test predictive value calculations for the confirmatory test. By utilizing a confirmatory test on all initial test-positive animals, the positive predictive value for a two-test-in-series strategy is much closer to 100% than for a single test strategy, meaning that there are fewer false-positive animals and subsequently a lower cost of false positive test results. The cost advantage of a two-step screening strategy compared to a single test strategy is greatest when the prevalence is low.

For conditions that are rare (low prevalence), the post-test predictive value of a positive test will be low, meaning that there will be many false positive results. In contrast, for conditions that are very common (high prevalence), the post-test predictive value of a negative test will be low, meaning that there will be many false negative results, particularly if using a single-test strategy. In both situations, a partial budget evaluation is likely to indicate little or no economic benefit for a testing strategy. However, because some of the relatively infrequent negative conditions of interest to veterinarians can have substantial production and economic costs when present (e.g., infectious disease causes of infertility or pregnancy loss, subfertile bulls in single-sire pastures, etc.), the cost of an infrequent but important condition can be better evaluated as an assessment of risk and cost of risk avoidance. Once the cost of the risk is quantified, the producer and veterinarian can determine the effects such an event would have on a confined period's cash flow, and can evaluate that effect with the cost of risk reduction.

5. Summary

Interpretation of diagnostic tests is an important component of clinical veterinary practice. Veterinarians use a variety of testing methods, including physical examination, laboratory tests, diagnostic imaging, and evaluation of performance records. Proper interpretation of test results involves close communication with the diagnostic laboratory and other experts, and an understanding of how commonness (prevalence) of the condition in question affects ones interpretation of the test results. An important component of diagnostic testing is the proper use of more than one test, either in series or in parallel. By using properly one or multiple tests, veterinarians can optimize their diagnostic accuracy to provide the greatest value, and concurrently determine factors that are producing the greatest impact on the economic return for the testing strategy.

References

[1] Rothman KJ, Greenland S. Causation and causal inference. In: Rothman KJ, Greenland S, editors. Modern epidemiology. 2nd

ed., Philadelphia, PA: Lippincott Williams & Wilkins; 1998 p. 7–28.

[2] Clemen RT. Modeling uncertainty—Monte Carlo simulation. In: Clemen RT, editor. Making hard decisions: an introduction to decision analysis. Belmont, CA: Duxbury Press; 1991, p. 167–95.

[3] Romatowski J. Problems in interpretation of clinical lab tests. J Am Vet Med Assoc 1994;205:1186–8.

[4] Larson RL, Pierce VL. Agricultural economics for veterinarians: partial budgets for beef cow herds. Compend Cont Educ Prac Vet 1999;21:S210–9.