

## **Johne's Disease in Cattle - Article 4**

*This article was adapted with permission from the original 1999-2000 series presented by the AABP Food Safety Committee. Content was edited and reviewed by the National Johne's Working Group and endorsed by the USAHA.*

### **Concepts for Interpretation of Johne's Disease Diagnostic Tests**

#### **Part 1 of 4 on the topic of Johne's Disease testing**

Prepared and edited by Don Hansen and Christine Rossiter from the AABP Food Safety Committee and the National Johne's Working Group

**Background.** Experience teaches that diagnosis is often an imperfect process resulting in a probability, rather than a certainty, of being right. The rule-out process narrows the field but sometimes leads to "it's most likely- -" rather than "I am sure it is - -". Johne's disease is a good example of this situation. Common symptoms of Johne's, including diarrhea, weight loss and drop in production, bring to mind several possibilities for differential diagnoses, including indigestion, salmonella, parasitism, internal abscess or other chronic disease processes. When faced with such a medical challenge, veterinarians and producers rely on a diagnostic laboratory to help establish a diagnosis and to assess the potential for a herd problem.

At present, ELISA and fecal culture are the most common tests used to detect subclinical Johne's infection in herds. The choice of test and strategy depends on the degree of confidence desired, what is to be accomplished in each situation and cost. AGID, ELISA and fecal culture are all useful to diagnose Johne's disease in an individual cow with clinical signs.

As one ponders the results of a diagnostic test, a number of questions about the test should arise. How accurate is this test? How is accuracy measured? What is the level of confidence that a positive or negative test result is correct?

#### **How accurate is the test?**

For our purpose, accuracy is defined as the ability of a test to correctly identify the disease status of the animal tested. With a test that is 100% accurate, a non-diseased animal will always have a negative test result, and a diseased animal will always have a positive test result. If a group of animals is placed into categories by true disease or infection status and diagnostic test results, they will fall into one of four groups: (Figure 1)

**Figure 1**

		DISEASE STATUS	
		Present	Absent
TEST RESULT	Positive	True Positive A	False Positive B
	Negative	False Negative C	True Negative D

- A) Diseased and test positive.
- B) Non-diseased and test positive.
- C) Diseased and test negative.
- D) Non-diseased and test negative.

All animals will be placed into either category A or D with a perfectly accurate test. However, few diagnostic tests are 100% accurate. Sometimes the infected animal sheds too few organisms for the culture techniques to detect and the test gives a false- negative result.

Salmonellosis and *Staph aureus* mastitis are such examples. At other times a non-pathogenic "look alike" organism may be mistaken for a pathogen, such as may happen with trichomonad identification in a case of suspected trichomoniasis. This test result would be a false-positive.

Diagnostic test accuracy is based on two measures of performance: the **sensitivity** and the **specificity** of the test. Together these measures describe how well tests detect the true disease or infection status of animals being tested. Sensitivity is the ability of a test to detect infected animals correctly with a positive result. Specificity is the test's ability to detect non-infected animals correctly with a negative result. The greater the sensitivity and specificity of the test, the greater the accuracy.

**How are the sensitivity and specificity values of a test determined?**

In the process of test validation, sensitivity and specificity values are determined by evaluating the test independently in **two separate groups** of animals. All animals in one group are confirmed infected with a “gold standard” test or combination of tests. All animals in the other group are confirmed free of infection. Both should otherwise represent the general population of interest where the test will be used, i.e. (U.S.) cattle population.

The larger and more representative each group is, the better the validation and the greater the confidence that the values describe how the test is expected to work in infected and non-infected individuals.

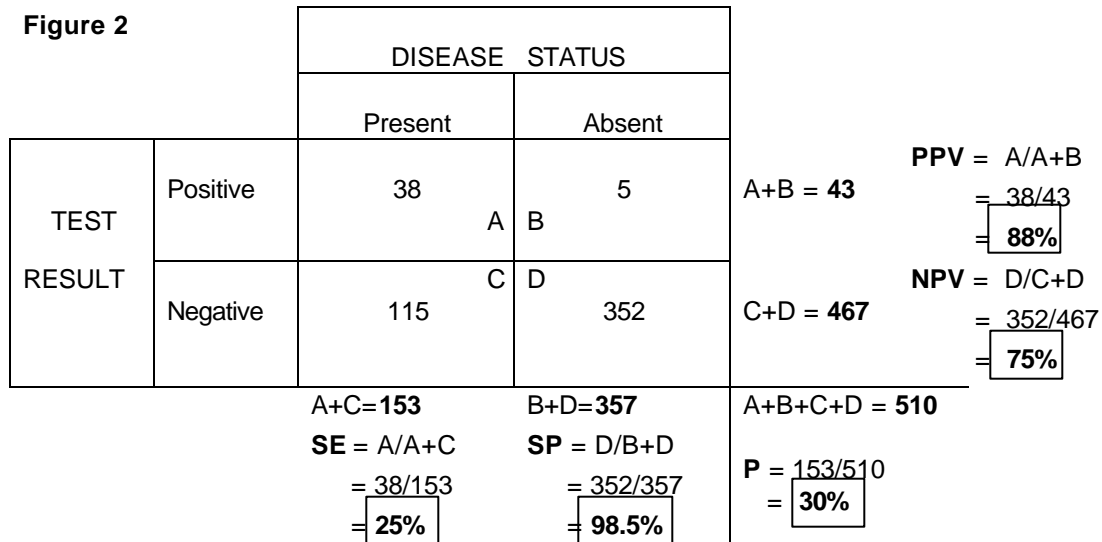
Sensitivity values for Johne’s tests are established by using the test in a well- described representative population of **infected cattle**. Their infected status must be confirmed by the best-available “gold standard” test(s) for *Mycobacterium avium paratuberculosis (Map)*. Preferably the infected cattle also represent the various stages of Johne’s infection. Because tests are unable to detect Stage I infection, these animals are often not represented in validation populations.

Sensitivity is the percent of all infected animals in the validation sample that have a positive test result (A/A+C, from Figure 1). Therefore, sensitivity (SE) becomes the expected probability or chance of receiving a positive test result in an infected animal. **Fewer false-negative test results are expected with higher sensitivity.**

See the 2X2 table in Figure 2 for an example calculation of sensitivity for an ELISA in a validation group of 153 infected cows that yields a SE equal to 25%.

Specificity values for Johne’s tests are established by testing a representative population of cattle that are **not infected** with *Map*. Prior to validation, all cattle in this group must be confirmed disease-free by history and negative herd tests (the best “gold standard” test).

Specificity of the test is the percent of the non-infected animals that correctly have a negative test result (D/B+D). Therefore, the specificity (SP) is the expected probability that the test gives a negative result in infection-free animals. **Fewer false-positive test results are expected with higher specificity.** See the example specificity calculation for an ELISA test in a group of 357 infection-free cows in Figure 2, which yields a SP of 98.5%.



The intracellular and slowly progressive nature of *Map* infection is largely responsible for the relatively low sensitivity of most Johne's tests. Recall the four stages of Johne's infection; current tests generally cannot detect early Stage I infection and miss many subclinically infected animals in Stage II.

However, tests will detect many infected cattle that are approaching (pre-clinical) Stage III and nearly 95% of clinical stage IV cattle. Thus, false negative test results are common for cattle in the early stages of Johne's infection, which includes mostly immature cattle.

The specificity values for ELISA and fecal culture are relatively high, meaning that false-positive results are relatively uncommon (0% to 3% of non-infected cattle). However, even though the number is low, it is a quandary to know how to interpret positive test results when they do occur in unknown or low-risk situations. Are they right or wrong? Most false-positive results come from interference by cross-reacting organisms, including *Mycobacteria* and *Corynebacteria*, which the animal was exposed to.

The population, distribution of stages of infection and the environment that the tested animals come from may influence the sensitivity and specificity of a test in a particular group of animals. This is true of published estimates and how they apply when the test is used in clients' herds. If only animals in early stages of infection (immature cattle) are tested, the sensitivity of Johne's tests is low (<20%). On the other hand, if all the animals are in Stage IV, the sensitivity is high (>90%). Published sensitivity values represent a population of infected animals in various stages.

Accurately defined Johne's disease populations are difficult to find. Therefore, it is important to recognize that sensitivity and specificity values for Johne's diagnostic tests are estimates that were determined from finite populations. Sensitivity and specificity of these tests does not necessarily represent all the cattle in the US, or all geographic regions of the country.

### A balance between sensitivity and specificity

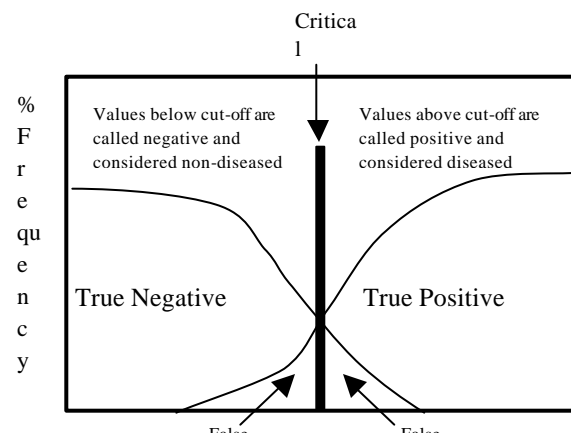
In the process of validating diagnostic tests that are measured on a continuous scale, e.g. antibody titers, a cut-off point is established and used to determine "positive" and "negative" results. The trade-off between sensitivity and specificity, when deciding the most accurate cut-off point, is inverse.

For example, an ELISA test produces a range of low to high values that are derived from Optical Density readings, which reflect antibody levels. At some point along this continuum, a cut-off is determined, above which the result values are considered positive. OD values below the cut-off point are considered negative. If the cut-off point is moved to a lower value on the scale, the sensitivity increases and more infected animals are identified as positive. However, if the specificity at the new cut-off decreases, more non-infected animals are identified as positive (See Figure 3).

Raising the cut-off point on the scale has the opposite effect. The specificity increases since only animals with high values are identified as positive. Non-infected animals with high enough values to be identified as positive (a false-positive result) are rare. But the sensitivity decreases at the same time, and more infected animals with elevated values in an intermediate range, are identified as negative.

Although serology tests usually report results as positive or negative, based on a single cutoff point, more information can be gained by examining how the actual values are distributed around the cut-off value. It is

useful to consider how the sensitivity, specificity and resulting decisions based on test results change when the cut-off is raised or lowered. Some test interpretations provide multiple cut-offs and probability of infection associated with each one. In practice, veterinarians also develop individual cut-offs to make decisions depending on the perceived accuracy of the test, the benefit or return if the decision is right, and the cost if it is wrong (See Figure 3).



**Figure 3.** Relative frequency of hypothetical titer values is graphed for two diseased and non-diseased populations. The critical cut-off point represents an optimal value for the diagnostic test.

### **What is the confidence in a positive test-result from a single animal?**

Even though the current Johne's tests are not 100% accurate, they are still useful tools for herd monitoring and disease control programs. However, when a positive test result returns from a lab, the question should arise, "Is this right?"

One useful way to assess the likelihood that a test-positive animal is truly infected, or a test-negative animal is truly not infected, is to consider the predictive value of the positive (PPV) or negative (NPV) test result. Predictive values can be calculated in a 2X2 table. However, it is actually an intuitive concept that takes into account:

1. The sensitivity (SE) of the test (% of infected cattle expected to test positive).
2. The specificity (SP) of the test (% of non-infected cattle expected to test negative).
3. The prevalence or expected number of infected animals in the test population before testing.

**What is prevalence?** Prevalence (P), also referred to as true prevalence, is the number of animals at a point in time that are infected with the disease or have the trait of interest, divided by the total number of animals in the population of interest. The apparent or test prevalence (AP) is the number of animals that test positive, divided by the total number of animals tested. Once determined, the AP can be used to calculate a more accurate estimate for P.

The calculation determines P from the AP by adjusting for the SE and SP of the test used. The simple formula is written as:

$$P = [AP + (SP - I)] / [SE + (SP - I)].$$

For example, after ELISA testing a herd, 5% of the herd has a positive ELISA result, so the AP is 5%. Using a sensitivity of 25% and a specificity of 98% for the ELISA (See Article 4 Part 2), the actual prevalence is calculated as follows:

$$P = [.05 + (.98 - 1)] / [.25 + (.98 - 1)]. \quad P = .13 \text{ or } 13\%.$$

The more accurate estimate of truly infected animals is therefore 13%, not the 5% that was detected as test-positive by ELISA.

### **Predictive value of a positive test**

The PPV is the percent of test-positive animals for which the result is correct or truly infected,  $A/(A+B)$ . It is affected by the disease prevalence in the herd, or the pre-test estimate of disease in the individual that is tested, i.e. a presumptive diagnosis or pre-test expected herd prevalence. The PPV is the probability that a positive test result agrees with true infection status. Its calculation requires estimates of SE and SP for the test and herd prevalence or probability of infection in an individual animal (P).

Prevalence and probability estimates are determined from existing information. Some options for making these estimates, in order of increasing accuracy, are:

1. Use an estimated prevalence based on herd history and risk assessment as estimated from pages A-2 through B-7 of the Manual for Veterinarians, Johne's disease Control Plan Guidelines or the NYSCHAP Johne's Disease Risk Assessment.
2. Use the presumptive diagnosis, i.e. an estimate of the probability that the animal may have Johne's disease or another disease(s).

3. Use a test prevalence estimate based on herd test results and calculate an estimate of the true prevalence (P).

The formula for calculating the predictive value of a positive test is:

$$PPV = SE \times P / (SE \times P) + [(1-SP) \times (1-P)].$$

For example, if SE = .25, SP = .985, P = .30, then PPV = 88%. At 30% prevalence, a positive test result has approximately an 88% chance of being correct. Another approach, which is more intuitive for some, is illustrated in Figure 4. It uses numbers in the original 2x2 table.

**How much confidence is there in a negative test-result for a single animal?**

As no commercial test for Johne’s disease can detect early-stage infections, it is not possible for a single negative test result to assure that an animal is not infected. Nevertheless, probability that a negative test result is correct, or the negative predictive value (NPV), can be calculated. For Johne’s tests, NPV values usually overestimate the true probability of being uninfected (because the SE of Johne’s tests is often overestimated and the numbers of test-negative animals are generally underestimated). However, the NPV values are still useful for individual animal and herd sample screening tests.

**Predictive value of a negative test**

The NPV is calculated as the percent of animals that are correctly identified with a negative test result. It also represents an estimate of the probability that a negative test result of an individual animal correctly indicates that it is not infected. Like the PPV, the NPV value for a negative test result varies depending on the prevalence or pre-test estimate of infection in the individual.

The NPV is calculated using the SE, the SP and a pre-test estimate of the prevalence of infection in the following formula:

$$NPV = SP \times (1-P) / [SP \times (1-P)] + [(1-SE) \times P].$$

For example, if SE = .25, SP = .985, P = .30, then NPV = 74%. At 30% prevalence, the negative test result has a 74% chance of being correct. See Figure 4 for the intuitive calculation example using numbers in a 2x2 table.

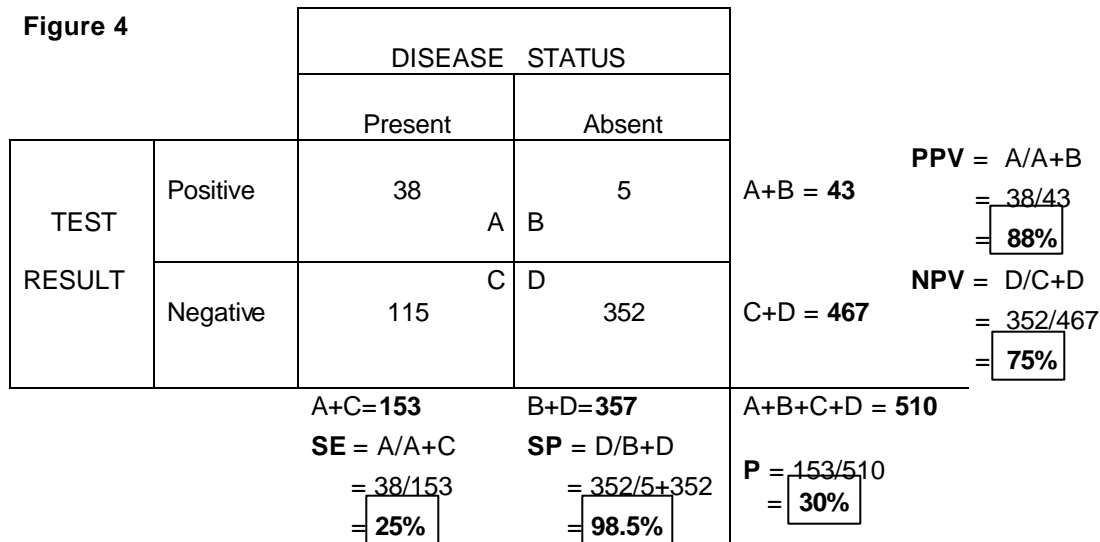


Table 1 has predictive values for several levels of prevalence and two reported estimates of SE for Johne’s ELISAs. Use the table to find the predictive values (PPV or NPV) for the ELISA or fecal culture tests based on estimates of the prevalence in the animal(s) tested.

Est. Prev.	ELISA25		ELISA60		CULTURE	
	PPV	NPV	PPV	NPV	PPV	NPV
1%	15%	99%	16%	99%	99%	99%
5%	48%	96%	51%	97%	100%	97%
10%	67%	92%	69%	95%	100%	94%
20%	82%	84%	83%	91%	100%	87%
30%	88%	75%	90%	85%	100%	80%
40%	92%	66%	93%	78%	100%	71%
50%	95%	57%	95%	71%	100%	63%

**Table 1.** For the generic Johne’s ELISA or fecal culture, the predictive value varies most dramatically with the prevalence (est. prev.) of infection in the herd. It can also be influenced by the SP of the test and sometimes by the test’s SE.

SE and SP values of 25% and 98% were used in the column labeled ELISA25. Those values were recommended by the National Johne’s Working Group to assess herds that participated in the Voluntary Johne’s Disease Herd Status Program (VJDHSP).

SE and SP values of 60% and 97% were used in the column labeled ELISA60. A commercial manufacturer of ELISA tests reports these values. When thinking about the performance of the ELISA across the ‘general’ population of infected cattle, these values should be considered an overestimate of the SE.

For fecal culture (CULTURE), SE and SP values of 40% and 99.9% were used. The National Johne’s Working Group for the VJDHSP also recommends these values.

Note that two important relationships between predictive values and disease prevalence (or probability estimates) may be used in practice.

For tests less than 100% specific, the predictive value of a positive test result decreases as the estimates of prevalence, or probability of infection decreases. In other words, the accuracy of a positive result is reduced with fewer infected animals.

Since all Johne’s tests are less than 100% sensitive, the predictive value of a negative test result decreases as the estimated herd prevalence of infection increases. The greater the number of infected animals, the greater the number of negative test results that are false-negative.

Understanding these relationships helps when interpreting positive and negative test results and assessing their potential impact for client’s decisions.

Understanding the importance of the predictive value of test results is an incentive for veterinarians to conduct a Johne’s risk-assessment with their clients. A working pre-test estimate of the herd prevalence or probability of infection is an important piece of information to evaluate the accuracy of individual test results and to better help clients control Johne’s disease. A good history and risk assessment, combined with clinical evaluation of individual cows, helps establish these estimates.

Veterinary Clinical Epidemiology by Ronald D. Smith, CRC Press, 1995 is a good reference for more discussion of predictive values and another predictor value called the Likelihood Ratio.