FLUORESCENCE POLARIZATION ASSAY (FP) TEST

Field Trial and Testing Results for Validation of the FP Test in Cattle, Bison, and Swine

Background:

The Cooperative National Brucellosis Program currently uses a number of tests in the serological diagnosis of brucellosis. Some tests, such as the Buffered Acidified Plate Antigen (BAPA) test, and the card test, are very economical and can be run in the field, yet are prone to subjectivity and variation between individual technicians. These tests, while relatively more sensitive than other tests, are lacking in specificity. Other tests, such as the Particle Concentration Fluorescence Immunoassay (PCFIA) and Complement Fixation Test (CFT), are not as sensitive. While they have the advantage of increased specificity, they must be run under controlled laboratory conditions. The PCFIA is comparatively more expensive, and the CF is a complicated test that requires a highly skilled technician.

The "gold standard" for the diagnosis of brucellosis is culturing of the bacteria from tissues from an infected animal. However, false negative culture results can occur for many reasons, including absence of the bacterium in the cultured tissues or insufficient numbers of the bacterium present to reproduce on growth media. Furthermore, some tests take days (e.g. Standard Tube Test) to weeks (e.g. microbiological culture) to produce a result, making them impractical for field testing or testing where livestock health authorities must make immediate decisions. These inadequacies can be largely overcome by use of many tests together, but they still often cause difficulties in test interpretation, especially when an immediate answer is needed. Consequently there is a constant quest for better tests.

There has been considerable interest in the development of a brucellosis test that is fast and easy to perform, economical, and is more accurate. The Fluorescence Polarization Test (FP) test, a test that has been used for years in human medicine, has been adapted for use in the field and the laboratory for the detection of brucellosis in cattle, bison, and swine.

APHIS has been evaluating the FP test since 1998 to determine whether it should be added to the battery of tests currently in use. The test was initially evaluated using established cut off points for positive and negative samples. Interpretation was subsequently modified to improve quality control, so that known positive and negative samples are tested first for machine calibration on every daily run, and the determination of a sample as positive, suspect, or negative is made based on a comparison to the known status samples.

The FP test was conducted in several laboratories to establish consistency between laboratories. It has been evaluated for use in cattle, bison, and swine. Results of the numerous test evaluation trials are included in this document. In addition, several versions of the instruments used to conduct the test were evaluated and validated to ensure that all instrument types utilized to conduct the test were equivalent.

System Overview:

Fluorescence polarization immunoassays were first developed in the 1970s, and are based on measuring the polarization of light caused by changes in molecular size as a result of antigen-antibody reactions. The technology has long been used in human clinical applications. FP is based on the rotational differences between a small fluorochrome labeled antigen molecule in solution and the antigen molecule complexed with its antibody. The theory is based on the fact that molecules naturally spin in a liquid medium. The rate of molecular spin is a function of the molecule's size. Larger molecules spin at a slower rate than smaller molecules do. A fluorescent dye label can be used to mark or tag a specific molecule. A beam of polarized light can determine quantitatively the rate of spin of the fluorescent molecule, and can detect any change in the rate of spin, and therefore the molecule's size. If the fluorescent dye tagged molecule (the reagent) finds and combines with the target molecule (antibody), the antigen-antibody complex that forms creates a larger molecule that spins at a slower rate. Fluorescence polarization therefore detects the binding of a tagged molecule to a target molecule.

The advantages of FP testing technology are many. The test is simple to perform, gives rapid results, is highly reproducible across laboratories and instruments, and reduces the human error and variability that occurs when reading agglutination tests such as the card test, the standard plate test, and other similar such test. The entire assay is done in solution, in a single tube with no precipitation or washing steps. Therefore, it is readily adaptable to field implementation.

Testing Overview:

The testing process is quite simple. A target sample is added to a test tube, and the first FP mP (millipolarization units) measurement is taken by the instrument. The test reagent is then added, and a second mP measurement is taken. The instrument calculates the polarization difference between the two readings. This change is compared to the values obtained on known negative samples to calculate the status of the test sample.

Validation Criteria:

The purpose of this validation is to determine the fitness of the fluorescence polarization (FP) diagnostic assay in diagnosing *Brucella abortus* in cattle, swine and bison. A validated assay should consistently provide test results that accurately predict the infection status of animals with some predetermined degree of statistical certainty.

The OIE has outlined 5 principles for validation of diagnostic assays (OIE reference). In summary, these are:

1. Determine the feasibility of the method.
2. Select the appropriate techniques and methods for performing the test.
3. Determine the performance characteristics of the test.
4. Continually monitor the performance of the test.
5. Maintain and enhance validation criteria during routine use.

The FP assay is a new test for brucellosis for which principles 1-3 are most applicable. Principles 4 and 5 are most applicable when validating diagnostic assays that are already in common use, but have not been previously validated.

By the time the USDA considers approving a diagnostic test as an official test, the test's feasibility and techniques must be thoroughly examined (Principles 1 and 2). Early feasibility studies should establish that the assay correctly identifies samples from known-infected and uninfected individuals. Early assay development and standardization also establishes the optimal concentration of reagents and measurement techniques such that the test agrees between replicates of known status. This early development work should also establish, through end-point dilutions, the lower limit of detecting antibody in serum (i.e., the analytical sensitivity). Furthermore, this work should explore potential cross-reactive antibodies that might result in false-positive results (i.e., the analytical specificity).

To approve a diagnostic assay, USDA must scrutinize its performance characteristics (Principle 3). According to OIE standards, it is desirable that 300 reference samples are examined from known-infected animals. The proportion of these samples classified positive by the diagnostic assay is an estimate of the test's diagnostic sensitivity. The OIE standards also suggest that 1000 samples from known-uninfected animals are also examined. The proportion of these samples classified as negative by the diagnostic assay is an estimate of the test's diagnostic specificity. Both diagnostic sensitivity and specificity depend on the selection of a cut-off value for an assay, such as the FP, with results on a continuous scale. When the frequency distribution of results from the infected and uninfected reference groups are compared, the cut-off value can be determined based on receiver operator characteristics analysis, visual inspection, or an algorithm that favors either sensitivity or specificity.

When selecting known-infected and known-uninfected samples, however, it is crucial that the samples represent the diversity of animal and environmental factors that can influence the test's performance. To incorporate this diversity, reference samples should be representative of the geographic area in which the test will be applied.

Another important performance characteristic is the diagnostic assay's precision. A precise assay results in a small amount of dispersion in its results for a repeatedly tested sample. This precision relates to repeatability within the same laboratory, and reproducibility between laboratories, for the same set of samples. As a general rule, OIE suggests at least 10-20 reference samples be examined repeatedly (2-3 repetitions) within the same laboratory, and between different laboratories, to evaluate precision.

Once a diagnostic assay is officially approved, there is a need to continually monitor the test's performance in the field (Principle 4). It is important that the predictive value of the test is evaluated so that users of the test can interpret the tests' results. For example, in very low prevalence situations the likelihood that a positive test indicates a truly infected animal (i.e., positive predictive value) tends to be very low, while the likelihood that a negative test indicates a truly uninfected animal (i.e., negative predictive value) tends to be very high. In contrast, a high prevalence situation results in a high positive predictive value (i.e., a high likelihood that a positive test indicates a truly infected animal) and a low negative predictive

value. Therefore, the classification of animals based on the diagnostic assay should include some consideration of predictive values.

Maintenance and enhancement of the diagnostic assay should also continue after the test is officially approved (Principle 5). Quality control is assessed via periodic proficiency testing of laboratories conducting the test. Quality of the reagents used in the assay must be frequently assessed. The test's performance should also be continually challenged under different circumstances and estimates of its diagnostic sensitivity and specificity revised to include the increased evidence.

This report considers the FP assay according to the first three OIE principles.

1. Feasibility of the method

The FP has been shown to be a highly accurate assay for detection of antibodies to *B. abortus* in bovine sera without detecting vaccination-induced antibodies (Nielsen et al., 1996[1]). A homogenous immunoassay, such as the FP, can be accomplished rapidly and does not require repetitive steps to wash away unbound reagents as other immunoassays require. The output of the test is objective because it does not require interpretation on the part of the technician running the sample. The ease and rapidity of this testing technology suggest it is highly adaptable to field application.

The theory of the FP technology is that large molecules rotate at a slower rate than smaller molecules, thus confining emitted light more to a single plane (more polarized) while small molecules rotate at a faster rate emitting more depolarized light. Polarized light intensity is measured both on the vertical and horizontal planes. The FP for *B. abortus* relies on an O-polysaccharide from *B. abortus* that is covalently linked with a fluorescein isothiocyanate tracer molecule. If antibody is present in a serum sample, a tracer-antibody complex would form, slowing rotation of the tracer and thereby increasing polarization of emitted light.

A preliminary study used the FP to test sera from infected and uninfected cattle in Canada (Nielsen et al., 1996). Infected cattle were those from which *B. abortus* had been cultured from milk or tissues. Serum samples were also assayed from *B. abortus* strain 19 vaccinated cattle at various times post vaccination. The FP test was shown to accurately classify uninfected cattle as negative, infected cattle as positive, and strain 19 vaccinated cattle as negative. This study also found that the FP test provided consistent results among repetitive analyses of four controls: strong positive, weak positive, negative and vaccinated.

[1] Nielsen K, D Gall, M Jolley, G Leishman, S Balsevicius, P Smith, P Nicoletti, and F Thomas. A homogeneous fluorescence polarization assay for detection of antibody to Brucella abortus. J Immunologic Methods 195:161-168, 1996.

2. Appropriate techniques and methods for performing test

The FP assay has been standardized to use a consistent concentration of reagents and measurement techniques such that the test agrees between replicates of known status. The process has been commercially developed by Viral Antigens, Incorporated, and licensed by USDA. Furthermore, the FP technology has already been developed for numerous other applications such as detecting illicit drugs and monitoring for drugs and other macromolecules.

Although the FP assay is a qualitative test, the nature of the technology provides some quantitative evidence of antibody levels in serum samples. A sample containing more antibodies will result in a larger FP reading relative to a sample containing fewer or no antibodies. A lower limit of detecting antibody was implied by two samples from known-infected cattle whose FP results were negative. In both of these cases, a small number of *B. abortus* bacteria were isolated from a single lymph node of the infected cattle (Nielsen et al. 1996).

A critical component of the FP assay is the antigen-tracer molecule conjugation. In one trial, six batches of this conjugate were prepared and tested. Results of the FP were equivalent regardless of the batch of conjugate used (Nielsen et al., 1996). The quantity of this conjugate to add to serum samples was determined by titrating until sufficient fluorescence intensity was achieved.

In repeated tests of sera of known status – strong positive, weak positive, negative, and vaccinated – little dispersion in FP results has been noted. For example, negative results averaged 75 mP, while 3 standard deviation units below and above the mean ranged from 60 mP to 90 mP (Nielsen et al., 1996).

Polarization readings are affected by temperature. It is reported that the polarization reading decreases by 3 mP for a temperature increase of 4 degrees C (Viral Antigens, Incorporated). To compensate for temperature effects, it is now recommended that negative controls should be run at regular intervals and sample results interpreted relative to the negative standard.

Four different fluorescent polarization reading instruments can be used to conduct the FP test. The Fluorescence Polarization Analyzer instrument (FPM-1, Jolley Consulting and Research Inc., Round Lake, IL) was initially used during FP trials. Subsequent work compared the FPM-1 with the Tecan Polarian, Tecan Ultra, and Sentry instruments. This work relied on serum panels developed by the USDA-ARS-National Animal Disease Center.

1) Comparing the FPM1 to the Tecan Polarian using bovine serum.

   Technicians from Missouri and Texas conducted trials on these two instruments using serum samples from cattle with known brucellosis status. Sixty nine culture-positive sera, including one RB51 culture positive serum, and one hundred fourteen negative control sera were evaluated. The culture

positive samples included 43 strain 19 samples, 1 RB51 sample, and 24 field strain samples, including biotypes 1, 2, and 4. The Tecan Polarian instrument was also evaluated using one culture positive specimen pipetted on all wells of one plate. This evaluation was to check the linearity and repeatability of the instrument.

The FPM1 apparatus uses a serum dilution of 1:50 and it was determined that the value at which the "suspect" range begins should be the value of the negative control plus ten calibration points above that. Since the Tecan Polarian FP testing protocol calls for a lower serum dilution of 1:20, making it more sensitive, it was determined that the positive range should start at 20 points above that of the average of the negative controls.

Both FP instruments detected all of the culture positives, for a sensitivity of 100%. The "reactor" reading for the Polarian averaged 160.58 points above that of the average of the negative controls; for the FMP1, 168.58 above the negative control was recorded. Both assays also picked up two negative control serums in the positive range; one at an average value of 51.51 above that of the negative control (in the "reactor" range), and the other at 11.9 above the negative control (a "borderline suspect"). Therefore, the specificity was calculated to be 98.3% for both instruments. Because the negative control samples were randomly drawn from an unknown serum pool of animals that were negative on all other serologic tests, it should be noted that those animals could have been incubating the disease and on the verge of spiking a titer on the traditional tests.

The Tecan Polarian instrument was unerring in its detection of animals from which field strain was cultured. Furthermore, the RB51 culture-positive serum sample was negative using both FP instruments. The difference in the test results between *B. abortus*, Strain 19, and *B. abortus* field strains (1, 2, and 4) was not statistically significant.

The repeatability and linearity of the test results across the microtiter plate of the Tecan Polarian was demonstrated by using the same culture positive sample in multiple wells of the microtiter test plate. One serum sample, which displayed titers of .06 on the PCFIA, 164 on the CF, and 250.9 on the Polarian was tested repeatedly in 92 wells. Out of the 96 wells on the microtiter plate, the remaining 4 wells were controls. The range of values for all runs was only 15.4 points, and averaged 254.41, or 169.06 calibrations above that of the negative control.

2) Comparing the FPM1 to the Tecan Polarion using bison serum.

Bison serum was tested using both the FPM1 and the Tecan Polarian instruments. There were 220 samples (50 culture positives and 170 culture negatives) tested on the FPM-1 and 189 samples (46 culture positive and 143 culture negatives) tested on the Polarion. In addition 218 samples were tested using the Buffered Acidified Plate Antigen (BAPA) test. For the FPM-1

instrument, sensitivity and specificity was 98.0 % and 96.4%, respectively. For the Polarion instrument, the sensitivity was 97.8% and the specificity was 96.5%. Among these samples, the BAPA test had a sensitivity of 86.0% and a specificity of 98.2%. Therefore, the two FP instruments demonstrated higher sensitivity and nearly equivalent specificity when compared with the BAPA test results.

3) Comparing the FPM1 to the Tecan Ultra and Polarion instruments using bovine serum; and comparing the FPM1 and Sentry instruments using bison serum.

Serum for these trials was obtained from animals used in NADC vaccination studies. This serum bank, while readily available, is not appropriate to use to calculate true sensitivity and specificity of a test, as some samples were collected very shortly after an animal had been challenged with strain 2308. Although there had not been sufficient time for the animal to mount a detectable antibody response, the challenged animals were considered culture positive. Therefore, the calculated sensitivity, specificity, and positive and negative predictive value will be considerably lower than in natural infected animals. However, this serum bank is appropriate to use to compare instruments, as the purpose of the trial is to observe consistency in readings across the various instruments.

i. Bovine Tecan Ultra and FPM1 instrument comparison

In this trial, the Tecan Ultra instrument was compared to the FPM1 instrument using bovine serum. A total of 62 culture positive and 27 culture negative bovine serum samples were evaluated. The sensitivities and specificities estimated using the two different instruments were equivalent on samples from these experimentally challenged animals.

ii. Bison Tecan Ultra, FPM1, and Sentry instrument comparison

In this trial, the Tecan Ultra and the Sentry instruments were compared to the FPM1 instrument using bison serum. There were 30 culture positives and 36 culture negative bison serum samples evaluated. The sensitivities and specificities were equivalent on all three instruments on the samples from these experimentally challenged animals.

3. Performance characteristics of the test

Testing Cattle

The performance of the FP in cattle has been extensively evaluated by independent researchers, as well as by the manufacturers of the test kits commercially available. In general, these evaluations have demonstrated that the FP rarely misclassifies uninfected cattle as positive. Therefore, this test has a high degree of specificity. Analysis of sera from known-infected cattle is confounded because the preferred gold standard of culturing *B. abortus* from tissues does not necessarily identify all infected cattle. Consequently, determining if the FP correctly classifies infected cattle as positive can be confusing because the true infection status of animals included in these studies is uncertain in some cases. Typically, animals positive on other serologic assays were assumed to be infected for the purposes of evaluating the FP. In some trials, the FP rarely misclassified so-called infected cattle as negative. Yet, in other trials some serologically-positive cattle were classified negative by the FP. Whether such results occurred because the FP incorrectly assayed these samples, or because the FP was correct while the other serologic tests were incorrect, is uncertain. In trials involving culture-positive cattle, however, the FP was shown to have a high sensitivity. Furthermore, most comparisons suggest the FP performs as well as, or better than, other serologic tools commonly used to diagnose brucellosis in cattle.

The FP was initially developed using reference sera from cattle known to be infected with *Brucella abortus* and from animals with no epidemiologic or serologic evidence of brucellosis. In a large multinational study, Nielson et al.[2] (1998) report their estimates of FP sensitivity and specificity from analyses performed in Argentina, Canada, Chile and Mexico, as well as the U.S. states of Iowa, Missouri, and Texas. APHIS did not include the data from Argentina and Chile because the cut-off values used in these countries were different than the others. Excluding these data, sera were tested from over 12,000 animals.

In this study, samples with results greater than 90 mP were considered positive. This cutoff value was based on receiver operator characteristics analysis of each dataset. This analytic algorithm essentially selected a cutoff value that maximized the sum of the resulting sensitivity and specificity. The procedure was repeated for all laboratories completing the analysis. Sera were judged to be from infected or uninfected cattle based on the results of buffered acidified plate antigen and complement fixation tests, therefore the results did not necessarily reflect true sensitivity and specificity. Instead of true sensitivity and specificity, the relative sensitivity (% FP positive of those positive to BAPA and CF tests) and relative specificity (% FP negative of those negative to BAPA and CF tests) were reported. Nevertheless, sera from 661 culture confirmed cases were analyzed by Canada with estimated sensitivity between 99% and 100% (Table 1).

[2] Nielsen K., D. Gall, M. Lin, C. Massangill, et al. Diagnosis of bovine brucellosis using a homogeneous fluorescence polarization assay. Veterinary Immunology and Immunopathology 66:321-329, 1998.

The results in Table 1 suggest that the FP correctly classifies serologically-negative cattle as negative, and serologically-positive or culture-positive cattle as positive. Although these results do not directly measure true sensitivity and specificity, the relative sensitivity and specificity are reasonable approximations.

Nielsen et al. (1998) also demonstrate that the FP does not cross react with antibodies to *B. abortus* strain 19. In Canada, 99% of 248 samples from cattle known to be vaccinated with *B. abortus* strain 19 were classified negative using the FP. In the Missouri laboratory, 87% of 241 such samples were classified negative using the FP.

Table 1. Summary of results from a multinational study of the FP assay as reported by Nielsen et al. (1998).

Source	Number of samples	Relative Sensitivity	Relative Specificity
Canada	8,669		100
	1,114		100
Mexico	255		97
U.S. – Iowa	7		100
U.S. – Missouri	433		97
	241		87
Total samples	10,719		
Canada	561[a]	99	
	100[a]	100	
Mexico	305	99	
U.S. – Iowa	29	65	
U.S. – Missouri	94	90	
U.S. – Texas	472	82	
Total samples	1,561		

[a] Sera were from cattle from which *B. abortus* was cultured.

Nielsen et al. (1998) also compared the performance characteristics of the FP with other serologic tests for brucellosis (Table 2). In most locations, the FP had the highest relative sensitivity and specificity of the tests examined. Nevertheless, the relative sensitivity estimated by the Iowa laboratory was lower than observed in the other locations. This result prompted re-examination of the protocol and equipment in the Iowa laboratory. It was determined that the reader used in those initial tests at the Iowa lab required a new light source, and one of the reading parameters, the lamp feed-back feature, was incorrectly set. Once the instrument was corrected, the trials were repeated and the relative sensitivity was estimated as 86% while the relative specificity remained 100%.

Table 2. Comparison of the FP assay to the card, buffered acidified plate antigen (BAPA), complement fixation test (CFT), particle concentration fluorescence immunoassay (PCFIA), and competitive enzyme immunoassay (CELISA) serologic tests for *Brucella abortus* completed in five locations. Source: Nielsen et al., 1998.

		Canada**	Mexico	U.S. - Iowa	U.S. - Missouri	U.S. - Texas
Relative Sensitivity	Card		98	89	82	
	BAPA	98		97		
	CFT	97		100		
	PCFIA	98		48	67	43
	CELISA	100				
	FP	99	99	65*	90	82
Relative Specificity	Card		69	100	34	
	BAPA	99		86		
	CFT	93		71		
	PCFIA	70		100	86	
	CELISA	100				
	FP	100	97	100	97	

* This result was based on a faulty instrument. After correcting the instrument, the relative sensitivity was estimated as 86%.
** Canadian sera were from known-infected and uninfected cattle. Therefore, the test parameters directly measure test sensitivity and specificity.

A commercially licensed kit for FP testing is marketed by Viral Antigens, Incorporated (VAI). To become licensed, this company completed trials in U.S. laboratories to evaluate the reproducibility of the test between labs and to characterize its performance in different populations of cattle. In contrast to previous research on the FP, VAI established cutoff values that reference the mP value of a negative control. A participating laboratory was instructed to complete three runs of a negative control serum and use the average of these results to characterize all other samples examined. Any sample reading 20 mP higher than the mean negative control mP is characterized as positive. Samples which assay between 10 and 20 mP higher than the negative control mean should be retested using 20 µL of sample (instead of the initial 10 µL). If the 20 µL repeat is > 20 mP higher than the mean negative control mP, it should be reported as positive. If the result is still in the 10 to 20 mP range above the mean negative control, it should be reported as suspect. Samples assaying < 10 mP above than the mean negative control are considered negative for *Brucella abortus* antibodies.

To evaluate reproducibility of results between laboratories, a panel of 11 sera was independently examined by laboratories in Missouri, Tennessee and Texas. The average reading for the negative control was 78, 84, and 87 in the Missouri, Tennessee and Texas labs, respectively. The three laboratories generated similar results for each of the 11 sera in the test panel (Figure 1). Results were reported as the difference between the sample mP value and the mP value of the negative control. Therefore, results could theoretically range from $-\infty$ to $+\infty$.

The Texas laboratory examined the same panel twice to determine the variability within their laboratory. There was a high degree of correlation between these independent runs of this serum panel suggesting that test results were highly repeatable within the same laboratory (e.g., Pearson's correlation coefficient was 0.994).

To further compare the three laboratories, another panel of 20 sera developed by USDA-National Veterinary Services Laboratory was assayed by the Missouri, Tennessee, and Texas laboratories (Table 3). This trial was also useful in comparing the performance of two polarization reading units; the FPM-1 used in Texas and the Sentry unit used in Missouri and Tennessee. The labs' FP classification agreed on all but two samples; samples 9 and 13 were classified as suspect by the Missouri lab while the Tennessee and Texas labs classified these samples as negative. It should be noted that suspect samples were not re-analyzed in this trial and it is likely that the Missouri lab would have re-classified samples 9 and 13 as negative if the 20 µL dilution had been assayed.

The FP assays completed at all three labs clearly identified the most sero-reactive samples (samples 4, 8, 11, 16, and 18 based on rivanol titers) as positive (Table 3). In fact, these samples had an average reading of 254 mP at the Texas laboratory compared to the average negative control reading of 88 mP. The FP also clearly identified the four sero-negative samples (samples 1, 7, 17, and 19) as negative. The average reading for these samples was 86 mP at the Texas laboratory. Nevertheless, there were eight samples that could be classified as sero-reactor titers from the NVSL serologic assays that were classified as negative on the FP analysis. Generally, these samples had low titers on the rivanol and complement fixation tests, and it is unclear if these samples originated from *B. abortus* strain 19-vaccinated cattle or *B. abortus*-infected cattle. If the samples reflected strain 19 vaccination, then many of the samples would be classified as negative according to 9 CFR part 78 and the FP results would agree with such a classification.

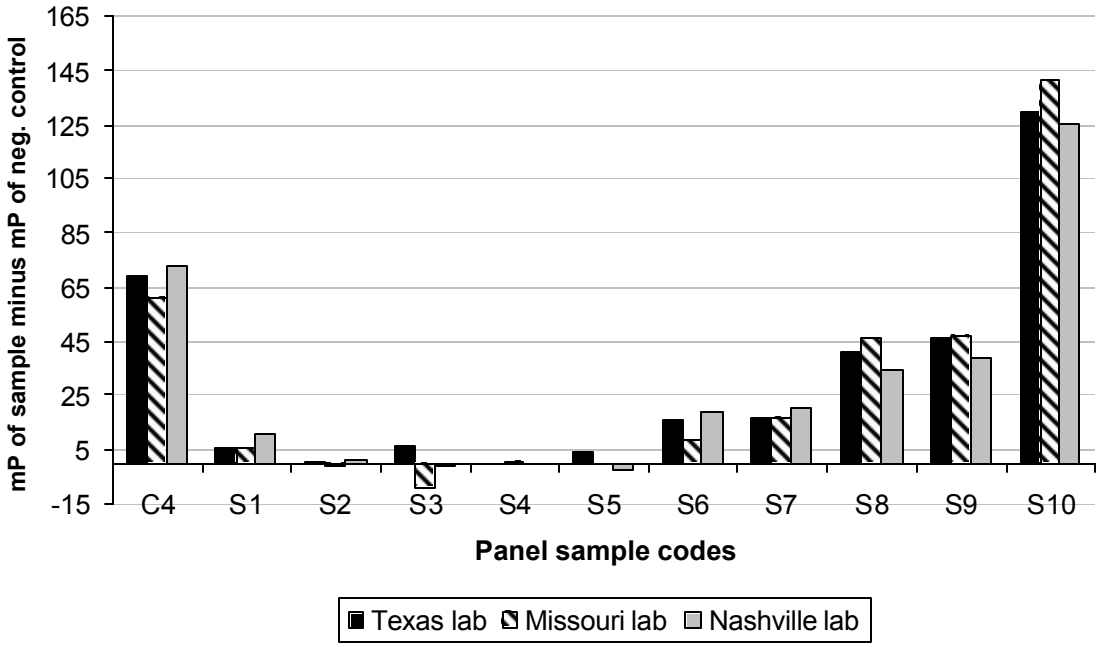

Figure 1. Comparison of results from three laboratories that analyzed the same 11 sera panel using FP.

Table 3. Comparison of FP analysis completed in Missouri, Texas and Tennessee laboratories with a USDA-NVSL serum panel subjected to a battery of brucellosis tests.

	Missouri	Texas	Tennessee	NVSL Tests				
Sample number	FP-Sentry	FP - FPM-1	FP - Sentry	BAPA	CARD	PCFIA	Rivanol	CF
1	N	N	N	N	N	N	-25	-10
2	N	N	N	P	N	N	-25	2+10
3	N	N	N	P	P	N	I25	3+10
4	P	P	P	P	P	P	+200	4+640
5	N	N	N	P	P	Suspect	I50	4+20
6	N	N	N	P	P	N	-25	2+10
7	N	N	N	N	N	N	-25	-10
8	P	P	P	P	P	P	+200	4+640
9	Suspect	N	N	P	P	N	-25	2+10
10	N	N	N	P	N	N	-25	2+10
11	P	P	P	P	P	P	+200	4+640
12	N	N	N	P	P	N	I25	3+10
13	Suspect	N	N	P	N	N	-25	4+10
14	P	P	P	P	P	Suspect	I25	2+20
15	N	N	N	P	P	N	I50	3+20
16	P	P	P	P	P	P	+200	4+640
17	N	N	N	N	N	N	-25	-10
18	P	P	P	P	P	Suspect	I200	2+40
19	N	N	N	N	N	N	-25	-10
20	P	P	P	P	P	Suspect	I25	2+20

A field trial of the FP was completed in Texas using submissions to that laboratory on September 10-18, 2002. A total of 500 serum samples were assayed using FP and PCFIA. Samples classified as negative to the PCFIA were assumed to be from uninfected cattle (n = 450), while samples classified as suspect or reactor to the PCFIA were assumed to be from infected cattle (n = 50). FP results were similarly classified as negative if the result was less than 10 mP units above the negative control, or positive if the result was greater than or equal to 10 mP units above the negative control.

Figure 2 illustrates the distribution of relative mP values for the so-called uninfected and infected populations in this Texas study. Excedance fraction is the probability that values are greater than the corresponding mP value. For PCFIA negative samples, the probability of a sample being greater than zero is nearly zero, yet 50% of the samples are greater than -6 mP (Figure 2). The abrupt, nearly vertical slope of the excedance fraction curve for PCFIA negative samples is indicative of a narrow probability distribution for this population. In contrast, the excedance curve for PCFIA positive samples has much more spread and indicates a broader probability distribution for this population. This curve suggests there is 100% probability that these samples are greater than 0 mP; about 50% probability that these samples are greater than 20 mP; and 0% probability that these samples are greater than 200 mP (Figure 2).

Excedance curves can be used to illustrate the specificity and sensitivity of alternative cutoff points. For example, a cutoff value of 10 mP in Figure 2 corresponds to an excedance fraction of 0% for the PCFIA-negative population. Therefore, none of these samples would be classified as positive by the FP at this cutoff value. Such a result implies perfect (100%) relative specificity. A cutoff value of 10 mP also corresponds to an excedance fraction of about 62% for the PCFIA-positive population. This means that 62% of samples from putatively infected cattle would result in FP readings of 10 mP or greater. Such a result implies a relative sensitivity of 62%. In general, the relationship between the excedance fraction of a particular cutoff value X, i.e., EF(x), and the relative sensitivity and specificity is:

$$\text{Relative sensitivity} = EF_{PCFIA+}(x)$$

$$\text{Relative specificity} = 1 - EF_{PCFIA-}(x)$$

From these Texas data it is evident that the cutoff values specified by VAI (i.e., 10 mP above the negative control for a suspect classification and 20 mP above the negative control for a positive classification) will maximize the specificity of the test and consequently avoid misclassifying uninfected animals as FP-positive. For the PCFIA-negative population, the average sample was about 6 mP below the negative control value. The standard deviation of this population was about 3 mP. A cutoff value can be conservatively selected to be at least three standard deviations greater than the average value from this negative population. If such a cutoff value were estimated from these Texas data, it would be 3 mP above the negative control [-6 mP + (3x3)] and the corresponding relative sensitivity and specificity would be approximately 70% and 98%, respectively. Therefore, these data could support a

lower cutoff value that would increase sensitivity somewhat at the expense of a slight reduction in specificity.

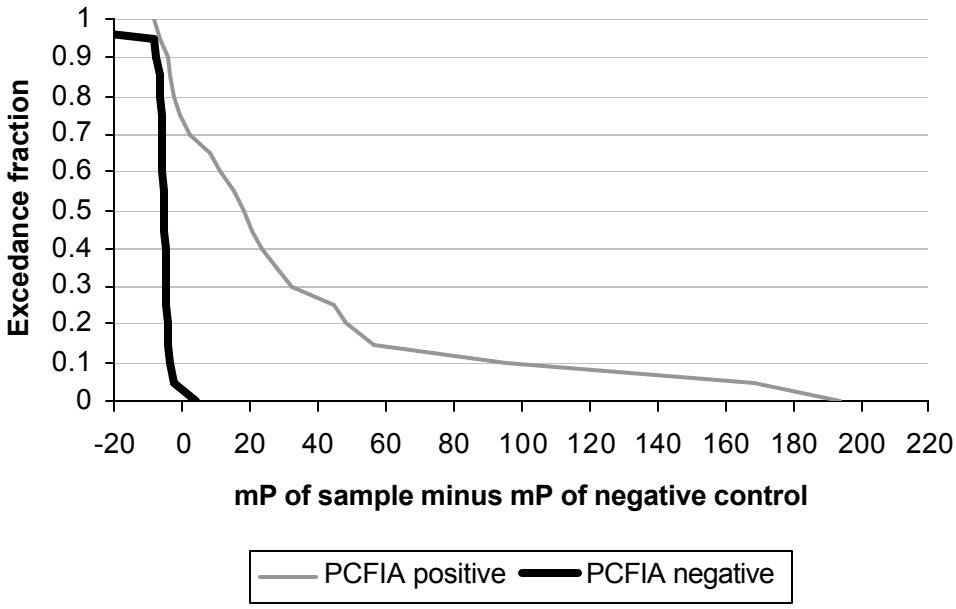

Figure 2. Excedance fraction curves of FP results (measured as difference between sample and negative control in millipolarization units) from 450 PCFIA-negative and 50-PCFIA positive samples assayed in Texas.

Another field trial of the FP was completed in Missouri using submissions to that laboratory on September 11-13, 2002. A total of 497 serum samples were assayed using FP, rapid agglutination plate test (RAP), and PCFIA tests. All these samples were classified negative on the RAP and PCFIA assays. Therefore, all samples were assumed to originate from uninfected cattle.

The excedance fraction curve generated from the Missouri data demonstrates zero probability of samples from uninfected cattle having FP results greater than 10 mP above the negative control (Figure 3). The average value from these data was about 5 mP below the negative control and the standard deviation was about 5 mP. If the cutoff value was three standard deviation units above the average, then these data suggest that 10 mP above the negative control is an appropriate cutoff value. In this case, the relative specificity of the FP would be 100%

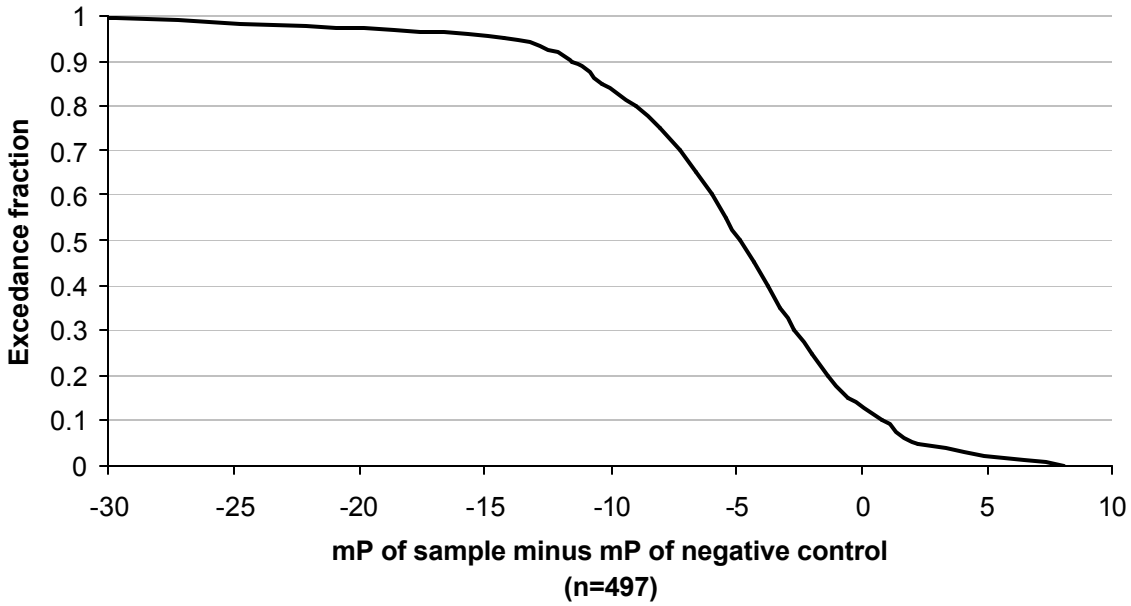

Figure 3. Excedance fraction curve of FP results (measured as difference between sample and negative control in millipolarization units) from 497 PCFIA-negative samples assayed in Missouri.

Another field trial of the FP was completed in Tennessee using submissions to that laboratory on October 2-3, 2002. A total of 100 serum samples were assayed using FP, rapid agglutination plate test (RAP), and card test. Although 20 samples were RAP positive, only two were confirmed positive using the card test. Samples that were negative to the RAP test or the card test were assumed to originate from uninfected cattle.

The exedance fraction curve generated from the Tennessee data demonstrates there is essentially no probability of samples from uninfected cattle having FP readings greater than the negative control (Figure 4). The average value from these data was about 11 mP below the negative control and the standard deviation was about 4 mP. If a cutoff value were three standard deviation units above the average, then these data suggest that 1 mP above the negative control would be the cutoff value between negative and suspect. In this case, the relative specificity of the FP would be nearly 100%.

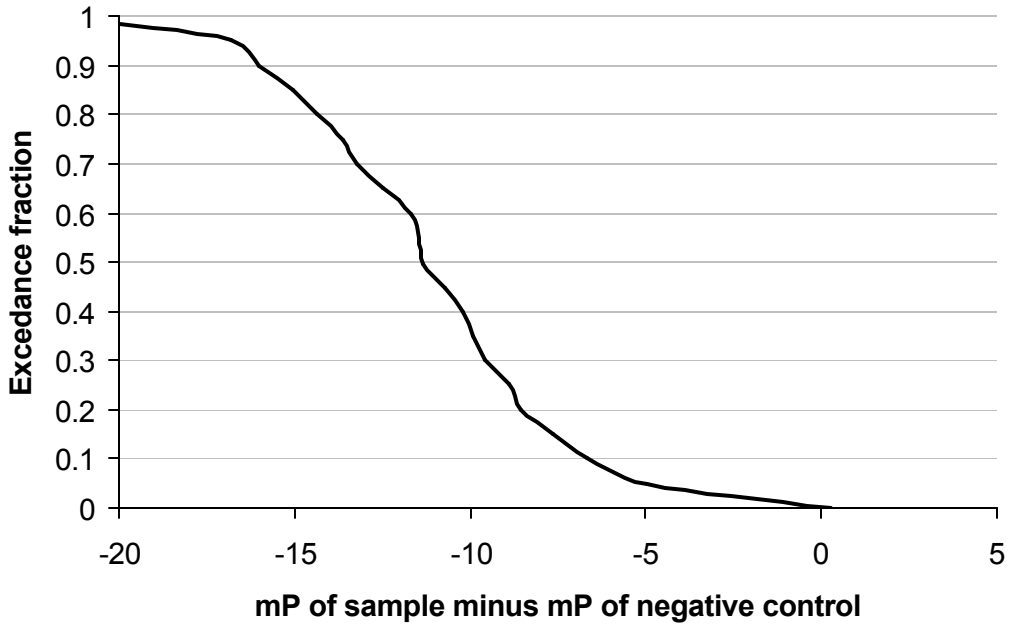

Figure 4. Excedance fraction curve of FP assay results (measured as difference between sample and negative control in millipolarization units) from 100 RAP or card-negative samples assayed in Tennessee.

Testing bison

The FP brucellosis assay has been extensively evaluated for use in bison. Like the FP in cattle, the interpretation of this test is confounded because the true infection status of bison is often unknown. Specificity of the FP assay (i.e., the likelihood that uninfected animals test negative) is best assessed in uninfected populations of bison. Evidence from known-uninfected populations in Canada and the U.S. demonstrates that the FP assay rarely misclassifies uninfected bison as positive. This high specificity is reassuring in populations of bison wherein infected and uninfected bison coexist. In these populations, FP-positive results will rarely include uninfected bison. Sensitivity of the FP test (i.e., the likelihood that infected bison test positive), however, is best assessed from bison within infected populations. For the purposes of evaluating sensitivity of the FP assay, bison can be considered infected based solely on culture, or a combination of culture and results of other serologic tests. Regardless of the basis for identifying an infected bison, the FP assay is shown to rarely classify infected bison as negative. In mixed populations of infected and uninfected bison, therefore, FP-negative results will rarely include infected bison.

To evaluate the use of the FP on bison serum, the FP was initially compared to other serological tests, including the buffered acidified plate antigen, complement fixation, and the indirect and competitive ELISA (IELISA and CELISA) tests on 38 samples from bison in which *Brucella abortus* had been isolated and 2,807 samples from

bison in areas with no brucellosis (Gall et al., 2000[3]). The sera from known-infected bison were collected from Wood Buffalo National Park (n=5), the U.S. (n=15), and Ontario (n=18). Using a cutoff value of 85 mP, this study determined that the FP assay had the highest specificity of the serologic tests (Table 4). Furthermore, the FP assay's estimated sensitivity was similar to the BPAT and CELISA sensitivities.

Table 4. Comparing the sensitivity and specificity of the FP assay with other serologic tests for brucellosis among infected and uninfected bison (Gall et al., 2000).

Test	%Sensitivity (n)	% Specificity (n)
BPAT	92 (38)	92 (1,000)
CFT	89 (38)	95 (2,807)
CELISA	92 (38)	98 (1,000)
IELISA	100 (38)	96 (1,044)
FP	92 (38)	99 (2,807)

In a subsequent blind study, bison sera from 214 animals of known culture status were tested following the same protocol. Of these samples, 54 were from bison in which *B. abortus* had been cultured, while the remaining 160 bison were culture-negative. The sensitivity and specificity of the FP assay estimated in this blind trial were similar to the estimates in the initial study (Table 5).

Table 5. A blind study comparison of sensitivity and specificity for FP assay and other serologic tests for brucellosis among infected and uninfected bison (Gall et al., 2000).

Test	%Sensitivity (n=54)	% Specificity (n=160)
BPAT	81	98
CFT	94	92
CELISA	96	94
IELISA	96	98
FP	96	98

Another study was conducted on serum collected from bison that exited Yellowstone National Park during the time period from February 1996 to March 2001 (Corso et al., 2003[4]). No specific criteria were used for selection of the animals in this study.

[3] Gall D., K. Nielsen, L. Forbes, D. Davis, et al. Validation of the fluorescence polarization assay and comparison to other serological assays for the detection of serum antibodies to Brucella abortus in bison. Journal of Wildlife Diseases 36(3):469-476, 2000.

[4] Corso B.A., JC Rhyan, LM Philo, IA Gardner, and MD Salman. Evaluation of Fluorescence Polarization Assay Test for Screening of Brucella Abortus Infection in Bison from the Greater Yellowstone Area, USA and A Privately Owned Bison Herd. In Proceedings of the 10th International Symposium on Veterinary Epidemiology and Economics, Vina del Mar, Chile, November 17-21, 2003.

They were generally animals that left the confines of YNP, and were subsequently tested and slaughtered according to state bison management plans. Serology and culturing of tissues were completed on 149 bison. These bison were classified as reactors, suspects, or negatives based on the results of a battery of standard serologic tests interpreted according to 9 CFR part 78. The serology used to classify bison, however, did not include the FP assay. There were 32 bison classified as serologic reactor and culture-positive; 82 bison classified as serologic reactor and culture-negative; and 35 bison classified as serologic negative and culture-negative.

By convention, FP results greater than 90 mP were considered positive in this study while results less than 90 mP were assumed negative. Furthermore, sensitivity was separately estimated from; 1) bison that were serologic reactors and culture-positive and, 2) bison that were serologic reactors but culture-negative. Specificity was only estimated from bison that were both serologically negative and culture-negative.

Excedance fraction curves for both serologic reactor groups are similar, but the mP values for the culture-positive group were consistently larger than the culture-negative group (Figure 5). For example, the average FP result from the sero-positive/culture-positive group was 261 mP, while the average from the sero-positive/culture-negative group was 237 mP. Using a 90 mP cutoff value, the sensitivity of the FP is estimated as 100% for both serologic reactor groups.

The excedance fraction curve for the sero-negative/culture-negative group demonstrates that the mP values for this group are substantially less than either of the serologic reactor groups (Figure 5). The average FP result for the sero-negative/culture negative group was 81 mP. Using a 90 mP cutoff value, the specificity of the FP is estimated to be about 83%. Nevertheless, the true status of the sero-negative/culture-negative group is uncertain because all bison from Yellowstone National Park are considered potentially exposed to brucellosis. Therefore, the specificity estimated from this population is suspect and likely an underestimate of the true specificity of the FP assay in bison.

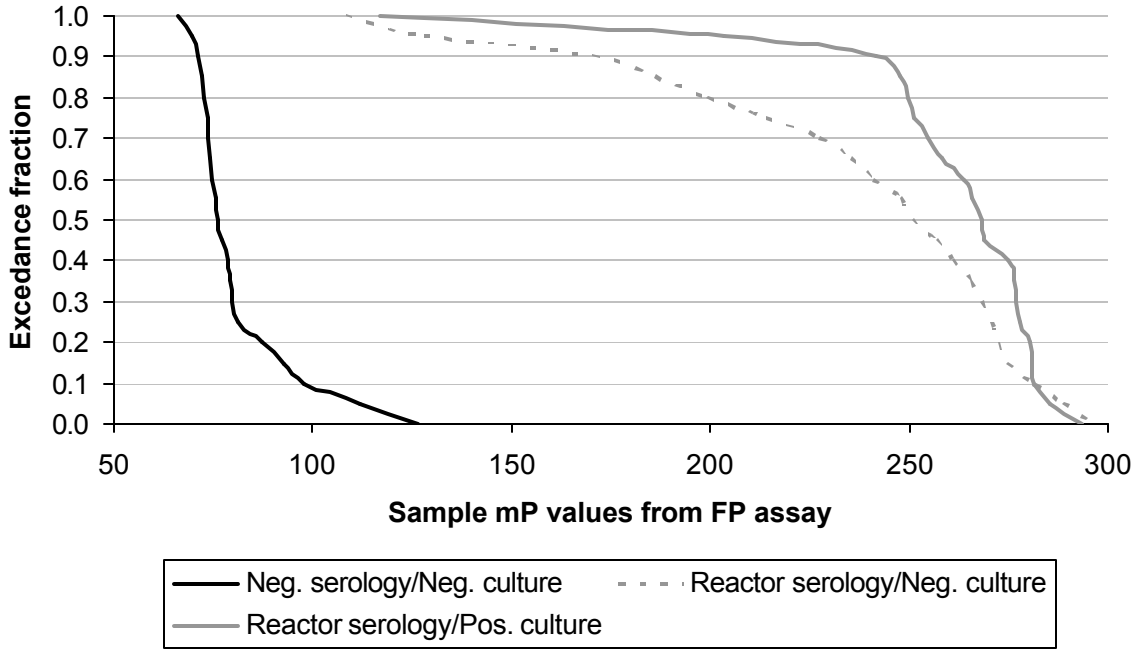

Figure 5. Excedance fraction curves from the FP assay results of 35 sero-negative/culture-negative bison, 82 sero-positive/culture-negative bison, and 32 sero-positive/culture-positive bison that exited Yellowstone National Park.

An additional 128 samples from bison in a historically brucellosis-negative herd were available for testing (Corso et al., 2003). These samples were all negative using a battery of tests including the BAPA, card, standard plate test, rivanol, and complement fixation. The excedance fraction curve for this set of samples predicts there is zero probability of sample values greater than 90 mP (Figure 6). Therefore, none of these uninfected bison would be misclassified as positive using the FP test. Furthermore, the results from the FP assay completely agreed with the results of the card test among these bison. Such findings are reassuring because the card test is a commonly-used screening test for brucellosis in bison.

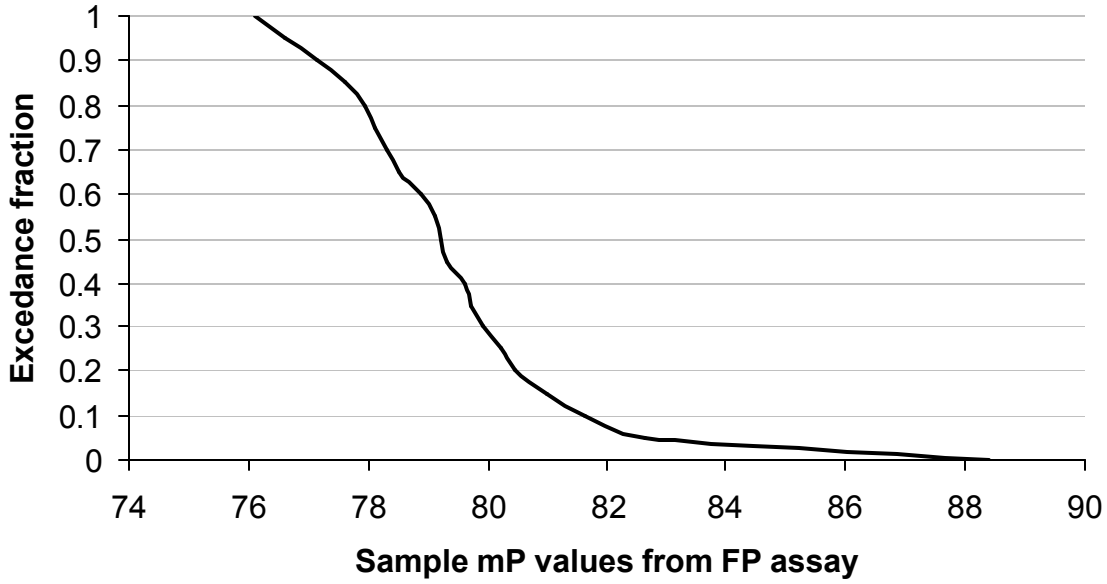

Figure 6. Excedance fraction curve of FP assay results from 128 sero-negative bison in a historically brucellosis-negative herd.

Between October, 2002, and March, 2003, the Montana State Veterinary Diagnostic Laboratory completed serologic examinations of 229 bison from Yellowstone National Park. This laboratory classified a sample as suspect on the FP if the result was between 15 mP and 20 mP higher than the negative control. Any sample reading 20 mP higher than the mean negative control mP was characterized as positive. Independent of the FP results, each sample was characterized as reactor, suspect or negative based on the 9 CFR part 78 interpretations of results of standard serology tests (i.e., the BAPA, card, standard plate test, standard tube test, rivanol, and complement fixation tests). To estimate relative sensitivity and specificity, a sample was assumed to originate from uninfected bison if the interpretation of standard serology was negative, otherwise the sample was assumed to originate from an infected bison.

The excedance fraction curve for those samples negative on standard serology (n=116) demonstrates a small probability that these samples will have mP values more than 15 mP above the negative control (Figure 7). Nevertheless, a small fraction of these samples had readings as high as 60 mP above the negative control. The excedance fraction curve for those samples classified reactors or suspects on standard serology (n=113) demonstrates that nearly all of these samples were more than 20 mP above the negative control. Using a cutoff value of 15 mP above the negative control, the estimated relative sensitivity and specificity of the FP assay were both 96%.

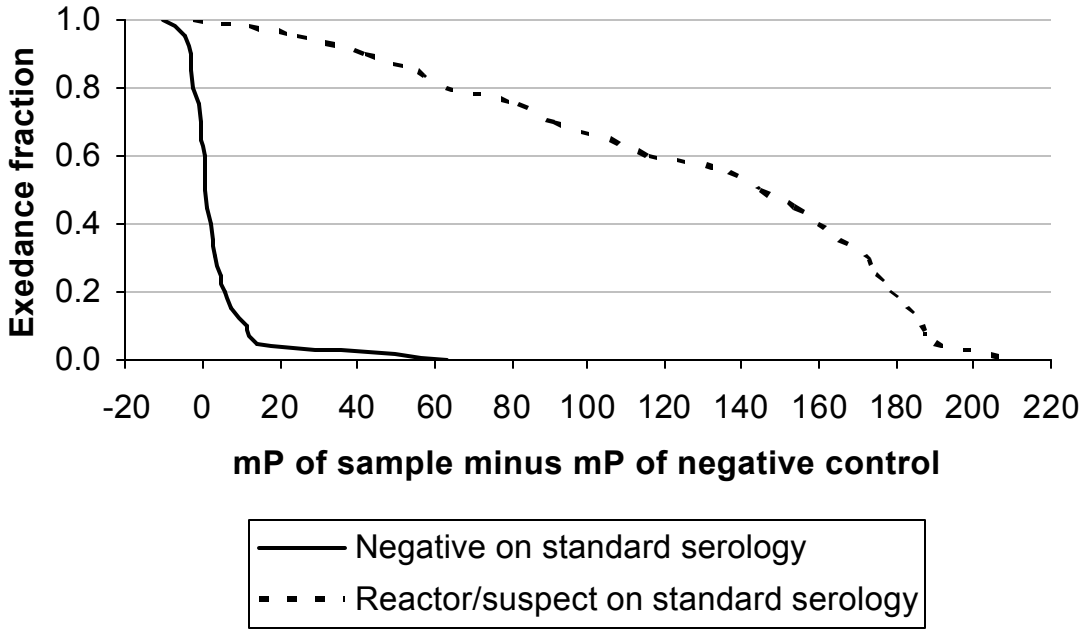

Figure 7. Excedance fraction curves of FP assay results from 116 bison that were negative to standard brucellosis serology, and 113 bison that were reactors or suspects on standard brucellosis serology. Samples were collected from Yellowstone National Park bison and analyzed at Montana State Veterinary Diagnostic Laboratory.

Testing swine

The conventional serological tests for brucellosis are generally not accurate in diagnosing brucellosis in swine. One of the reasons may be that the tests were developed for diagnosis on *Brucella abortus* in cattle and were not validated for the detection of *Brucella suis* in swine. Serologic tests are not accurate on an individual basis, and the conventional tests are only suitable for herd diagnosis. The complement fixation test is not as effective for diagnosis of swine brucellosis as it is for brucellosis in ruminants.

The FP assay was evaluated in a multinational study using sera from known-uninfected and known-infected swine (Nielsen K., 1999)[5]. A total of 401 samples from *B. suis* culture-positive swine from Argentina, Chile, Mexico, and the U.S. (Florida and Louisiana) were assayed using the buffered acidified plate antigen, complement fixation, indirect enzyme immunoassay (IELISA), competitive enzyme immunoassay (CELISA), and FP tests. A total of 14,037 samples from Canadian swine were similarly tested. Porcine brucellosis has never been diagnosed in Canada, so these samples were confidently assumed to originate from uninfected swine.

[5] Nielsen K, D Gall, P Smith, A Vigliocco, et al. Validation of the fluorescence polarization assay as a serological test for the presumptive diagnosis of porcine brucellosis Veterinary Microbiology 68:245-253, 1999.

An optimal cutoff value of 84 mP for the FP test was established using receiver operator characteristics analysis. The sensitivity and specificity of the FP assay corresponding to this cutoff value was 94% and 97%, respectively (Table 6). These parameters compared favorably with the other serologic tests completed on these samples. For example, FP sensitivity was comparable with the IELISA test and better than that estimated for the CELISA test.

Table 6. Comparing the sensitivity and specificity of the FP assay to other serologic tests for brucellosis among infected and uninfected swine (Nielsen, 1999).

Serologic test	Sensitivity (%) (n = 401)	Specificity (%) (n = 14,307)
BAPA	77	96
CF (ac +)*	93	95
CF (ac -)*	58	100
IELISA	94	98
CELISA	91	97
FP	94	97

* Complement fixation test with anticomplementary sera assumed to be positive (ac +) or negative (ac -) reactions.

The FP assay in swine was subsequently tested at NVSL using a panel of sera from 98 swine experimentally challenged intravenously with either *Brucella suis* biovars 1 or 3. Of these swine, 64 were culture positive for *Brucella sp*. while 34 were culture negative. All sera were examined using the buffered antigen plate antigen, rapid automated presumptive, card, complement fixation, particle concentration fluorescence immunoassay, and FP tests.

This panel was a challenging one in which to determine sensitivity and specificity because, in some cases, there was insufficient time for antibody development at the time the serum samples were collected. Therefore, the sensitivity and specificity are lower than expected on all tests (Table 7). Nevertheless, the FP assay had the highest specificity among all of the tests performed.

Table 7. Comparing the sensitivity and specificity for the FP assay with other serologic tests among experimentally *B. suis*-inoculated swine (USDA-NVSL). To estimate sensitivity, sera from culture-positive swine were evaluated. To estimate specificity, sera were from culture-negative swine were evaluated.

Serologic test	Sensitivity (%) (n = 64)	Specificity (%) (n = 34)
BAPA	61	71
RAP	55	74
Card	41	76
CF	86	44
PCFIA	27	79
FP	47	85

Conclusions:

The preceding evaluation of the FP test for brucellosis in cattle, bison, and swine satisfies principles outlined by OIE for validation of diagnostic assays. In fact, this evaluation of a serologic assay for brucellosis is unprecedented. Many of the standard brucellosis tests have been approved under much less scrutiny than this FP test. To its credit, the FP assay has a transparent trail of published research that documents its accuracy and repeatability when used for cattle, bison, and swine.

The data from the various studies discussed herein clearly show that the FP test performs as well as or better than the conventional serological tests. In addition, it also is faster to perform, and the automated process removes much of the human error that potentially occurs with standard agglutination tests. Therefore, the FP test is approved as a screening or confirmatory brucellosis test for cattle, bison, and swine. As with other brucellosis tests, however, the results of the FP assay should be considered in the context of the epidemiologic evidence when evaluating disease status of an animal or herd.

Continued monitoring of the FP assay subsequent to its approval as an official test is recommended by the OIE. With the continued use of this test in the U.S. will come improved understanding of its use in controlling and eradicating brucellosis. Future research objectives might include determining optimal combinations of tests for detecting infected animals and not detecting uninfected animals, as well as determining optimal cutoff values for different population-prevalence levels. For example, it may be feasible in the future to maximize positive or negative predictive values in high or low prevalence situations by adjusting the cutoff value used for the test.

As with any diagnostic test, the FP assay requires that users carefully adhere to the testing protocol. Quality control is standardized for the FP through the use of negative and positive control sera. At a minimum, these controls must be analyzed daily. Furthermore, as new instruments become available, their accuracy must be evaluated before they are adopted for use in this assay.