

Comparison of serologic tests for detection of *Brucella* infections in cattle and water buffalo (*Bubalus bubalis*)

Geoffrey T. Fosgate, DVM, PhD; Abiodun A. Adesiyun, DVM, MPH, PhD;
David W. Hird, DVM, MPVM, PhD; Wesley O. Johnson, PhD; Sharon K. Hietala, PhD;
Gerhardt G. Schurig, DVM, PhD; Joseph Ryan, BVSc

Objective—To estimate sensitivity and specificity of 4 commonly used brucellosis screening tests in cattle and domestic water buffalo of Trinidad, and to compare test parameter estimates between cattle and water buffalo.

Animals—391 cattle and 381 water buffalo.

Procedure—4 *Brucella*-infected herds (2 cattle and 2 water buffalo) and 4 herds (2 of each species) considered to be brucellosis-free were selected. A minimum of 100 animals, or all animals > 1 year of age, were tested from each herd. Serum samples were evaluated for *Brucella*-specific antibodies by use of standard plate agglutination test (SPAT), card test (CT), buffered plate agglutination test (BPAT), and standard tube agglutination test (STAT). A Bayesian approach was used to estimate sensitivity and specificity of diagnostic tests without the use of a gold standard, assuming conditional independence of tests.

Results—Sensitivity and specificity estimates in cattle, respectively, were SPAT, 66.7 and 98.9; CT, 72.7 and 99.6; BPAT, 88.1 and 98.1; and STAT, 80.2 and 99.3. Corresponding test estimates in water buffalo, respectively, were SPAT, 51.4 and 99.3; CT, 90.4 and 99.4; BPAT, 96.3 and 90.7; and STAT, 75.0 and 98.8. Sensitivity of the CT and specificity of the BPAT were different between cattle and water buffalo with at least 95% probability.

Conclusions and Clinical Relevance—Brucellosis serologic test performance varied by species tested, but BPAT had the highest sensitivity for screening cattle and water buffalo. Sensitivity and specificity of more than 2 screening tests can be estimated simultaneously without a gold standard by use of Bayesian techniques. (*Am J Vet Res* 2002;63:1598–1605)

these islands, and recognition of *B abortus*-infected livestock necessitated adoption of a test-and-slaughter program to control and eventually eradicate this infection from domestic animals. In Trinidad and Tobago, the buffered plate agglutination test (BPAT) was used as the official livestock screening procedure for brucellosis. This test has been studied extensively for use in cattle³⁻⁹ and has been used in many countries for control of brucellosis in livestock.¹⁰ The BPAT is 1 of the screening tests recommended by the International des Epizooties for diagnosis of brucellosis in cattle.¹¹

Brucellosis has been recognized in cattle and domestic water buffalo (*Bubalus bubalis*) populations in Trinidad.² Although *Brucella abortus* infection has been recorded in domestic water buffalo of other countries,¹²⁻¹⁷ there have been few reports^{17,18} critically evaluating brucellosis diagnostic tests in this species. In published studies,¹⁹⁻²⁹ testing procedures used for cattle were performed in water buffalo without verifying their diagnostic accuracy for this species. Accurate screening tests are important for success of brucellosis control programs, and these tests should be evaluated for target animal populations. Several screening tests have been used in brucellosis control programs in other countries, including the standard tube agglutination test (STAT), standard plate agglutination test (SPAT), and card test (CT),^a as well as the BPAT.^{3,30,31} The USDA has made recommendations for brucellosis testing protocols for cattle and bison,³² but no such guidelines had been reported for domestic water buffalo.

Development of screening test guidelines for brucellosis is difficult because of the nature of *Brucella* infection in livestock. Traditional methods for test evaluation require a perfect reference test (gold standard) for determining the true infection status of evaluated animals. Results of a gold standard test are used to place study animals into infected and noninfected groups. Accuracy, as defined by the sensitivity (probability of a positive result for an infected animal) and specificity

Trinidad and Tobago were classified free of brucellosis by the Office International des Epizooties in 1984.¹ These 2 Caribbean islands were considered free of this disease until the isolation of *Brucella abortus* in 1999.² No other *Brucella* spp were known to exist on

Received Feb 4, 2002.

Accepted Apr 30, 2002.

From the Departments of Veterinary Anatomy and Public Health, College of Veterinary Medicine, Texas A&M University (Fosgate). The Departments of Medicine and Epidemiology (Hird) and Statistics (Johnson), and the California Animal Health and Food Safety Laboratory (Hietala), School of Veterinary Medicine, University of California, Davis, CA 95616; the Faculty of Medical Sciences, School of Veterinary Medicine, University of the West Indies, Champs Fleurs, Trinidad and Tobago (Adesiyun); Virginia Maryland Regional College of Veterinary Medicine, Virginia Tech University, Blacksburg, VA 24061 (Schurig); and the Veterinary Diagnostic Laboratory, Ministry of Agriculture, Lands and Marine Resources, Champs Fleurs, Trinidad and Tobago (Ryan).

Results of this study were presented at the Conference of Research Workers in Animal Diseases, November 12, 2001.

Supported by the Center for Food Animal Health and Formula Funds from the Agricultural Experiment Station, School of Veterinary Medicine, and the Jastro Shields Memorial Foundation at the University of California, Davis, Calif.

The authors thank Nadira Seepersadsingh and Margit Pool for technical assistance.

Address correspondence to Dr. Fosgate.

(probability of a negative result for a noninfected animal) of the screening test can then be estimated directly. Diagnostic test evaluation is difficult for brucellosis because of the lack of a perfect gold standard.³³

Bacterial isolation of *Brucella* organisms from an animal is a good test to confirm infection, but bacterial culture results are often negative for infected animals,¹⁹ making it a poor criterion for noninfected animals. In addition, animals with positive bacterial culture results for *Brucella* organisms may not adequately represent the target population, because acutely infected animals are more likely to have positive culture results than chronically infected animals.³³ Results of serologic tests may also vary according to stage of infection, resulting in incorrect sensitivity estimates. Use of bacterial culture as a gold standard could also result in incorrect specificity estimates because of the misclassification of some infected animals as noninfected. Bacterial culture has other limitations because it often necessitates collection of specimens at slaughter, and extensive culling of animals is not always appropriate or feasible given local economic and agricultural conditions.

Methods have been developed to evaluate diagnostic tests when information from a gold standard test is not available.³⁴⁻³⁸ Analyses have been performed that assume conditional independence between tests and others that adjust for conditional dependence in sensitivity and specificity estimates. Conditional independence between diagnostic tests means that, conditional on true infection status, the result of the first test has no bearing on the probability that the second test will have positive or negative results. In other words, sensitivity and specificity of the second test will not be influenced by results by the first test (ie, positive, negative, suspicious). Diagnostic tests that measure similar analytes, for example serum antibody concentrations, may not be conditionally independent.^{35,39} The results of tests may be dependent because they are measuring the same or relatively similar quantities. An extreme example of dependent test results is that obtained by performing the same test on specimens in duplicate. The results would certainly not be conditionally independent. Conditional dependence in diagnostic tests can be accounted for by including sensitivity and specificity covariance terms in the statistical model.³⁵

Bayesian techniques have been used to estimate sensitivity and specificity of diagnostic tests without the use of a gold standard.³⁶⁻³⁸ These methods can be used for conditional independence models and can be adapted to include conditional dependence. They rely on the premise that researchers have prior knowledge (information) about test characteristics (sensitivity and specificity) and infection prevalence that was acquired before the current experiment or data collection process. This prior knowledge can then be modified or updated, using results of the current experiment. Updating prior information with experimental data results in formation of posterior summary information, formally obtained through use of Bayes' theorem, which is central to these methods.

The purpose of the study presented here was to estimate sensitivity and specificity of 4 commonly used brucellosis screening tests (SPAT, CT, BPAT, and STAT)

in cattle and domestic water buffalo of Trinidad. Secondary objectives were to determine significant differences in sensitivity and specificity estimates between cattle and domestic water buffalo, and to estimate *Brucella* infection prevalence in the tested animal populations.

Materials and Methods

Animal selection—During 1999, 8 herds in Trinidad were selected for evaluation of SPAT, CT, BPAT, and STAT. Four of these herds were composed exclusively of domestic water buffalo, and the other 4 composed of only cattle. The chosen herds were scheduled for sample collection as part of the active brucellosis surveillance system instituted by the government of Trinidad and Tobago. Reason for selection by the government was based on previous positive results for brucellosis or the necessity to determine brucellosis status in herds that had not previously been screened.

The eight tested herds included the 4 largest cattle and water buffalo production units on the island. One large Holstein dairy farm and 1 large water buffalo farm declined to participate in our study. Of the participants, herd 1 was from a small dairy farm raising Holstein and Holstein-cross cattle and was on the premises where brucellosis was first detected serologically on the island. Herd 2 was from a small dairy farm of Holstein and Holstein-cross cattle determined to be brucellosis-free on the basis of results of previous government screening. Herd 3 was also determined to be brucellosis-free on the basis of results of previous government screening and consisted of *Bos indicus* cattle raised for meat production under pasture management conditions. Herd 4 was from an extensively managed *Bos indicus* beef cattle farm considered *Brucella*-infected on the basis of cattle with positive serologic test results at previous government sampling. Herds 5 and 6 were extensively managed domestic water buffalo raised for meat and were considered to be brucellosis-free as a result of no history of abortions in these closed herds (although no previous testing had been performed). Herds 7 and 8 were extensively managed water buffalo raised for meat production with the same ownership, but in different areas of Trinidad. Both were considered infected with *Brucella* spp on the basis of results of previous government serologic testing. Subsequent work by our research team resulted in the isolation of *B abortus* from animals of 1 water buffalo and from both cattle farms considered to be infected at the time of sample collection.² Blood samples were obtained from water buffalo and cattle on study farms from April through June 1999 as part of the brucellosis-screening program instituted by the Trinidad and Tobago government.

Large farms used pasture management techniques, and animals were driven into holding pens the day of blood sample collection. Animals were sent through a chute and restrained for blood collection. Blood samples were collected from all animals > 1 year of age as they passed through the chute until at least 100 samples had been obtained. Multiple days were often necessary for collection, because facilities were inadequate for efficient animal restraint. Blood samples were collected from animals on small farms (< 100 animals) in a single day, and all animals > 1 year of age were included in our study. Whole blood samples were allowed to clot overnight at 4°C before centrifugation and harvesting of sera. All samples were stored in 3 aliquots at -40°C before initiation of serologic testing.

Diagnostic testing—Serum samples were tested for *B abortus*-specific antibodies by use of 4 brucellosis-screening tests in a blinded fashion by a single evaluator. The evaluator was not aware of the source of the sample (species and farm) and the results of previous serologic tests that may have already been performed. Sample identification numbers were

randomly assigned at the time of serum separation and storage. Serologic tests (ie, SPAT, CT, BPAT, and STAT) were performed by use of USDA S-1119 antigens and published USDA protocols.^b Serologic test results were determined in accordance with the Cooperative Brucellosis Eradication Program.³² Brucellosis vaccination had not been practiced in Trinidad, thereby eliminating complication of residual vaccinal titers on test interpretation.

General Bayesian method—We define θ as the collection of parameters (sensitivity, specificity, prevalence) we wish to estimate in our study. Prior scientific knowledge about θ can then be represented by the density function, $P(\theta)$. Posterior knowledge is formed as a conditional density for the unknown parameters given the data and is obtained via Bayes' theorem⁴⁰:

$$P(\theta | \text{data}) = \frac{P(\text{data} | \theta) * P(\theta)}{P(\text{data})}$$

where $P(\text{data} | \theta)$ is the likelihood function, $L(\theta)$, and $P(\text{data})$ is the probability of the data.

The denominator is constant for the analysis and is only important to ensure that the total probability on integration is equal to 1; therefore, the formula can be rewritten as follows:

$$P(\theta | \text{data}) \propto P(\text{data} | \theta) * P(\theta)$$

The posterior density is proportional to the product of the likelihood of the data, $L(\theta)$, multiplied by the prior density, $P(\theta)$. The mean or median of the posterior density can be used as a point estimate, and corresponding intervals can be generated to reflect variability around this estimate. A probability interval, analogous to a confidence interval but with a direct probability interpretation, is obtained by finding the values (lower, upper) where the area under the graph of $P(\theta | \text{data})$ between these points is 0.95 (or any other selected value). The true value is interpreted as being between these upper and lower bounds with 95% probability on the basis of the investigator's prior knowledge in conjunction with current data. These calculations can be performed by use of analytic calculus when $P(\theta | \text{data})$ is in the form of a recognizable density function (eg, corresponding to normal or β density functions).

Simulation techniques can be used to solve complicated statistical problems, including those that result in unrecognizable posterior densities for which analytic calculus is not possible. Markov chain Monte Carlo (MCMC) methods have been previously developed, specifically, to solve these types of problems.^{41,42} The MCMC samples are obtained iteratively, and resulting iterates (values for each parameter of interest) are dependent on previous selections, thus resulting in a chain of values. These MCMC values constitute a sample from the density we are interested in finding. When a large sample is obtained, medians and percentiles of the MCMC iterates serve as point estimates and probability intervals for the posterior densities, respectively.

General diagnostic test model—The basic model will be described for a simplified situation involving only 2 tests, STAT and BPAT, without a gold standard. The actual 4-diagnostic-test model used in our study is an extension of this model, but considers the results from 4 diagnostic tests simultaneously.

Lack of a gold standard creates a situation where sensitivity and specificity of diagnostic tests cannot be estimated directly, because true infection status of each animal is unknown. Available information includes the number of animals with each of the 4 possible positive (+) and negative (-)

test result patterns as follows: + STAT, + BPAT results; + STAT, - BPAT results; - STAT, + BPAT results; and - STAT, - BPAT results. The total count for each test pattern is the sum of the unobserved (latent) number of infected and noninfected animals for that category. Using the law of total probability, probability of belonging to a particular category is calculated through addition of the 2 components corresponding to infected and noninfected animals (**Appendix**).

For a single animal population, the observed 2×2 table of total counts (cross-classified test results) follows a multinomial distribution with unknown probabilities. Each cell probability is dependent on the infection prevalence and diagnostic test parameters (2 sensitivities, 2 specificities, and covariance terms). Probability for any given cell is a weighted average with contributions corresponding to infected animals (weighted by prevalence) and noninfected animals (weighted by $1 - \text{prevalence}$). Conditional dependence between diagnostic test results is allowed by including covariance terms in probability statements for each multinomial cell. Removal of these terms results in the conditional independence model. Separate covariance terms are fitted for infected (sensitivity covariance) and noninfected animals (specificity covariance).^{35,39} Sensitivity covariance is the difference between the probability of both tests having positive results for an infected animal and the product of the corresponding test sensitivities. Specificity covariance is similarly defined for noninfected animals on the basis of negative results on both tests. The range of possible values for these covariance terms is restricted, because all cell probabilities must be between zero and 1. The limits are based on sensitivity and specificity of the 2 tests,^{35,39} and appropriate constraints can be added into computer simulation models.

Prevalence of *Brucella* infection is an important component of the multinomial model, and this analysis accounts for variation in herd (population) prevalence by assuming independent multinomial samples from each herd (population). Prevalence is allowed to vary across populations, but diagnostic characteristics are assumed to be the same. Therefore, the evaluated tests are assumed to have equal accuracy across all tested populations. Estimation of diagnostic test parameters in this manner assumes that the prevalence of infection is different for each population. Violation of this assumption can lead to invalid estimates.^{34,35,38} To prevent this problem from occurring, the 2 herds of cattle considered to be brucellosis-free before sample collection were considered as a single population. The 2 herds of water buffalo considered to be brucellosis-free were also grouped together before analyzing the data. Therefore, the data were considered to be from 6 populations, 3 cattle and 3 water buffalo, and independent analyses were performed for each species.

Bayesian model—Beta probability distributions are often used to incorporate prior scientific information, $P(\theta)$, in analyses of diagnostic tests.³⁶⁻³⁸ The $P(\theta)$ for the present model were elicited to encompass prior knowledge for all unknown components (Table 1). Review of the literature identified a single study⁴ that reported sensitivity and specificity estimates for all 4 of the screening tests that we investigated. Methods used in that paper were excellent and were therefore used as a starting point for determination of the most likely values for sensitivity and specificity of the tests in our study. Point estimates reported in that study were slightly increased or decreased on the basis of knowledge of test procedures and the antibody isotypes detected by the tests. Upper and lower bounds for the sensitivity and specificity of all tests were then determined on the basis of the experience of the authors. True values for the test parameters were thought to be between these upper and lower bounds with 95% certainty. The most likely value was used as the mode, and the 95% intervals were used to determine spread

of the β distribution to allow for uncertainty in these point estimates. A commercially available computer program^c was used to plot β distributions to find those most consistent with the reported literature and the authors' experience.

Analysis of cattle and water buffalo data used the same prior probabilities for the diagnostic tests, but different population infection prevalences. The 6 populations of animals were considered to have zero, low, moderate, or high *Brucella* infection prevalence on the basis of the knowledge of the government veterinarians who provided routine and emergency medical services to these herds in addition to collecting samples for brucellosis screening. Interviews were performed before sample collection for our study, and prevalence information was therefore based on previous serologic testing. Herds that were not screened previous to initiation of our study were considered noninfected by government veterinarians because they were closed herds (no introduction of new animals) with no history of abortions. Modes of the prior prevalence distributions were assumed to be 0.00, 0.05 (low), 0.20 (moderate), and 0.33 (high). Beta prior probability distributions with large variances were chosen to account for the uncertainty in these estimates.

Simulation techniques using available software^d were used to estimate posterior probability distributions for all models under evaluation. The primary model was composed of information from all 4 brucellosis tests and 3 populations each of cattle and water buffalo. Diagnostic test results were considered from a multinomial sample of each population with 16 possible cells (4 dichotomous tests, 2⁴ = 16 test result patterns). The initial model assumed conditional independence of the diagnostic tests.^e

Computer code was designed so that water buffalo and cattle models were implemented independently, but they were nested so that at each iteration of the Monte Carlo procedure, values for all parameters in both species were generated. Sensitivities, specificities, and population infection prevalences were produced at each iteration. Cattle and water buffalo values were subtracted from each other at each iteration for all diagnostic test parameters. Distributions of these differences were used to evaluate significance between species' estimates. Within each species, sensitivities and specificities of the 4 tests were compared by subtracting pairwise iterate values (eg, difference 1 = BPAT sensitivity - CT sensitivity) and evaluating the resulting distributions of differences.

Plots of model parameter iterates (sensitivity, specificity, and prevalence) were monitored for trends in successive iterations to determine when convergence was achieved. Convergence is the point when parameter values (iterates) no longer drift over successive iterations and appear to be fluctuating randomly without any discernible pattern. The number of iterations necessary to reach convergence is termed the burn-in phase of the analysis, and estimates obtained during this period should not be used for inference making. The model appeared to converge at 10,000 iterations, but to be conservative, the first 20,000 iterations were discarded as the burn-in phase. Inferences were made on the basis of 200,000 iterations after burn-in. Median values and percentiles of the posterior distributions were used as point estimates and probability intervals, respectively.

The assumption of conditional independence among brucellosis serologic tests in cattle and water buffalo was evaluated as part of a sensitivity analysis. A model was created that included all possible 2-way sensitivity and specificity dependencies (6 each) among the 4 brucellosis tests. Models that incorporated pairwise dependencies thought most likely to be important as a result of the biologic nature of the tests under evaluation were also created. All models were implemented with the same informative prior probability distributions used in the primary 4-diagnostic-test conditional independence model. The 4 tests were also examined in 2-test

Table 1—Informative prior probability distribution for parameters in a Bayesian model used for estimating sensitivity and specificity for 4 brucellosis serologic tests in cattle and domestic water buffaloes (*Bubalus bubalis*) in which the true infection status was unknown

Parameters	Populations* and tests	Prior probability distribution (β)		95% Probability interval	
		Mode	Median		
Prevalence	Population 1	1.9, 18	0.05	0.08	0.011, 0.254
Prevalence	Population 2	1, 35	0.00 [†]	0.02	0.001, 0.100
Prevalence	Population 3	1.7, 3.8	0.20	0.28	0.037, 0.708
Prevalence	Population 4	1, 35	0.00 [†]	0.02	0.001, 0.100
Prevalence	Population 5	3.05, 5.25	0.33	0.36	0.098, 0.697
Prevalence	Population 6	1.7, 3.8	0.20	0.28	0.037, 0.708
Sensitivity	SPAT	13, 4.9	0.75	0.73	0.505, 0.899
Specificity	SPAT	50, 2.5	0.97	0.96	0.881, 0.992
Sensitivity	CT	18.4, 8.35	0.70	0.69	0.504, 0.845
Specificity	CT	60, 2.2	0.98	0.97	0.907, 0.995
Sensitivity	BPAT	17, 5	0.80	0.78	0.581, 0.918
Specificity	BPAT	35, 1.7	0.98	0.96	0.867, 0.996
Sensitivity	STAT	8.8, 4.5	0.69	0.67	0.399, 0.879
Specificity	STAT	27, 1.8	0.97	0.95	0.827, 0.993

Prior probabilities were the same for cattle and water buffalo.
^{*}Populations 1-3 are cattle and 4-6 are domestic water buffalo. [†]Mode is zero at 2 decimal places.
 SPAT = Standard plate agglutination test. CT = Card test. BPAT = Buffered plate agglutination test. STAT = Standard tube agglutination test.

only analyses incorporating conditional dependencies between the test pairs. All 6 possible 2-way analyses were performed by use of the same informative prior information as the primary model. Results of interest were point estimates and 95% probability intervals.

Sensitivity analysis was also performed to evaluate consistency of point estimates and intervals when changing prior information for the model. The effect of prior information should be evaluated by repeating all analyses with noninformative prior probabilities. A noninformative prior probability distribution corresponds to a situation where all values in the range of interest are equally likely. A β (1, 1) or uniform distribution has all values between 0 and 1 equally likely and is an appropriate noninformative prior for diagnostic test characteristics and infection prevalences, because their values must fall within this range. The sensitivity analysis was performed by use of β (1, 1) distributions for all sensitivity and specificity values in conjunction with the previously mentioned informative prior probabilities (Table 1) for the infection prevalences. Another analysis, which used noninformative prior probabilities for the prevalences but informative prior probabilities for the sensitivity and specificity of the tests, was also performed. Use of noninformative prior densities for all parameters in the model at the same time was not possible because of complexity of the problem and insufficient information in the data to estimate all parameters.⁴³ Models with high order terms, beyond pairwise dependencies, were also not considered for the same reason.

Results

Brucellosis testing—Seven hundred seventy-two animals were serologically evaluated for *B abortus*-specific antibodies by use of SPAT, CT, BPAT, and STAT. Serologic data were obtained for 391 cattle from 3 populations and 381 domestic water buffalo from 3 populations and were cross-classified by diagnostic test result pattern and animal population (Table 2).

Bayesian analysis—Posterior analysis yielded point

Table 2—Brucellosis serologic test result patterns for cattle and domestic water buffalo (*Bubalus bubalis*) populations in which the true infection status was unknown

Test result pattern	Diagnostic test result				No. of animals with test result pattern					
					Cattle populations			Water buffalo populations		
	SPAT	CT	BPAT	STAT	1	2	3	4	5	6
1	N	N	N	N	35	201	122	148	31	66
2	N	N	N	P	1	0	0	2	0	0
3	N	N	P	N	0	0	7	0	21	10
4	N	N	P	P	2	0	2	0	1	0
5	N	P	N	N	0	0	0	0	0	0
6	N	P	N	P	0	0	0	0	0	0
7	N	P	P	N	0	0	2	0	14	7
8	N	P	P	P	0	0	2	0	25	5
9	P	N	N	N	0	2	0	0	0	0
10	P	N	N	P	0	1	0	0	0	0
11	P	N	P	N	0	0	0	0	0	0
12	P	N	P	P	0	0	0	0	0	0
13	P	P	N	N	0	0	0	0	0	0
14	P	P	N	P	0	0	0	0	0	0
15	P	P	P	N	0	0	0	0	1	0
16	P	P	P	P	1	0	13	0	23	27
Total	NA	NA	NA	NA	39	204	148	150	116	115

SPAT, CT, BPAT, and STAT results are recorded as positive or negative in accordance with the Cooperative Brucellosis Eradication Program.³²
P = Positive. N = Negative. NA = Not applicable.
See Table 1 for remainder of key.

Table 3—Sensitivity and specificity estimates for 4 common brucellosis serologic tests by use of Bayesian methods for cattle and domestic water buffalo (*Bubalus bubalis*) in which the true infection status was unknown

Serologic test*	Species	Sensitivity-% (95% probability interval)	Specificity-% (95% probability interval)
SPAT	Cattle	66.7 [†] (50.9, 80.4)	98.9 (97.5, 99.6)
	Water buffalo	51.4 [†] (42.4, 60.3)	99.3 (98.1, 99.9)
CT	Cattle	72.7 [†] (58.7, 84.2)	99.6 (98.6, 99.9)
	Water buffalo	90.4 [†] (83.2, 95.3)	99.4 (98.1, 99.9)
BPAT	Cattle	88.1 [†] (76.2, 95.6)	98.1 [†] (96.4, 99.3)
	Water buffalo	96.3 [†] (92.1, 98.7)	90.7 [‡] (86.9, 94.1)
STAT	Cattle	80.2 (64.2, 91.5)	99.3 (98.0, 99.9)
	Water buffalo	75.0 (66.0, 82.7)	98.8 (97.2, 99.7)

*SPAT, CT, BPAT, and STAT estimates using recommended cutoff values in accordance with the Cooperative Brucellosis Eradication Program.³²
[†]Differences between cattle and water buffalo diagnostic test parameters are not zero with at least 90% probability. [‡]Differences between cattle and water buffalo diagnostic test parameters are not zero with at least 95% probability.
See Table 1 for remainder of key.

estimates and probability intervals for sensitivity and specificity of the 4 brucellosis serologic tests under evaluation (Table 3). For cattle, the BPAT had the highest sensitivity (88.1). Specificities of all tests were estimated to be good (> 98.0) for diagnosis of *Brucella* infection in cattle. Sensitivity of the BPAT (96.3) was the highest of the 4 tests for water buffalo, but the specificity of this test (90.7) was estimated to be the lowest of the 4. The specificities of all tests, except for the BPAT, were estimated to be good (> 98.0) in the water buffalo. Sensitivity of the CT and specificity of the BPAT were different between cattle and water buffalo, with at least 95% probability. Species differences in sensitivities of the SPAT and BPAT were found to be at least 90% likely.

Sensitivity of the BPAT was better than the SPAT in

Table 4—Prevalence estimates for *Brucella* infection by use of Bayesian methods in 6 populations of cattle and domestic water buffalo (*Bubalus bubalis*) in which the true infection prevalence was unknown

Population	Species	Prevalence (95% probability interval)
1	Cattle	0.080 (0.027, 0.168)
2	Cattle	0.005 (0.000, 0.021)
3	Cattle	0.140 (0.089, 0.203)
4	Water buffalo	0.004 (0.000, 0.020)
5	Water buffalo	0.564 (0.470, 0.659)
6	Water buffalo	0.342 (0.260, 0.431)

cattle, with at least 95% probability on the basis of the MCMC procedure. No other differences for sensitivity or specificity were likely at this probability level among cattle estimates. In domestic water buffalo, sensitivity of the SPAT was lower than all other tests. Sensitivity of the STAT was also different from all other serologic tests under evaluation. Specificity of the BPAT was also lower than all other tests in water buffalo, with at least 95% probability. All possible comparisons were evaluated with the model, and reported probability levels were based on each individual pairwise test.

Prevalence of *Brucella* infection was higher in domestic water buffalo populations, compared with cattle (Table 4). The highest prevalence in water buffalo was 56% in population 5, whereas the highest cattle prevalence was a relatively low 14% in population 3.

Two-test only (pair-wise) conditional dependence model analyses gave results consistent with the primary, 4-test conditional independence model (data not shown). Four-test conditional dependence models that were evaluated gave different results than the primary model, and estimates for some parameters were not biologically plausible (data not shown). These models yielded infection prevalence estimates of 1.0 for some populations, and sensitivity and specificity estimates of certain

tests were < 0.4 . Convergence was also a problem as some models were unable to converge even after 200,000 iterations, and bimodal distributions of covariance parameters were observed. Sensitivity analysis by use of non-informative priors for the serologic test parameters and infection prevalences yielded posterior distributions similar to results of the primary model (data not shown).

Discussion

A test-and-slaughter program had been instituted in Trinidad for the control of brucellosis in cattle and domestic water buffalo. Accurate screening tests are important for success of this control program, and the SPAT, CT, BPAT, and STAT were evaluated in our study for this reason. These tests were chosen because they are easily implemented without specialized equipment, and they have been used in other countries for the eradication of *Brucella* infection in livestock.

The BPAT was determined to be the most efficient test to screen for brucellosis in cattle and water buffalo of Trinidad, because it had the highest sensitivity of the 4 tests evaluated, and this is the most important criterion for eradication programs. Specificities of all 4 tests were estimated to be good in the tested cattle populations, and this has importance for the overall cost of eradication programs.

Based on our findings, the recommended brucellosis-screening test for cattle and water buffalo of Trinidad is the BPAT. However, this test does not function with the same degree of accuracy in these 2 species. Specificity of this test was better in cattle, compared with water buffalo, which will result in a higher percentage of false-positive test results for water buffalo. Use of a confirmatory test with high specificity for water buffalo is important to reduce the number of water buffalo improperly classified as infected. Sensitivity of the BPAT was calculated to be better in water buffalo, compared with cattle, but this difference was only 90% probable on the basis of the MCMC procedure. Screening tests with high sensitivity result in fewer false-negative classifications and can therefore be more efficient at finding infected animals for the test-and-slaughter program. Difference in diagnostic test parameters (sensitivity and specificity) across species is important to consider when developing disease control programs. Research performed in 1 species may not be directly applicable to another without verification of findings in target animal populations.

Selection bias may have been a problem in our study because only a fraction of the animals were tested from the larger herds. Population list frames were not available, and handling facilities were inadequate for the efficient inspection of all animals. This may have biased the estimates of population prevalence, because tested animals may not have been representative of the herd as a whole. However, *Brucella* prevalence estimates were higher in domestic water buffalo populations, compared with cattle. The cattle population with the highest prevalence had the same ownership as the farms of water buffalo with high *Brucella* prevalences. These 3 herds were managed similarly, and it is possible that the epidemiologic characteristics of brucellosis in domestic water buffalo may be different from what has been learned through extensive research with cattle. Prevalences were not compared, and

observed results may simply reflect variability in sample collection or the presence of bias in the estimates; nevertheless, they suggest a difference that should be further evaluated because the water buffalo is an important food producing animal in Trinidad and worldwide.

Bayesian iterative simulation techniques can be used to estimate sensitivity and specificity of diagnostic tests in the absence of a gold standard test. These methods are useful for situations when a gold standard test is too expensive or invasive to perform. Some diseases, such as brucellosis, rely on bacterial isolation to confirm infection, and this may not be effective at determining true status in all animals. Bacteriologic methods are seldom, if ever, 100% effective at recovering viable organisms from infected animals. Animals yielding isolates may be different in stage or degree of infection, compared with infected animals that have negative bacterial culture results. False-positive culture results are not likely to occur, but false-negative results may have a meaningful effect on sensitivity and specificity estimates.

Diagnostic tests that measure similar analytes, for example serum antibody concentrations, may not be conditionally independent, and dependence terms are easy to incorporate into a 2 diagnostic test model. Brucellosis agglutination tests are analytically sensitive to different antibody isotypes to varying degrees.^{44,45} The standard agglutination tests (STAT and SPAT) are more sensitive to IgM and IgG₂ antibody classes. The acid buffered agglutination tests (BPAT and CT) are more sensitive to the presence of IgG₁. Work in cattle has determined that IgG₁ is the antibody class associated with chronic brucellosis infections.⁴⁵ Immunoglobulin M is the predominant antibody of the primary immune response and is efficient at agglutination as a result of its pentamer structure. This binding and agglutinating efficiency can lead to false-positive reactions. The acidified buffered antigen tests were developed to minimize false-positive reactions caused by this antibody class.^{30,44,45} Similar work has not been performed in water buffalo to determine whether acid buffering of these tests prevents IgM binding and agglutination in this species.

Models that incorporated conditional dependence among the tests would be expected to give better results than conditional independence models, because all tests measure agglutinating antibody concentrations and are therefore expected to be conditionally dependent. However, complex problems involving results from more than 2 tests become increasingly difficult to incorporate appropriate dependence terms. Four-diagnostic-test models must allow for the possibility of 3- and 4-way dependency terms among diagnostic tests. The saturated model (including all possible 2-, 3-, and 4-way dependency terms) can have too many unknown parameters to estimate in relationship to the amount of data (degrees of freedom). Evaluating all possible models that incorporate these sensitivity and specificity dependency terms is difficult, if not impossible, as more diagnostic test results are added. It is unclear why the 4-test conditional independence model performed better than the evaluated dependence models. Multiple zero counts for certain test patterns in conjunction with possible complex dependencies among these 4 diagnostic tests may have prevented the determination of a better model. Further investigations

are necessary to provide insight into the incorporation of conditional dependence terms in such models when results from more than 2 diagnostic tests are used.

Bayesian techniques were chosen for use in the diagnostic test model over maximum likelihood methods that have become increasingly popular. Data were provided (Table 2) so that the interested reader could evaluate other methodologic approaches to solving this problem. The authors first performed a likelihood-based analysis, but estimates of sensitivity and specificity appeared to be highly dependent on the choice of best guesses for the starting points of the iterative algorithms. The complex nature of the likelihood function resulting from the large number of unknown parameters for estimation resulted in the possibility that only local maxima were being found, rather than the global maximum. This implies that standard large sample theory inferences based on the maximum-likelihood results would not apply. Conditional independence models solved by use of the likelihood-based approach did not adequately predict the data on the basis of rejection of the likelihood ratio test for model fit. However, more than half (35/48 for cattle; 34/48 for water buffalo) of the expected cell counts were < 1.0, invalidating the usual theory used for interpretation of the likelihood ratio statistic. Use of prior information in a Bayesian approach helps to guide the analysis in the direction of biologically plausible estimates and results in a valid statistical inference regardless of the sample size or existence of a multimodal likelihood.

The evaluated models assumed equal diagnostic potential (sensitivity and specificity) of the serologic tests in all 3 populations within each species. Potential violation of this assumption should always be evaluated when performing such analyses. Sensitivity of a test may be associated with certain animal-level factors, such as concurrent health problems, pregnancy, nutritional status, and other situations that may result in immunosuppression and poor humoral antibody response secondary to *Brucella* infection. The first author was present at all sample collections, and this was not thought to be a source of bias in the present investigations. Another major consideration is the stage of infection (acute vs chronic) of the animals in the tested herds. Sensitivities of these screening tests are expected to be higher in acutely infected versus chronically infected animals where circulating antibodies may have declined below the analytic thresholds. Therefore, if the tested herds had varying proportions of animals in the various stages of infection, then this assumption of equal diagnostic potential across populations may have been violated. *Brucella abortus* infection was first recognized in Trinidad during 1998, and it is unknown whether time from first exposure until sample collection had been different for the tested farms. The lack of history of abortion storms on the tested farms suggested that *B abortus* might have already been endemic on the farms at the time of first recognition in 1998. On the basis of these observations, there was no evidence of violation of the assumption in the model of equal diagnostic accuracy across tested populations.

Sensitivity and specificity are important aspects of diagnostic tests used for screening and controlling disease in populations. Knowledge of test accuracy is

imperative when designing eradication programs. Estimates of sensitivity and specificity can be obtained through use of various study designs and analytic methods. Traditional approaches use a gold standard for determination of diagnostic test characteristics. However, a gold standard test is often not available as a result of study limitations or inability to determine true infection status in an economically and ethically feasible manner. The MCMC iterative techniques can be used to estimate accuracy of diagnostic tests when a gold standard is not available by use of results from more than 2 serologic tests. Such procedures can also be used to statistically compare estimates of sensitivity and specificity for different tests and different animal species. These methods may be most appropriate for evaluation of tests in wildlife and other species of animals where traditional diagnostic test research has been limited as a result of economical, ethical, and feasibility reasons.

^aBrewer Diagnostic Kits, Becton Dickinson Microbiology Systems, Cockesville, Md.

^bProtocols provided by Drs. J. Payeur and B. Martin USDA/NVSL, Ames, Iowa.

^cRisk Version 4.0, Palisade Corporation, Newfield, NY.

^dWinBUGS Version 1.3, MRC Biostatistics Unit, Cambridge, UK.

^eWinBUGS program code available upon request to Dr. Fosgate, Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, Calif.

References

1. Blajan L, Melendez LV. Contribution of the OIE to controlling animal brucellosis on a world-wide scale. *Rev Sci Tech* 1984;3:167-189.
2. Fosgate GT, Adesiyun AA, Hird DW, et al. Isolation of *Brucella abortus* biovar 1 from cattle and water buffalo of Trinidad. *Vet Rec* 2002;151:272-273.
3. Pietz DE, Cowart WO. Use of epidemiologic data and serologic tests in bovine brucellosis. *J Am Vet Med Assoc* 1980;177:1221-1226.
4. Stemshorn BW, Forbes LB, Eaglesome MD, et al. A comparison of standard serological tests for the diagnosis of bovine brucellosis in Canada. *Can J Comp Med* 1985;49:391-394.
5. Dohoo IR, Wright PF, Ruckerbauer GM, et al. A comparison of five serological tests for bovine brucellosis. *Can J Vet Res* 1986;50:485-493.
6. Nielsen KH, Kelly L, Gall D, et al. Comparison of enzyme immunoassays for the diagnosis of bovine brucellosis. *Prev Vet Med* 1996;26:17-32.
7. Molnár E, Molnár L, Vale WG. Value of different serological tests in the diagnosis of bovine brucellosis in the Amazonian region. *Acta Vet Hung* 1998;46:199-210.
8. Nielsen K, Gall D, Lin M, et al. Diagnosis of bovine brucellosis using a homogeneous fluorescence polarization assay. *Vet Immunol Immunopathol* 1998;66:321-329.
9. Nielsen K, Gall D, Smith P, et al. Fluorescence polarization assay for the diagnosis of bovine brucellosis: adaptation to field use. *Vet Microbiol* 2001;80:163-170.
10. Stemshorn BW. Recent progress in the diagnosis of brucellosis. *Dev Biol Stand* 1984;56:325-340.
11. Standards Commission, International Office of Epizootics. *Manual of standards for diagnostic tests and vaccines: lists A and B diseases of mammals, birds and bees*. 3rd ed. Paris: Office international des épizooties, 1996;242-255.
12. Zaki R. *Brucella abortus* infection among buffaloes in Egypt. *J Comp Pathol* 1948;58:73-79.
13. Mathur TN. A study of human brucellosis based on cultures isolated from man and animals. *Indian J Med Res* 1968;56:250-258.
14. Mohan RN. Diseases and parasites of buffaloes. *Vet Bull* 1968;38:647-659.
15. Panjarathinam R, Jhala CI. Brucellosis in Gujarat State. *Indian J Pathol Microbiol* 1986;29:53-60.
16. Pal M, Jain HS. Anthroponozoonotic role of *Brucella abortus*. *Int J Zoonoses* 1986;13:246-248.

17. Nicoletti P. An evaluation of serologic tests used to diagnose brucellosis in buffaloes (*Bubalus bubalis*). *Trop Anim Health Prod* 1992;24:40–44.

18. Sandhu KS, Joshi DV. Comparative study in cattle and buffaloes for evaluation of various diagnostic tests for brucellosis. *Indian J Pathol Microbiol* 1993;36:458–465.

19. Rao CK. Epizootiology, diagnosis and control of brucellosis in India. *Bull Off Int Epizoot* 1970;73:3–7.

20. Subhargkassen S. Brucellosis in Thailand. *Bull Off Int Epizoot* 1970;73:9–15.

21. Joshi DV, Prakash O. Prevalence of brucellosis in man and animals. *Indian J Pathol Bacteriol* 1971;14:96–101.

22. Nandgoankar D, Narayana Rao PL. A survey of the incidence of bovine brucellosis in the integrated milk project area and in some of the dairy farms of Andhra Pradesh. *Indian Vet J* 1971;48:12–18.

23. Randhawa AS, Kalra DS, Kapur MP. Some sero-epidemiologic observations on brucellosis in humans and animals. *Indian J Med Sci* 1974;28:133–138.

24. Sharma VD, Sethi MS, Yadav MP, et al. Sero-epidemiologic investigations on brucellosis in the states of Uttar Pradesh (U.P.) and Delhi (India). *Int J Zoonoses* 1979;6:75–81.

25. Mathias LA, Pinto AA. Serological diagnosis of brucellosis in water buffaloes (*Bubalus bubalis*): comparison among complement fixation, serum agglutination and rose bengal plate test. *Int J Zoonoses* 1983;10:122–126.

26. Akhtar S, Afzal M, Ali S, et al. Effect of reactor retention on the spread of brucellosis in Jersey cattle and buffalo herds. *Rev Sci Tech* 1990;9:1179–1185.

27. Akhtar S, Mirza MA. Rates of seroconversion in the progeny of *Brucella abortus* seropositive and seronegative cattle and buffalo. *Rev Sci Tech* 1995;14:711–718.

28. Sharma JK, Saini SS. Sero-prevalence of brucellosis among farm animals of Punjab. *Indian Vet J* 1995;72:881–882.

29. Isloor S, Renukaradhya GJ, Rajasekhar M. A serological survey of bovine brucellosis in India. *Rev Sci Tech* 1998;17:781–785.

30. Nielsen K. A brief review of diagnosis of bovine brucellosis by detection of antibody. *Archivos De Medicina Veterinaria* 1995;27:9–17.

31. Alton GG. Institut national de la recherche agronomique (France). *Techniques for the brucellosis laboratory*. Paris: Institut national de la recherche agronomique, 1988;123–136.

32. United States Animal and Plant Health Inspection Service. Brucellosis eradication uniform methods and rules: effective May 6, 1992. Raleigh, NC: US Department of Agriculture Animal and Plant Health Inspection Service, 1992;21–49.

33. Cheville NF, McCullough DR, Paulson LR, et al. *Brucellosis in the greater Yellowstone area*. Washington, DC: National Academy Press, 1998;1–41.

34. Hui SL, Walter SD. Estimating the error rates of diagnostic tests. *Biometrics* 1980;36:167–171.

35. Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics* 1985;41:959–968.

36. Dendukuri N, Joseph L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics* 2001;57:158–167.

37. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol* 1995;141:263–272.

38. Enøe C, Georgiadis MP, Johnson WO. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med* 2000;45:61–81.

39. Gardner IA, Stryhn H, Lind P, et al. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. *Prev Vet Med* 2000;45:107–122.

40. Gelman A, Carlin JB, Stern HS, et al. *Bayesian data analysis*. 2nd ed. London: Chapman & Hall, 1997;3–27.

41. Carlin BP, Louis TA. *Bayes and Empirical Bayes methods for data analysis*. 2nd ed. Boca Raton, Fla: Chapman & Hall/CRC, 2000;121–191.

42. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov chain Monte Carlo in practice*. Gilks WR, Richardson S, Spiegelhalter DJ, eds. Boca Raton, Fla: Chapman & Hall/CRC, 1998.

43. Johnson WO, Gastwirth JL, Pearson LM. Screening without a “gold standard”: the Hui-Walter paradigm revisited. *Am J Epidemiol* 2001;153:921–924.

44. Allan GS, Chappel RJ, Williamson P, et al. A quantitative comparison of the sensitivity of serologic tests for bovine brucellosis to different antibody classes. *J Hyg (Lond)* 1976;76:287–298.

45. Mittal KR, Tizard IR. Agglutination tests and their modifications in the diagnosis of bovine brucellosis. *Comp Immun Microbiol Infect Dis* 1983;6:1–8.

Appendix

Steps in the analysis of test results for infected and noninfected animals

Step 1—Observed numbers of animals cross-classified by diagnostic test results (positive [+]) or negative [–] are the sum of the unobserved (latent) numbers of infected and noninfected animals

	STAT		
	+	–	
BPAT +	w	x	w + x
BPAT –	y	z	y + z
	w + y	x + z	n

BPAT = Buffered plate agglutination test

STAT = Standard tube agglutination test

w = Number infected (w') plus noninfected (w'')

x = Number infected (x') plus noninfected (x'')

y = Number infected (y') plus noninfected (y'')

z = Number infected (z') plus noninfected (z'')

n = Total number of infected (n') plus noninfected (n'')

Step 2—Probability statements for *Brucella*-infected animals

	STAT		
	+	–	
BPAT +	w'	x'	w' + x'
BPAT –	y'	z'	y' + z'
	w' + y'	x' + z'	n'

$P(w') = (Se_{STAT}) * (Se_{BPAT}) + \gamma_{Se}$

$P(x') = (1 - Se_{STAT}) * (Se_{BPAT}) - \gamma_{Se}$

$P(y') = (Se_{STAT}) * (1 - Se_{BPAT}) - \gamma_{Se}$

$P(z') = (1 - Se_{STAT}) * (1 - Se_{BPAT}) + \gamma_{Se}$

Se_{STAT} = Sensitivity of STAT

Se_{BPAT} = Sensitivity of BPAT

γ_{Se} = Sensitivity covariance between STAT and BPAT

Step 3—Probability statements for *Brucella*-noninfected animals

	STAT		
	+	–	
BPAT +	w''	x''	w'' + x''
BPAT –	y''	z''	y'' + z''
	w'' + y''	x'' + z''	n''

$P(w'') = (1 - Sp_{STAT}) * (1 - Sp_{BPAT}) + \gamma_{Sp}$

$P(x'') = (Sp_{STAT}) * (1 - Sp_{BPAT}) - \gamma_{Sp}$

$P(y'') = (1 - Sp_{STAT}) * (Sp_{BPAT}) - \gamma_{Sp}$

$P(z'') = (Sp_{STAT}) * (Sp_{BPAT}) + \gamma_{Sp}$

Sp_{STAT} = Specificity of STAT

Sp_{BPAT} = Specificity of BPAT

γ_{Sp} = Specificity covariance between STAT and BPAT

Step 4—Probability statements for observed data are the weighted sum of probabilities for infected and noninfected animals

	STAT		
	+	–	
BPAT +	w	x	w + x
BPAT –	y	z	y + z
	w + y	w + z	n

$P(w) = [P(w') * \text{prevalence}] + [P(w'') * (1 - \text{prevalence})]$

$P(x) = [P(x') * \text{prevalence}] + [P(x'') * (1 - \text{prevalence})]$

$P(y) = [P(y') * \text{prevalence}] + [P(y'') * (1 - \text{prevalence})]$

$P(z) = [P(z') * \text{prevalence}] + [P(z'') * (1 - \text{prevalence})]$