A Comparison of Five Serological Tests for Bovine Brucellosis

I.R. Dohoo, P.F. Wright, G.M. Ruckerbauer, B.S. Samagh, F.J. Robertson and L.B. Forbes*

ABSTRACT

Five serological assays: the buffered plate antigen test, the standard tube agglutination test, the complement fixation test, the hemolysis-in-gel test and the indirect enzyme immunoassay were diagnostically evaluated. Test data consisted of results from 1208 cattle in brucellosis-free herds, 1578 cattle in reactor herds of unknown infection status and 174 cattle from which *Brucella abortus* had been cultured.

The complement fixation test had the highest specificity in both nonvaccinated and vaccinated cattle. The indirect enzyme immunoassay, if interpreted at a high threshold, also exhibited a high specificity in both groups of cattle. The hemolysis-in-gel test had a very high specificity when used in nonvaccinated cattle but quite a low specificity among vaccinates.

With the exception of the complement fixation test, all tests had high sensitivities if interpreted at the minimum threshold. However, the sensitivities of the standard tube agglutination test and indirect enzyme immunoassay, when interpreted at high thresholds were comparable to that of the complement fixation test.

A kappa statistic was used to measure the agreement between the various tests. In general the kappa statistics were quite low, suggesting that the various tests may detect different antibody isotypes. There was however, good agreement between the buffered plate antigen test and standard tube agglutination test (the two agglutination tests evaluated) and between the complement fixation test and the indirect enzyme immunoassay when interpreted at a high threshold.

With the exception of the buffered plate antigen test, all tests were evaluated as confirmatory tests by estimating their specificity and sensitivity on screening-test positive samples. Tests were evaluated alone or in conjunction with one other test with the results interpreted in series or parallel. The complement fixation test and the indirect enzyme immunoassay (interpreted at a high threshold) had the highest specificities of individual tests but most combinations of tests resulted in a high specificity provided that the results were interpreted in series. If two tests were interpreted in parallel, only the combinations of complement fixation test and indirect enzyme immunoassay (at a high threshold) following indirect enzyme immunoassay screening resulted in an acceptable level of diagnostic specificity.

The sensitivities of the confirmatory tests in screening-test positive samples were similar to the sensitivities estimated from all the culture positive sera. The method of interpretation of pairs of results did not have as large an impact on sensitivity as it did on specificity.

It is recommended that either the buffered plate antigen test or indirect enzyme immunoassay be used as a screening test. Either the complement fixation test or the indirect enzyme immunoassay is appropriate for use as a confirmatory test in situations requiring a high specificity. The latter test has some technical advantages.

Key words: Bovine, brucellosis, serology, sensitivity, specificity, tests, screening.

RÉSUMÉ

Cette expérience portait sur les cinq épreuves sérologiques suivantes : l'épreuve de l'agglutination rapide avec un antigène-tampon, l'épreuve de l'agglutination lente standard, l'épreuve de la déviation du complément, l'épreuve de l'hémolyse sur gel et l'épreuve immuno-enzymatique indirecte. Elle visait à évaluer leur efficacité comme méthodes de diagnostic de la brucellose. Les auteurs les utilisèrent à cette fin sur 1208 bovins appartenant à des troupeaux exempts de brucellose, sur 1578 issus de troupeaux qui comptaient des réacteurs, mais dont ils ignoraient le statut d'infection, et sur 174 desquels ils avaient isolé *Brucella abortus*.

L'épreuve de la déviation du complément se révéla la plus spécifique, tant chez les sujets vaccinés que chez les autres. L'épreuve immuno-enzymatique indirecte s'avéra aussi très spécifique, chez les deux groupes de sujets précités, lorsqu'on l'interprétait à un seuil élevé. L'épreuve de l'hémolyse sur gel afficha une grande spécificité, chez les sujets non vaccinés, contrairement à ce qui se produisit chez les vaccinés.

À l'exception de l'épreuve de la déviation du complément, toutes les épreuves précitées manifestèrent une grande sensibilité, lorsqu'on les interprétait au seuil minimal. La sensibilité de l'épreuve de l'agglutination lente standard et celle de l'épreuve immuno-enzymatique indirecte s'avérèrent toutefois comparables à celle de l'épreuve de la déviation du complément, lorsqu'on les interprétait à un seuil élevé.

Les auteurs utilisèrent une méthode statistique kappa, pour mesurer la

*Animal Pathology Division, Agriculture Canada, 2255 Carling Avenue, Ottawa, Ontario K1A 0Y9 (Dohoo), Animal Diseases Research Institute, Agriculture Canada, 801 Fallowfield Road, Nepean, Ontario K2H 8P9 (Wright, Ruckerbauer, Samagh, Robertson) and Animal Pathology Laboratory, Agriculture Canada, 116 Veterinary Road, Saskatoon, Saskatchewan S7N 2R3 (Forbes). Present address of Dr. Dohoo: Department of Health Management, Atlantic Veterinary College, University of P.E.I., Charlottetown, Prince Edward Island C1A 4P3.

Submitted July 18, 1985.

concordance entre les diverses épreuves précitées. Les résultats ainsi obtenus se révélèrent plutôt bas, indice qu'elles détecteraient différents isotypes d'anticorps. Ils constatèrent toutefois une bonne concordance entre les deux épreuves d'agglutination, à savoir: celle de l'agglutination rapide avec un antigène-tampon et celle de l'agglutination lente standard, de même qu'entre l'épreuve de la déviation du complément et l'épreuve immuno-enzymatique indirecte, lorsqu'ils les interprétaient à un seuil élevé.

À l'exception de l'épreuve de l'agglutination rapide avec un antigène-tampon, les auteurs utilisèrent toutes les épreuves précitées, à titre d'épreuves confirmatives, en évaluant leur spécificité et leur sensibilité sur des échantillons positifs de tamisage. Cette évaluation se fit sur une base individuelle et comparative, et l'interprétation des résultats, en série ou en parallèle. L'interprétation à un seuil élevé de l'épreuve de la déviation du complément et de l'épreuve immuno-enzymatique indirecte donna la spécificité la plus élevée, à titre individuel; la plupart des combinaisons d'épreuves manifestèrent aussi une spécificité élevée, en autant que l'interprétation des résultats se faisait en série. L'interprétation en parallèle des résultats de deux épreuves à la fois, révéla que seulement la combinaison de l'épreuve de la déviation du complément et de l'épreuve immuno-enzymatique indirecte, à un seuil élevé, après un tamisage par l'épreuve de l'agglutination rapide avec un antigène-tampon, ainsi que celle de l'épreuve de la déviation du complément et de l'épreuve de l'agglutination lente standard, à un seuil élevé, après un tamisage par l'épreuve immuno-enzymatique indirecte, donnaient un niveau acceptable de spécificité diagnostique.

La sensibilité des épreuves confirmatives, avec les échantillons positifs de tamisage, se révéla comparable à celle qu'obtinrent les auteurs, à partir du sérum de tous les sujets desquels ils avaient isolé *B. abortus*. La méthode d'interprétation des résultats de paires d'épreuves n'afficha pas un impact aussi important sur la sensibilité que sur la spécificité.

Les auteurs recommandent d'utiliser, pour le tamisage, l'épreuve de l'agglutination rapide avec un antigène-tampon ou l'épreuve immuno-enzimatique indirecte. L'utilisation de cette dernière ou de l'épreuve de la déviation du complément, à titre d'épreuves confirmatives, donne de bons résultats, dans les cas qui requièrent une spécificité élevée. L'épreuve immuno-enzymatique indirecte offre aussi certains avantages techniques.

Mots clés: brucellose bovine, sérologie, épreuves de sensibilité et de spécificité, tamisage.

INTRODUCTION

Prior to the utilization of a new test in a diagnostic program, it is desirable that the specificity and sensitivity of the test be determined and compared to standard tests. The specificity of a test is the probability of the test returning a negative result when applied to a nonaffected animal (1). A high level of specificity ensures that a test does not incorrectly classify too many nonaffected animals as positive. The sensitivity of a test is the probability of the test returning a positive result when applied to an affected animal (1). A test with high sensitivity will prevent an excessive number of affected animals escaping detection.

The sensitivity and specificity of a test may vary among populations of animals as the characteristics of the population change. One factor which is known to influence the specificity of serological tests for *Brucella abortus* is the prevalence of vaccination in the tested population (2). Consequently it is important to know the specificity of the tests in both vaccinated and nonvaccinated populations.

It is common practice to use a rapid, economical and highly sensitive test as a screening test and then apply one or more additional confirmatory tests to those samples with a positive result. In this case it is important to know the sensitivity and specificity of the confirmatory tests when applied only to screening test positive samples. If two confirmatory tests are used the results can be interpreted in series or in parallel and the method of interpretation affects the sensitivity and specificity of the testing program. Series interpretation means that a sample is given a positive test result only if it is positive on both confirmatory tests. This results in the testing program having a high level of specificity. Parallel interpretation results in a sample being declared positive if it has a positive result on either of the two confirmatory tests. This diminishes the specificity of the testing programs but increases the sensitivity.

This paper describes the evaluation of five tests for the serological diagnosis of *B. abortus*. For two of the tests, multiple positive thresholds were evaluated.

MATERIALS AND METHODS

SEROLOGICAL ASSAYS

All serological tests were performed at the Animal Diseases Research Institute, Nepean, Ontario by personnel carrying out routine diagnostic testing. The origin of the samples was unknown to those carrying out the tests.

Buffered Plate Antigen Test (BPAT) — This test was performed as previously described by Angus and Barton (3). The BPAT antigen (Agriculture Canada, Animal Diseases Research Institute, Microbiology Service Section, P.O. Box 11300, Station H, Nepean, Ontario K2H 8P9) consisted of an 11% suspension of *B. abortus* strain 1119-3 stained with crystal violet and brilliant green and buffered to pH 3.63. Thirty μL volumes of antigen were mixed by hand with 80 μL volumes of test sera on a glass plate. After it was rocked with a circular motion, the plate was incubated in a humid chamber at 37°C for eight minutes. Immediately prior to reading, the plates were again rocked with a circular motion to ensure a uniform suspension of reagents. Reactions were visually scored as either negative (–) in the absence of agglutination or positive (+) in the presence of partial or complete agglutination.

Standard Tube Agglutination Test (STAT) — This test was performed

essentially as described by Malkin (4). A stock suspension of 8% *B. abortus* strain 413 (Agriculture Canada, Animal Diseases Research Institute, Microbiology Service Section, P.O. Box 11300, Station H, Nepean, Ontario K2H 8P9) was diluted with 0.5% phenolyzed physiological saline to a working concentration of 1/200 and dispensed in 2.0 mL volumes into tubes. Test serum dilutions of 1/25, 1/50, 1/100 and 1/200 were prepared by the addition of an appropriate volume of neat serum into each tube. Primary incubation at 37°C for 24 hours was followed by secondary incubation at room temperature for 18 hours. Following incubation, each tube in the series was visually scored (1+ to 4+) on the basis of clarity of the supernatant in the serum-antigen mixture. Agglutinated complexes were resuspended from the bottom of the tube by gentle tapping with the finger and depending on their fragility, the test was finally scored (e.g. water clear supernatant with large intact complexes was scored as 4+). Agglutination reactions of 3+ or greater at test dilutions of 1/25 (STAT 1) and 1/50 (STAT 2) were independently considered as minimum seropositive thresholds. These reactions represent agglutinating activities of approximately 30 and 60 IU/mL, resprectively (5).

Complement Fixation Test (CFT) — This test was performed essentially as described by Samagh and Boulanger (6); however, rather than using tubes, the assay was performed in microtiter plates (U-bottom, #001-010-2201, Dynatech Laboratories Inc., 900 Slaters Lane, Alexandria, Virginia 22314) with one-quarter the volume of all reagents. Test sera, heat inactivated at a dilution of 1/5, were serially diluted (twofold) in veronal buffer, pH 7.3, within the wells of the microtiter plate such that the final volume of each dilution was 25 μL. Equivalent volumes of antigen (1/200 of STAT stock) and guinea pig complement (three corrected hemolytic units), each in veronal buffer, were mixed with the serum dilutions and the plates were incubated at 4°C for 18 hours. After addition of a 50 μL volume of a 1.25% suspension of sensitized sheep erythrocytes in veronal buffer (prepared initially as a 2.5% suspension prior to sensitization), the plates were twice incubated at 37.5°C for 20 minutes with a five minute period of plate shaking (Plate shaker, Titertek, Flow Laboratories Inc., 7655 Old Springhouse Road, McLean, Virginia 22102) at room temperature between incubations. After centrifugation (Centrifuge, IEC Model UV, International Equipment Company, 300 Second Avenue, Needham Heights, Massachusetts 02194) of the plates, each well in the dilution series was visually scored on the basis of hemolysis and cell-button formation. A hemolytic reaction of 50% or less at a dilution of 1/5 was considered as the minimum seropositive threshold. Appropriate controls were included for each serum tested, as well as, other reagents.

Enzyme Immunoassay (EIA) — This assay was performed as described by Nielsen and Wright (2). Smooth lipopolysaccharide antigen from *B. abortus* strain 413, 1 μg/mL in 0.06 M carbonate buffer, pH 9.6, was coated onto microtiter plate wells (Microplates (flat bottom), Linbro/Titertek #76-331-05, Flow Laboratories Inc., 7655 Old Springhouse Road, McLean Virginia 22102) at room temperature for 18 hours. Test sera (1/100) and peroxidase-conjugated rabbit anti-bovine IgG (heavy and light chain (1/2000) (#3202-0082, Coopermedical Inc., One Technology Court, Malvern, Pennsylvania 19355) in 0.01 M phosphate-buffered saline, pH 7.2, plus 0.05% Tween 20 were separately incubated in the wells at room temperature for three hours and one hour respectively. The substrate/chromogen consisted of 4 mM H_2O_2 plus 1mM ABTS (Chromogen 2,2'-Azinobis (3-ethylbenzthiazolinesulfonic acid), #A1888, Sigma Chemical Company, P.O. Box 14508, St. Louis, Missouri 63178) in 0.05 M citrate-NaOH buffer, pH 5.0. After a calculated reaction interval (10.4 minutes average), chromogen optical density (OD) was photometrically (Microplate Reader, Multiskan, Flow Laboratories Inc., 7655 Old Springhouse Road, McLean, Virginia 22102) measured at 414 nm. Four OD values, 0.220 (EIA 1), 0.260 (EIA 2), 0.300 (EIA 3) and 0.340 (EIA 4) were independently considered as minimum seropositive thresholds.

Hemolysis-in-Gel Test (HIGT) — This was performed as described by Ruckerbauer *et al* (7). A 5% suspension of bovine erythrocytes (J-antigen negative) in phosphate buffered saline, pH 7.2, was sensitized with alkali-treated smooth lipopolysaccharide from *B. abortus* strain 413 (8,9) at a concentration of 250 μg/mL of suspension. Sensitized erythrocytes (0.75%) and guinea pig complement (10% v/v) were incorporated into a low-temperature gelling, 1% agarose gel (Type VII, #A4018, Sigma Chemical Company, P.O. Box 14508, St. Louis, Missouri 63178) in veronal buffer, pH 7.3. Test sera were pipetted in 20 μL volumes into wells cut into the gel (Immunodiffusion plates, 16-well template and well-cutter, Miles Laboratories Inc., P.O. Box 2000, Elkhardt, Indiana 46514). After primary incubation at 4°C for 18 hours and a secondary incubation at 37°C for two hours, zones of hemolysis were measured and expressed as millimeters diameter. A zone of hemolysis of 6 mm was considered to be the minimum seropositive threshold. Appropriate controls were included for every serum tested as well as other reagents.

A summary of criteria for seropositive designation for the five assays described is presented in Table I.

SAMPLE SELECTION

Three groups of samples, referred to as negative herd samples, culture positive samples and reactor herd samples, were used in this study. The negative herd samples were obtained from 22 Ontario dairy herds, free-listed for brucellosis on the basis of annual serological certification in both 1982 and 1983. Herds become free-listed if there are no BPAT reactors in the herd or if any BPAT reactors are subsequently declared negative by the use of STAT and CFT confirmatory tests. A total of 1209 sera were collected from these herds and it was assumed that all of these samples were from cattle free of infection with *B. abortus*. Of these cattle, 80 (6.6%) had a history of calfhood vaccination and were between the ages of 18 months and five years. All cattle over five years of age would have been recorded as nonvaccinates.

Culture positive samples were sera obtained from cattle from whose tissues or fluids an isolation of *B. abortus* had been made. A total of 174 sera were available with 159 being from western Canadian beef cattle. These animals were selected for culturing on the basis of a confirmed isolation from a reactor in their herd of origin and were not necessarily serologically positive themselves. The other 15 samples were from eastern Canadian dairy cattle which were selected for culture on the basis of a serological reaction. All sera were obtained at or near the time of slaughter of the cattle.

Reactor herds were defined as any herd having, on initial bleeding, at least two BPAT positive reactors which were also seropositive on either the CFT or the STAT 2 or demonstrated low grade reactions (i.e. the minimum threshold criteria shown in Table I) on both of these assays. Under routine diagnostic conditions selective repeat bleedings of STAT and/or CFT reactors and their herdmates are normally requested. When available, sera from all repeat bleedings (whole herd or reactors only), were included in the reactor herd group. A total of 22 herds were classified as reactor herds and isolation of *B. abortus* was attempted in eight of these herds either prior to or during the study. Positive cultures were obtained from 19 cows in three herds. A total of 3440 sera were obtained from 1578 cows with 1004 (29.2%) sera being from cattle with a history of calfhood vaccination against brucellosis.

TABLE I. Summary of Criteria for Seropositive Designation of Several Diagnostic Tests for Bovine Brucellosis

Assay	Positive Threshold	Comment
BPAT agglutination	+	Partial or complete agglutination
STAT 1	≥ 1/25	Minimum 3+ agglutination
STAT 2	≥ 1/50	Minimum 3+ agglutination
CFT point	≥ 1/5	50% or less hemolytic end point
EIA 1	≥ 0.220	Optical density at 414 nm
EIA 2	≥ 0.260	Optical density at 414 nm
EIA 3	≥ 0.300	Optical density at 414 nm
EIA 4	≥ 0.340	Optical density at 414 nm
HIGT	≥ 6	Ring diameter in mm

DATA MANAGEMENT AND ANALYSIS

The data were stored, edited and manipulated using a computer based database management package (Datatrieve, Version 2, Digital Equipment Corp., Maynard, Massachusetts).

Estimates of each tests specificity in vaccinated cattle, nonvaccinated cattle and all animals combined, were determined using the data from the negative herd samples. Estimates of the minimum possible specificity of the tests when used in vaccinated cattle were also obtained by assuming that all positive reactions among vaccinates in the reactor herds were false positives. If any of these positive reactions occurred in sera from infected cows (true positives) then the estimate of specificity would be too low. Confidence intervals were calculated for all point estimates using an approximate method based on the distribution of the F-statistic (10).

Estimates of the sensitivity of each test were determined from the results obtained from the sera of culture positive cattle. Relative sensitivities for all tests, except the BPAT, were obtained from reactor herd data. All samples from these herds which had a positive BPAT were identified and the proportion of these samples which were positive for each of the other tests was determined. In order to avoid biasing the estimates of relative sensitivity in favour of tests with low specificities the number of false positives expected to be found (based on the estimates of specificity determined above) was subtracted from the total number of positives observed. The method of calculation was as follows:

The level of agreement among all pairs of tests was determined in two ways. Firstly the percentage of samples which yielded the same result on each test was determined and secondly agreement was determined using the kappa statistics (11). Kappa is a measure of agreement that is adjusted for agreement due to chance. It ranges from $-\infty$ to +1. A value of $-\infty$ represents perfect disagreement, 0 represents no agreement above what would be expected due to chance alone and +1 represents perfect agreement.

In order to further evaluate the various tests as confirmatory tests, their sensitivity and specificity were determined using samples that had previously tested positive on a screening test. Tests were evaluated by themselves or in conjunction with one other test applied in series or in parallel. For the evaluation of specificity, all samples from the negative herds that were BPAT or EIA 1 positive were identified. These two groups of

# of nonvaccinates tested x (1 - specificity (in nonvaccinates))
= expected # false positives amongst nonvaccinates
= EFPnv

# of vaccinates tested x (1 - specificity (in vaccinates))
= expected # false positives amongst vaccinates
= EFPv

Total # of false positives expected = EFPnv + EFPv = EFP

Relative Sensitivity = (observed # positive — EFP)/Total # tested

TABLE II. Estimates of the Specificity of Nine Serological Tests for Bovine Brucellosis Based on Sera from Several Sources[a]

Test	Negative Herds All Cows (%) N[b] = 1208	Negative Herds Nonvaccinates (%) N = 1128	Negative Herds Vaccinates (%) N = 80	Reactor Herds Vaccinates (%) N = 1004
BPAT	98.7[c] (97.9, 99.2)[d]	98.8 (98.0, 99.4)	92.5 (84.2, 97.2)	93.4 (91.7,94.9)
STAT 1	93.3 (91.7, 94.6)	93.9 (92.3, 95.2)	85.0 (75.1, 92.0)	69.0 (65.3,72.8)
STAT 2	99.3 (98.7, 99.7)	99.5 (98.9, 99.8)	97.5 (90.0, 99.7)	94.9 (93.3,96.2)
CFT	100.0 (99.7,100.0)	100.0 (99.7,100.0)	100.0 (95.4,100.0)	99.3 (98.6,99.7)
HIGT	98.3 (97.4, 98.9)	99.8 (99.4,100.0)	76.3 (65.4, 85.1)	79.7 (76.5,83.0)
EIA 1	98.4 (97.6, 99.1)	99.0 (98.3, 99.5)	90.0 (81.0, 95.6)	84.1 (81.0,86.9)
EIA 2	98.9 (98.2, 99.4)	99.5 (98.9, 99.8)	91.3 (82.6, 96.5)	92.4 (90.5,94.0)
EIA 3	99.4 (98.8, 99.8)	99.8 (99.4,100.0)	93.8 (85.9, 98.0)	94.1 (92.5,95.5)
EIA 4	99.7 (99.2, 99.9)	99.9 (99.5,100.0)	96.3 (89.3, 99.2)	96.2 (94.8,97.3)

[a]See text for description of sources of sera
[b]N = number of sera
[c]Point estimate
[d]95% confidence interval in brackets

samples were used separately to assess the specificity of various tests or pairs of tests. Samples from the culture positive group that were BPAT or EIA 1 positive were used in a similar manner for the evaluation of sensitivity.

RESULTS

SPECIFICITY, SENSITIVITY AND AGREEMENT

Table II presents the estimates of specificity for each test based on several different groups of samples. The CFT had no false positive results among samples from the negative herds (specificity = 100%) and the highest specificity based on vaccinated cattle in the reactor herds. With the exception of STAT 1, all tests had relatively high ($\geq$ 98.8%) specificities among nonvaccinated cattle. Only the CFT and EIA 4 had estimates of specificity among vaccinated cattle over 96% in both the negative herd and reactor herd samples. Since there were only 80 vaccinated cattle in the negative herds the confidence intervals for those estimates are much wider than those for other estimates in the table.

Table III presents the estimates of actual and relative sensitivities of each test. The STAT 1, EIA 1, HIGT and BPAT all had actual sensitivities over 95%. The STAT 1 and 2 and EIA 1 had the highest relative sensitivities. Confidence intervals for the actual sensitivities are quite wide since they were based on a sample size of only 174.

Table IV presents the percentages of agreement and the kappa statistics for the pairwise comparison of the tests. Percent agreement ranged from 69.3% to 98.6% and was generally highest for pairs of tests with a high level of specificity. The kappa statistic reveals that the two agglutination tests BPAT and STAT 2 had a high level of agreement, as did the CFT with EIA 3 and EIA 4. All other combinations of tests had relatively poor agreement (kappa $<$ 0.4). However, all of the kappa statistics were significantly different from zero.

CONFIRMATORY TEST EVALUATIONS

The specificities obtained when various confirmatory tests (either alone or in pairs) were applied to screening test positive samples from negative herds are given in Table V. Due to the very small number of

TABLE III. Estimates of the "Actual" and "Relative" Sensitivities[a] of Nine Serological Tests for Bovine Brucellosis

Test	"Actual" Sensitivity (%) N[b] = 174	"Relative" Sensitivity (%) N = 290
BPAT	95.4[c] (91.0,98.0)[d]	
STAT 1	98.3 (95.0,99.6)	82.8 (77.8,87.1)
STAT 2	93.1 (88.2,96.4)	54.8 (48.4,61.4)
CFT	92.9 (87.9,96.3)	16.5 (12.9,22.0)
HIGT	96.0 (91.8,98.4)	17.0 (13.2,22.6)
EIA 1	96.6 (92.6,98.7)	34.7 (29.2,42.0)
EIA 2	94.8 (90.4,97.6)	24.7 (19.6,31.7)
EIA 3	94.3 (89.7,97.2)	22.6 (17.9,29.3)
EIA 4	92.5 (87.5,96.0)	18.8 (14.7,24.8)

[a]"Actual" sensitivities were calculated using test results from sera from culture positive cattle. "Relative" sensitivities were calculated using data from BPAT positive samples from reactor herds (see text for method of calculation)
[b]N = number of sera
[c]Point estimate
[d]95% conficence interval in brackets

TABLE IV. Agreement Among Nine Serological Tests for Bovine Brucellosis as Measured by Percent Agreement (above the diagonal) and a Kappa Statistic[a] (below the diagonal)

	BPAT	STAT 1	STAT 2	CFT	HIGT	EIA 1	EIA 2	EIA 3	EIA 4
BPAT		76.9	89.5	93.0	87.8	89.5	90.9	92.0	92.3
STAT 1	0.31		—	71.7	69.3	71.2	71.6	71.7	71.5
STAT 2	0.61	—		95.0	93.6	89.4	92.4	93.5	93.8
CFT	0.26	0.07	0.33		93.6	92.0	96.1	97.6	98.6
HIGT	0.30	0.07	0.30	0.25		89.7	92.5	93.6	93.8
EIA 1	0.29	0.15	0.27	0.23	0.33		—	—	—
EIA 2	0.28	0.11	0.30	0.39	0.37	—		—	—
EIA 3	0.30	0.10	0.30	0.51	0.39	—	—		—
EIA 4	0.28	0.08	0.30	0.62	0.36	—	—	—	

[a]Test results from sera from reactor herds were used in calculations. See text for description of kappa

samples from negative herds that were positive on either the BPAT or EIA 1 (19 samples for each) the confidence intervals for all estimates are quite wide. However, as expected, series interpretation of results from any two tests resulted in a much higher specificity than parallel interpretation of the same two test results. In samples that were BPAT positive, CFT and EIA 4 had the highest individual specificities and in general, pairs of tests incorporating these tests had relatively high specificities when interpreted in series or in parallel.

The sensitivities of various combinations of confirmatory tests when applied to the culture positive samples which were also screening test positive are given in Table VI. Parallel interpretation results in high sensitivities and series interpretation yields somewhat lower values. However, all combinations of tests resulted in sensitivities in excess of 94% in BPAT positive samples and values over 92% in EIA 1 positive samples.

DISCUSSION

SPECIFICITY

The estimates of specificity based on all the data from all cattle in the negative herds are fairly precise because of the relatively large sample

TABLE V. Estimates of the Specificity of Selected Tests for Bovine Brucellosis when Applied to Screening Test Positive Samples from Negative Herds. Tests were Interpreted Alone or in Series or in Parallel with One Other Test[a]

	BPAT Positive Sera[b]		EIA 1 Positive Sera[c]	
Tests Interpreted Alone				
CFT	100.0[d]	(81.6, 100.0)[e]	100.0	(80.3, 100.0)
STAT 2	84.2	(60.1, 96.6)	89.5	(66.4, 98.7)
HIGT	79.0	(54.4, 94.0)	57.9	(34.5, 80.2)
EIA 3	89.5	(66.4, 98.7)	—	
EIA 4	94.7	(74.2, 99.9)	—	
Tests Interpreted in Series				
CFT and STAT 2	100.0	(81.6, 100.0)	100.0	(91.1, 100.0)
CFT and HIGT	100.0	(81.6, 100.0)	100.0	(91.1, 100.0)
CFT and EIA 3	100.0	(81.6, 100.0)	—	
CFT and EIA 4	100.0	(81.6, 100.0)	—	
STAT 2 and HIGT	100.0	(82.4, 100.0)	94.7	(74.2, 99.9)
STAT 2 and EIA 3	100.0	(82.4, 100.0)	—	
STAT 2 and EIA 4	100.0	(82.4, 100.0)	—	
HIGT and EIA 3	88.9	(65.1, 98.6)	—	
HIGT and EIA 4	94.7	(74.2, 99.9)	—	
Tests Interpreted in Parallel				
CFT or STAT 2	88.9	(65.1, 98.6)	94.1	(71.1, 99.9)
CFT or HIGT	77.8	(52.7, 93.6)	58.8	(32.7, 81.4)
CFT or EIA 3	88.9	(65.1, 98.6)	—	
CFT or EIA 4	94.4	(72.4, 99.9)	—	
STAT 2 or HIGT	63.2	(38.1, 83.4)	52.6	(28.9, 74.8)
STAT 2 or EIA 3	73.7	(43.9, 90.9)	—	
STAT 2 or EIA 4	79.0	(54.4, 94.0)	—	
HIGT or EIA 3	79.0	(54.4, 94.0)	—	
HIGT or EIA 4	79.0	(54.4, 94.0)	—	

[a]See text for description of method of interpretation
[b]Sera from the 13 nonvaccinates and 6 vaccinates in negative herds which tested positively on the BPAT
[c]Sera from the 11 nonvaccinates and 8 vaccinates in negative herds which tested positively on EIA 1
[d]Point estimate
[e]95% confidence interval in brackets

TABLE VI. Estimates of the Sensitivity of Selected Tests for Bovine Brucellosis when Applied to Screening Test Positive and Culture Positive Samples. Tests were Interpreted Alone or in Series or in Parallel with One Other Test[a]

	BPAT Positive Sera[b]		EIA 1 Positive Sera[c]	
Tests Interpreted Alone				
CFT	95.7[d]	(91.3, 98.3)[e]	95.2	(90.7, 97.9)
STAT 2	97.0	(93.1, 99.0)	94.1	(89.3, 97.1)
HIGT	97.0	(93.1, 99.0)	98.8	(95.8, 99.9)
EIA 3	96.4	(92.3, 98.7)	—	
EIA 4	95.8	(91.4, 98.3)	—	
Tests Interpreted in Series				
CFT and STAT 2	95.1	(90.5, 97.9)	92.8	(87.7, 96.2)
CFT and HIGT	95.1	(90.5, 97.8)	94.6	(89.9, 97.5)
CFT and EIA 3	95.7	(91.3, 98.3)	—	
CFT and EIA 4	95.7	(91.3, 98.3)	—	
STAT 2 and HIGT	94.6	(89.9, 97.5)	92.9	(87.8, 96.3)
STAT 2 and EIA 3	94.6	(89.9, 97.5)	—	
STAT 2 and EIA 4	94.6	(89.9, 97.5)	—	
HIGT and EIA 3	95.8	(91.4, 98.3)	—	
HIGT and EIA 4	95.2	(90.6, 97.9)	—	
Tests Interpreted in Parallel				
CFT or STAT 2	98.2	(94.7, 99.6)	96.4	(92.3, 98.7)
CFT or HIGT	98.8	(95.6, 99.9)	100.0	(97.8, 100.0)
CFT or EIA 3	97.5	(93.8, 99.3)	—	
CFT or EIA 4	96.9	(92.9, 99.0)	—	
STAT 2 or HIGT	99.4	(96.7, 100.0)	100.0	(97.8, 100.0)
STAT 2 or EIA 3	98.8	(95.7, 99.9)	—	
STAT 2 or EIA 4	98.2	(94.8, 99.6)	—	
HIGT or EIA 3	97.6	(93.9, 99.3)	—	
HIGT or EIA 4	97.6	(93.9, 99.3)	—	

[a]See text for description of method of interpretation
[b]166 sera from the culture positive cattle which tested positively on the BPAT
[c]168 sera from the culture positive cattle which tested positively on EIA 1
[d]Point estimate
[e]95% confidence interval in brackets

size on which they are based. They indicate the specificity of each test when applied to a population of cattle in which vaccination is infrequent. No false positive reactions were produced by the CFT (i.e. specificity = 100%) and the EIA 4, EIA 3 and STAT 2 all had specificities over 99%. Estimates of the specificity of the BPAT, CFT and STAT are likely to be upwardly biased because herds are given free-listed status (and hence were selected for inclusion in the negative herd group) on the basis of negative reactions on those tests.

Two estimates of the specificity of each test when applied to vaccinated cattle are available. The estimates based on vaccinates in the negative herds have the advantage that all of the animals were almost certainly free of infection. However, these estimates were based on only 80 sera and consequently they have wide confidence intervals. The estimates based on data from vaccinates in the problem herds are much more precise but it is possible that some of the positive test results occurred in sera from infected cattle. If this was the case, the estimates of specificity would be conservative. However, the low prevalence of infection in Canada and the partially protective nature of the vaccine make it unlikely that many of these cattle were infected. Consequently, these are probably the best estimates of the specificities of the tests among vaccinated animals. With the exception of the estimates for the STAT 1 and to a lesser extent EIA 1 the two sets of estimates were generally in close agreement.

The CFT had the highest specificity in both vaccinated and nonvaccinated cattle. The EIA 3, EIA 4 and HIGT performed very well in nonvaccinated cattle (specificity = 99.8, 99.9 and 99.8% respectively) but of these three tests only the EIA 4 had an estimate of specificity over 95% in both groups of vaccinated cattle. The results for the BPAT, STAT 2 and CFT in nonvaccinated animals were very similar to those reported previously (12). However, the previous study reported higher specificities for all three tests among vaccinated cattle than were found in this study. This may be due in part to the manner in which reactor herds were chosen for inclusion in this study. Designation of all cattle over five years of age as nonvaccinated may have lowered the estimates of specificity in that group of cattle.

SENSITIVITY

The best estimates of the sensitivity of the tests were provided by the data from the culture positive samples (i.e. "actual sensitivity"). With the exception of the CFT, all tests performed well on these samples with estimates of sensitivity being 98.3, 96.6, 96.0 and 95.4 for the STAT 1, EIA 1, HIGT and BPAT respectively. These results are substantially higher than those previously reported (12).

The "relative sensitivities" presented in Table III are the proportion of the BPAT positive sera from the

reactor herds which were also positive on one of the other tests. As such, they do not measure the true sensitivity of the tests but do provide some additional data about the ability of each test to detect positive cattle. When estimating the relative sensitivity of a test, it is desirable that the test used to identify the samples as positive, have a high level of sensitivity. Previous work identified the BPAT as the most sensitive of the standard tests (12). Adjustment of the relative sensitivities according to the false positive rate (i.e. 1-specificity) will remove some or all of the bias in favour of tests with low specificities.

The high relative sensitivities for STAT 1 and STAT 2 are probably due to the high level of agreement among the agglutination tests (BPAT and STAT) and not due to a high sensitivity of STAT 1 and 2 *per se*. For the other six tests, the only major discrepancy in the ranking of the tests by the "actual" and relative sensitivities was for the HIGT. It had a high ranking according to the "actual" sensitivities (second) but a low one according to the "relative" sensitivities (fifth).

AGREEMENT

Data from the reactor herds were used to assess agreement (Table IV). This was done because these herds provided the best mixture of positive and negative results. With the exception of the values for STAT 1, all percent agreement values were over 80% and many were over 90%. However, many of the samples contributing to this apparent high level of agreement were negative on both tests being compared. Due to the large number of sera in this category, a high level of agreement would be expected, due to chance alone. In order to adjust for the level of agreement due to chance, the kappa statistic was calculated for each pair of tests. It provides a more realistic evaluation of the level of agreement.

There was good agreement between the two agglutination tests, BPAT and STAT. There was also good agreement between the CFT and EIA 3 and EIA 4 and all of those three tests were ones with relatively high specificity. Agreement between all other pairs of tests was poor. Since the reactor herds probably contained a number of false positive reactors, the lack of agreement suggests that it is not the same samples which provide the false positive reactions on all tests. This supports the idea that it is beneficial to apply more than one test in problem situations and interpret the results according to the objectives of the control program.

CONFIRMATORY TESTS

The main requirements of a screening test are that it be rapid and economical since it will have to be applied to a large number of samples and that it be sensitive so that few positive animals are missed. The test must also have a reasonable level of specificity to prevent large numbers of samples having to be tested on confirmatory tests or a large number of animals rebled. Based on data obtained from the previous analyses and the experience of the authors, the BPAT and EIA 1 were selected as potential screening tests.

The CFT, STAT 2, HIGT, EIA 3 and EIA 4 were chosen for further evaluation as confirmatory tests, primarily because of their relatively high specificities. However, the EIA 3 and EIA 4 were not considered as possible confirmatory tests following EIA 1 screening.

The evaluation of the tests in this project has assumed that all sera were independent units. In reality, some cows contributed multiple sera, cows are members of herds and regulatory veterinarians have information such as herd history to consider along with test results when making decisions about reactors. This, along with the fact that there is generally poor agreement among the various tests suggests that it is desirable to carry out two confirmatory tests and interpret the results in light of other data and in accordance with the objectives of the testing program.

Evaluation of the specificity of the various confirmatory tests is difficult since it requires a group of samples which are known to be negative but which test positive on a screening test. Only 19 samples from the negative herds tested positive to each of the BPAT and EIA 1 screening tests. Consequently the estimates of specificity have very wide 95% confidence intervals. However, with the exception of the CFT, all tests when interpreted alone had considerably lower point estimates of specificity when compared to estimates obtained from the entire population of cattle in the negative herds. The two groups of 19 sera that were BPAT and EIA 1 positive contained sera from six and eight vaccinates respectively. This larger proportion of vaccinates would account, in part, for the lower specificities. However, in general, the estimates of specificity based on the screening test positive samples were even less than those estimates obtained from only vaccinated cattle in the reactor herds. This supports the necessity for determining the sensitivity and specificity of a test within the population to which the test is applied.

Interpretation of two tests in series usually resulted in a very high specificity but interpretation in parallel appears to produce false positive rates that range from 11 to 37% following BPAT screening and 6 to 47% following EIA 1 screening. The exception to this was the combination of CFT and EIA 4 which had a false positive rate of only 6% following BPAT screening. This suggests that care should be taken in the parallel interpretation of two tests.

The estimation of the sensitivity of the selected confirmatory tests (or pairs of tests) was carried out using culture positive samples that had also tested positive on one of the two screening tests. Since sample sizes of 165 and 168 were available for the BPAT and EIA 1 positive groups respectively, the confidence intervals are narrower than those for the specificities.

Since most of the sera from culture positive groups were included in these analyses, the estimates of the sensitivity of each test alone are similar to those presented earlier and are generally high. Since the sensitivities of each test taken alone remained high, the overall sensitivity of pairs of tests interpreted in series or in parallel also remained high. Series interpretation of two tests did not have as deleterious effect on the sensitivity of the testing program as parallel interpretation had on the specificity.

Selection of a test or group of tests for use in a screening program should

be based upon the attributes of the test(s) and the objectives of the program. One approach to guide the selection of a test is to identify the minimum acceptable specificity and of the tests meeting that criteria, choose the one with the highest sensitivity (or vice versa). For example, if a screening program requires specificities of 98% and 90% amongst nonvaccinates and vaccinates respectively, then EIA 1 (sensitivity = 96.6%) and BPAT (sensitivity = 95.4%) could be considered as screening tests. Another approach to the selection of tests is the use of a relative operating characteristic analysis (13). This technique is beyond the scope of this paper.

Based on the data from this study, either the BPAT or EIA 1 appear to be acceptable screening tests. Given the potential upward bias in the estimates of the sensitivity of the BPAT, the EIA 1 may be the test of choice. The enzyme immunoassay has the additional advantages that it is quantitative at a single dilution (and hence different thresholds could be used for vaccinates and nonvaccinates), results are determined objectively and it can be automated and linked to electronic data processing systems. It suffers the disadvantage that it requires a much greater capital investment than the BPAT.

Assuming that high specificity is a requirement of a confirmatory test, then the CFT or EIA 4 are the tests of choice. Following BPAT screening, the CFT had the highest specificity (100% vs 94.7% for the EIA 4) but as discussed earlier, the estimate for the CFT is likely upwardly biased by the manner in which the negative herd samples were identified. Also, these estimates of specificity must be interpreted with caution given the small sample sizes on which they were based. In addition to the advantages described above, the enzyme immunoassay has the advantages that it is not affected by anticomplementary activity and can be carried out on hemolyzed samples.

ACKNOWLEDGMENTS

Full responsibility for the data base design, creation and manipulation belonged to D. Gall. His excellent efforts in this area were greatly appreciated. The assistance of A. Boyd, B. Kirk, L. Kelly, M. Mallory and M. Masner in the serological testing of sera samples is likewise acknowledged. Dr. J. Stevens produced all of the antigens used in the tests. Dr. M. Garcia, with assistance from D. Martin and H. Kelly, was responsible for the bacteriological investigations conducted in the reactor herds. The assistance of the field staff of the Food Production and Inspection Branch in the collection of the appropriate samples is gratefully acknowledged.

REFERENCES

1. **MARTIN SW.** The evaluation of tests. Can J Comp Med 1977; 41: 19-25.
2. **NEILSEN K, WRIGHT PF.** Enzyme immunoassay and its application to the detection of bovine antibody to *Brucella abortus*. Agriculture Canada, Animal Diseases Research Institute, Nepean, Ontario, 1984.
3. **ANGUS RD, BARTON CE.** The production and evaluation of a buffered plate antigen for use in a presumptive test for brucellosis. In: Proceedings of the Third International Symposium on Brucellosis, Algiers, Dev Biol Stand 1983; 51: 349-356.
4. **MALKIN KL, TAILYOUR JM, BHATIA TRS, ARCHIBALD RM, DORWARD WJ.** A serological survey for brucellosis in Canadian swine. Can J Comp Med 1968; 32: 598-599.
5. **STEMSHORN BW, NEILSEN KH, SAMAGH BS, FORBES LB, INGRAM DG.** Evaluation of an enzyme-labeled antiglobulin test for anti-*Brucella* immunoglobulin G among 3 cattle populations. Am J Vet Res 1980; 41: 1779-1784.
6. **SAMAGH BS, BOULANGER P.** Adaptation of the Technicon Autoanalyser II for an automated complement fixation test for the detection of antibodies to *Brucella abortus* in bovine serums. Proc 21st Annu Meet Am Assoc Vet Lab Diagnost 1978: 347-366.
7. **RUCKERBAUER GM, STEMSHORN BW, NEILSEN KH.** An hemolysis-in-gel test for anti-*Brucella* antibody in cattle serum. In: Butler JE, ed. The ruminant immune system. New York: Plenum Press, 1981: 782-783.
8. **STILLER JM, NIELSEN KH.** Affinity purification of bovine antibodies to Brucella abortus lipopolysaccharide. J Clin Microbiol 1983; 17: 323-326.
9. **RUCKERBAUER GM, GARCIA MM, RIGBY CE, ROBERTSON FJ, SAMAGH BS, STEMSHORN BW.** An hemolysis-in-gel test for bovine brucellosis. Dev Biol Stand 1984; 56: 513-520.
10. **OSTLE B, MENSING RW.** Statistics in medicine. 3rd ed. Ames, Iowa: Iowa State U Press, 1975.
11. **FLEISS JL.** Statistical methods for rates and proportions. New York: John Wiley & Sons, 1973.
12. **STEMSHORN BW, FORBES LB, EAGLESOME MD, NIELSEN KH, ROBERTSON FJ, SAMAGH BS.** A comparison of standard serological tests for the diagnosis of bovine brucellosis in Canada. Can J Comp Med 1985; 49: 391-394.
13. **ERDREICH LS, LEE ET.** Use of relative operating characteristic analysis in epidemiology. A method for dealing with subjective agreement. Am J Epidemiol 1981; 114: 649-662.