Serological diagnosis of bovine brucellosis: a review of test performance and cost comparison

D. Gall & K. Nielsen

Canadian Food Inspection Agency, Animal Diseases Research Institute, Brucellosis Centre of Expertise, 3851 Fallowfield Road, Ottawa, Ontario, K2H 8P9, Canada

Submitted for publication: 10 May 2004
Accepted for publication: 2 August 2004

Summary
The authors reviewed over 50 publications in which the sensitivity and specificity values of assays used for the detection of exposure to Brucella abortus had been examined. The sum of the sensitivity and specificity values for each test was averaged to give a performance index (PI) and allow for a comparison between the different methodologies. A score of 200 was perfect. Based on the PI, the buffered antigen plate agglutination test (BPAT) rated highest (PI = 193.1) among the conventional tests. This indicates better accuracy than the other conventional tests including the Rose Bengal test (PI = 167.6) and the complement fixation test (PI = 172.5). Overall, the primary binding assays, including the fluorescence polarisation assay (PI = 196.4), the indirect enzyme-linked immunosorbent assay (PI = 189.8) and the competitive enzyme-linked immunosorbent assay (PI = 188.2), were more accurate than the conventional tests, except for the BPAT. In addition, a fee comparison suggested that the primary binding tests were price competitive with conventional tests for the diagnosis of brucellosis and, therefore, had a better combined cost/efficiency rating.

Keywords

Introduction

The development of the first agglutination test for the detection of antibody to Brucella infection was reported by Wright and Smith (81) over 100 years ago. Since then a great deal of work has been done to improve diagnostic methods and accuracy, culminating in the production of primary binding assays and polymerase chain reaction (PCR) procedures.

Briefly, primary binding assays directly measure the interaction of antibody and antigen while conventional serological tests, such as acidified agglutination tests or the complement fixation test (CFT), measure secondary phenomena such as the agglutination or activation of complement (43). Depending on the sensitivity (i.e. the ability of a test to correctly identify an animal with disease) and specificity (i.e. the ability of a test to correctly identify healthy animals or animals not having the disease), tests can be used to screen for, or confirm, disease. Traditionally, screening tests are inexpensive, fast and highly sensitive, but not necessarily highly specific. Confirmatory tests are required to be both sensitive and specific (70). The buffered antigen plate agglutination test (BPAT), enzyme-linked immunosorbent assays (ELISA) and the fluorescence polarisation assay (FPA) are appropriate screening tests (52) since they are all highly sensitive and specific, making them ideal tests for use in international trade. Additionally, the competitive enzyme-linked immunosorbent assay (CELISA) and the FPA have the capability to distinguish between animals vaccinated with the widely used Brucella abortus strain 19 and animals infected with B. abortus (27, 38, 41, 42).

Sera from animals from which B. abortus has been isolated are a good source of defined sera, which can be used as
reference sera to determine the sensitivity of a test. In the absence of bacterial isolation, other methods, such as another serological test or combination of tests with known sensitivity and specificity estimates can be used to define the status of animals, providing the necessary reference sera to determine the sensitivity and specificity of the new test relative to the other tests or combination of tests (21, 30). However, the specificity of a test cannot usually be determined by bacteriological isolation because some animals that yield negative culture results are in fact infected (6, 29). Reasons for this may be the condition of the tissues submitted, improper storage of tissues, not selecting an appropriate variety of tissues, not selecting a sufficient amount of tissues, or selecting samples from uninfected tissues. Estimates of specificity can be determined using populations of animals known to be free of disease.

For many regulatory agencies and diagnosticians it is difficult to determine from the scientific literature which tests are appropriate for use in control programmes (i.e. ongoing diagnostic monitoring programmes which detect positive animals and then implement control measures such as quarantine or slaughter), eradication programmes (i.e. ongoing programmes to reduce the prevalence or incidence of the disease in the population) and surveillance programmes (i.e. continuous surveys of populations focused on maintaining disease free status of the population, or case finding for disease control purposes and the provision of information to formulate policies). With a large number of test methodologies available for the diagnosis of brucellosis, a review of published sensitivity and specificity values that summarises the accuracy and cost effectiveness of the various tests would be a useful tool (Table 1).

The purpose of this study was to review the sensitivity and specificity values of the various brucellosis tests that have been published in peer reviewed journals, to determine which tests were consistently more accurate for the diagnosis of brucellosis and which tests were most cost-effective. Depending on the study, the sensitivity and/or the specificity values could contain data from animals vaccinated with *B. abortus* strain 19. Due to the different methodologies and experimental or validation designs, the review was limited to those sensitivity and specificity values published in peer-reviewed publications. Some earlier publications combined results and did not distinguish between animals which tested positive for antibody to *B. abortus* due to infection and those which tested positive due to previous vaccination with *B. abortus* strain 19. In addition, the sample sizes for the defined reference populations varied amongst publications, which would have an effect on the sensitivity and specificity values obtained and, in turn, on the performance index (PI).

Where available, data from advertised internet websites and personal communication with other organisations (government funded, university or private sector laboratories) were compiled to assess the relative unit fee per sample per test. Some laboratories used reagents produced in-house to reduce their costs while others used commercially available kits, which usually increased their costs. None of the laboratories indicated the manner in which their fee schedule was determined. However, factors including labour costs and overhead costs (e.g. heating, lighting and maintenance) would have been included. Most of the laboratories were not accredited or certified for ISO 17025 or 9001, respectively, which would also have had an impact on costs, quality and the international harmonisation of the results.

**Conventional serological tests**

Conventional serological tests such as the 2-mercaptoethanol (2ME) test, the Card test, the CFT for serum and milk (mCFT), the milk ring test, the plate agglutination test, the Rivanol test (RIV), the Rose Bengal test (RBT), and the tube agglutination test (TAT) were performed using reagents produced locally or purchased from other sources (some of the tests used standard procedures and reagents and others used modified standard procedures and/or reagents) (2, 3, 4, 22, 28, 32, 33, 51, 52, 63, 74, 75). The standardisation of reagents and use of control sera varied between the reviewed publications. Earlier publications described procedures that followed the internationally accepted laboratory practices of the time while more recent evaluations followed OIE (World Organisation for Animal Health) or United States Department of Agriculture procedures.

**Precipitation tests**

The agar gel immunodiffusion test (AGID), radial immunodiffusion test (RID), haemolysis-in-gel test (HIGT) and indirect haemolysis test (IHLT) were conducted according to described procedures (8, 24, 54, 55, 59).

**Materials and methods**

### Internet search criteria

Using internet search engines, a literature review of all tests with published sensitivity and specificity values for the detection of exposure to *Brucella* spp. in cattle was undertaken. The majority of the sensitivity results were based on data from animals from which *Brucella* spp. were isolated, while the specificity results were based on animals that were free of infection (i.e. animals from which *Brucella* spp. were not isolated or those without clinical, epidemiological or serological evidence of infection).
Table I
List of references cited for each serological test for bovine brucellosis included in this study

<table>
<thead>
<tr>
<th>Tests</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-mercaptoethanol test (2ME)</td>
<td>Lord et al. (24), Nicoletti (34), Saravi et al. (66) and Stemshorn et al. (70)</td>
</tr>
<tr>
<td>Agar gel immunodiffusion test (AGID)</td>
<td>Lord et al. (24)</td>
</tr>
<tr>
<td>Allergic skin test (AST)</td>
<td>Pouillot et al. (56) and Saegerman et al. (62)</td>
</tr>
<tr>
<td>Blood fluorescence polarization assay (bld FPA)</td>
<td>Niels et al. (47, 50)</td>
</tr>
<tr>
<td>Buffered antigen plate agglutination test (BPAT)</td>
<td>Dohoo et al. (9), Fosgate et al. (10), Nielsen et al. (40, 41, 43), Nielsen and Gall (48), Nielsen et al. (50), Ruckerbauer et al. (60), Samartino et al. (65), Saravi et al. (66), Stemshorn et al. (70), Uzal et al. (76, 77) and Vanzini et al. (79)</td>
</tr>
<tr>
<td>Bulk milk tank fluorescence polarization assay (bmFPA)</td>
<td>Gall et al. (13)</td>
</tr>
<tr>
<td>Card test</td>
<td>Fosgate et al. (10), Greenlee et al. (14), Hall and Confer (16), Huber and Nicoletti (19), Jones et al. (23), Lord et al. (24), Nicoletti (34), Nicoletti and Tanya (35), Nielsen et al. (46) and Stemshorn et al. (70)</td>
</tr>
<tr>
<td>Competitive enzyme-linked immunosorbent assay (CELISA)</td>
<td>Fosgate et al. (11), Gall et al. (12), McGiven et al. (25), Nielsen et al. (41, 43), Nielsen and Gall (48), Nielsen et al. (50), Rylatt et al. (81), Samartino et al. (65) and Stack et al. (69)</td>
</tr>
<tr>
<td>Complement fixation test (CFT)</td>
<td>Diaz et al. (8), Dohoo et al. (9), Greenlee et al. (14), Hall et al. (15), Hall and Confer (16), Huber and Nicoletti (19), Jones et al. (23), Lord et al. (24), McGiven et al. (25), Nicoletti (34), Nicoletti and Tanya (35), Nielsen et al. (41, 43, 47), Nielsen and Gall (48), Nielsen et al. (50), Paweska et al. (53), Ruckerbauer et al. (60), Rylatt et al. (61), Saravi et al. (66), Stack et al. (69), Stemshorn et al. (70), Sutherland (71), Sutherland et al. (72), Sutherland and MacKenzie (73), Uzal et al. (76, 77) and Van Aert et al. (78)</td>
</tr>
<tr>
<td>Fluorescence immunoassay (RIA)</td>
<td>Hall et al. (15) and Hall and Confer (16)</td>
</tr>
<tr>
<td>Fluorescence polarization assay (FPA)</td>
<td>Dajer et al. (7), McGiven et al. (25), Nielsen et al. (42, 46, 47), Nielsen and Gall (48) and Samartino et al. (64)</td>
</tr>
<tr>
<td>Haemolysis-in-gel test (HIGT)</td>
<td>Dohoo et al. (9) and Ruckerbauer et al. (60)</td>
</tr>
<tr>
<td>Indirect enzyme-linked immunosorbent assay (IELISA)</td>
<td>Abalos et al. (1), Dohoo et al. (9), Gall et al. (12), Hall and Confer (16), McGiven et al. (25), Nielsen et al. (35), Nicoletti and Tanya (35), Nielsen et al. (40), Nielsen et al. (41), Nielsen and Gall (48), Nielsen et al. (50), Niels et al. (43), Paweska et al. (53), Renukardhy et al. (57), Romero et al. (58), Rylatt et al. (61), Samartino et al. (65), Saravi et al. (66), Sutherland (71), Uzal et al. (76, 77), Van Aert et al. (78) and Vanzini et al. (79)</td>
</tr>
<tr>
<td>Indirect haemolysis test (IHLT)</td>
<td>Sutherland (71), Sutherland et al. (72) and Sutherland and MacKenzie (73)</td>
</tr>
<tr>
<td>Milk complementation test (mCFT)</td>
<td>Hunter and Allen (20)</td>
</tr>
<tr>
<td>Milk indirect enzyme-linked immunosorbent assay (mIELISA)</td>
<td>Gall et al. (13), Nielsen et al. (44), Nielsen and Gall (48), Nielsen et al. (49), Nielsen et al. (50), Vanzini et al. (79) and Vanzini et al. (80)</td>
</tr>
<tr>
<td>Milk fluorescence polarization assay (mFPA)</td>
<td>Gall et al. (13), Nielsen and Gall (48) and Nielsen et al. (50)</td>
</tr>
<tr>
<td>Milk ring test (MRT)</td>
<td>Huber and Nicoletti (19), Hunter and Allen (20), Nicoletti (34), Nielsen and Gall (48) and Vanzini et al. (80)</td>
</tr>
<tr>
<td>Particle concentration fluorescence immunoassay (PCFA)</td>
<td>Greenlee et al. (14), Nicoletti and Tanya (35) and Nielsen et al. (46)</td>
</tr>
<tr>
<td>Plate agglutination test (PAT)</td>
<td>Fosgate et al. (10), Nielsen et al. (36), Ruckerbauer et al. (60) and Stemshorn et al. (70)</td>
</tr>
<tr>
<td>Polymerase chain reaction (PCR)</td>
<td>Bricker et al. (5), Hamdy and Amin (17) and Romero et al. (58)</td>
</tr>
<tr>
<td>Radial immunodiffusion (RID)</td>
<td>Diaz et al. (8) and Jones et al. (23)</td>
</tr>
<tr>
<td>Rivanol test (RIV)</td>
<td>Dajer et al. (7), Hall et al. (15), Hall and Confer (16), Huber and Nicoletti (19), Jones et al. (23), Lord et al. (24), Nicoletti (34) and Nicoletti and Tanya (35)</td>
</tr>
<tr>
<td>Rose Bengal test (RBT)</td>
<td>Dajer et al. (7), Diaz et al. (8), Hunter and Allen (20), Samartino et al. (65), Saravi et al. (66), Stemshorn et al. (70), Sutherland (71), Sutherland et al. (72), Sutherland and MacKenzie (73) and Van Aert et al. (78)</td>
</tr>
<tr>
<td>Tube agglutination test (TAT)</td>
<td>Dohoo et al. (9), Fosgate et al. (10), Hall et al. (15), Hunter and Allen (20), Lord et al. (24), McGiven et al. (25), Nicoletti (34), Nielsen et al. (36), Ruckerbauer et al. (60), Stemshorn et al. (70), Sutherland (71), Sutherland et al. (72) and Van Aert et al. (78)</td>
</tr>
</tbody>
</table>
Primary binding tests

Indirect enzyme-linked immunosorbent assays (IELISAs) (for the detection of antibodies in serum and milk) and CELISAs with the ability to distinguish vaccinated animals from animals infected with *Brucella* spp. were performed according to standard procedures, some with modifications as described by the respective authors (37, 39, 41, 45, 51, 52, 61, 69).

Fluorescence polarisation assays used for serum (FPA), milk (mFPA) and bulk milk tank samples (bmFPA) and whole blood (BldFPA) were performed as described in the relevant papers (13, 25, 42, 47, 49, 50). Likewise, fluorescence immunoassays (FIA) and particle concentration fluorescence immunoassays (PCFIA) were conducted according to described procedures (15, 35, 68).

Isolation of organisms

Most of the published evaluations used some variation of the standard procedure for the isolation of *Brucella* spp. (2, 3, 51, 52).

Polymerase chain reactions

Polymerase chain reactions (PCRs) PCRs were conducted according to described protocols (5, 17, 58).

Skin tests

Allergic skin tests (AST) were conducted as described in the relevant papers (56, 62).

Data analysis

The PI was used as a measure of accuracy and was calculated by adding the sensitivity and specificity of each test used in each study. Performance indices and their associated mean specificity and sensitivity values are tabulated in Table II. Similar to Youden's index (83), the PI is one way to summarise test accuracy in a single numeric value. The average of all PI values for each test was charted in ascending order as shown in Figure 1. Also calculated for each PI average, and shown in Table II, are the standard deviation (SD) and percent coefficient of variation (%CV), which are measures of uncertainty.

Relative unit cost per sample per test was calculated using the average fee from different sources converted to a common denominator value to facilitate comparison.

Results

Figure 1 and Table II represent data used to calculate the accuracy of each diagnostic test. The average PI for each
The test was calculated using two or more PIs, except for the tests for which there was only one publication (i.e. the AGID test, bmFPA, the mCFT and bacterial isolation). The mean PI value for the RBT (n = 12,146) was based on eleven PIs calculated from eleven individual sets of sensitivity and specificity values extracted from the literature between 1972 and 1999. Performance index values for each of the 11 sets of sensitivity and specificity values are depicted in Figure 2. Similarly, the mean PI value for the CFT (n = 28,537) was based on 38 PIs calculated from 38 sets of sensitivity and specificity values reported in the literature between 1969 and 2003 (Fig. 3). Performance index data calculated from 37 sets of sensitivity and specificity values published between 1984 and 2003, and used to determine the mean PI value for the IELISA (n = 60,985), are presented in Figure 4.

The PCFIA (n = 2,436) was identified as the least accurate assay while the FPA (n = 39,934) performed with the highest accuracy (Figure 1). Performance index estimates for the PCFIA and FPA were based on six and seven sets of sensitivity and specificity data, respectively.

### Table II

<table>
<thead>
<tr>
<th>Test</th>
<th>Total number of samples tested</th>
<th>Mean of sensitivities</th>
<th>Mean of specificities</th>
<th>Mean PI</th>
<th>SD</th>
<th>% CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Particle concentration fluorescence immunoassay (n = 6)</td>
<td>2,436</td>
<td>91.8</td>
<td>46.7</td>
<td>138.5</td>
<td>21.6</td>
<td>15.6</td>
</tr>
<tr>
<td>Culture (n = 1)</td>
<td>102</td>
<td>46.1</td>
<td>100.0</td>
<td>146.1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Card (n = 11)</td>
<td>6,434</td>
<td>91.0</td>
<td>55.2</td>
<td>146.2</td>
<td>31.3</td>
<td>21.4</td>
</tr>
<tr>
<td>Rivianol test (RIV) (n = 12)</td>
<td>4,845</td>
<td>89.6</td>
<td>63.1</td>
<td>152.7</td>
<td>25.3</td>
<td>16.6</td>
</tr>
<tr>
<td>Milk ring test (n = 6)</td>
<td>7,328</td>
<td>89.5</td>
<td>74.5</td>
<td>164.0</td>
<td>22.2</td>
<td>13.6</td>
</tr>
<tr>
<td>Indirect haemolysis test (n = 3)</td>
<td>970</td>
<td>91.0</td>
<td>75.4</td>
<td>166.4</td>
<td>40.9</td>
<td>24.6</td>
</tr>
<tr>
<td>Rose Bengal test (n = 11)</td>
<td>12,146</td>
<td>81.2</td>
<td>86.3</td>
<td>167.6</td>
<td>24.8</td>
<td>14.8</td>
</tr>
<tr>
<td>Tube agglutination test (n = 14)</td>
<td>9,534</td>
<td>75.9</td>
<td>95.7</td>
<td>171.6</td>
<td>20.1</td>
<td>11.7</td>
</tr>
<tr>
<td>Complement fixation test (n = 38)</td>
<td>28,537</td>
<td>89.0</td>
<td>83.5</td>
<td>172.5</td>
<td>24.3</td>
<td>14.1</td>
</tr>
<tr>
<td>Plate agglutination test (n = 4)</td>
<td>2,754</td>
<td>77.1</td>
<td>96.0</td>
<td>173.1</td>
<td>8.5</td>
<td>4.9</td>
</tr>
<tr>
<td>Milk complement fixation test (n = 1)</td>
<td>2,147</td>
<td>89.0</td>
<td>86.0</td>
<td>175.0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Allergic skin test (n = 3)</td>
<td>3,811</td>
<td>78.3</td>
<td>99.7</td>
<td>178.0</td>
<td>14.8</td>
<td>8.3</td>
</tr>
<tr>
<td>Haemolysis-in-gel test (n = 3)</td>
<td>2,450</td>
<td>80.1</td>
<td>93.3</td>
<td>179.4</td>
<td>27.2</td>
<td>15.2</td>
</tr>
<tr>
<td>2-mercaptoethanol test (n = 4)</td>
<td>7,693</td>
<td>88.4</td>
<td>91.5</td>
<td>179.9</td>
<td>21.3</td>
<td>11.8</td>
</tr>
<tr>
<td>Polymerase chain reaction (n = 5)</td>
<td>539</td>
<td>82.0</td>
<td>98.6</td>
<td>180.6</td>
<td>21.4</td>
<td>11.9</td>
</tr>
<tr>
<td>Radial immunodiffusion (RID) (n = 2)</td>
<td>1,525</td>
<td>90.6</td>
<td>90.0</td>
<td>180.6</td>
<td>18.2</td>
<td>10.1</td>
</tr>
<tr>
<td>Fluorescence immunoassay (n = 3)</td>
<td>680</td>
<td>88.1</td>
<td>94.6</td>
<td>182.7</td>
<td>16.9</td>
<td>9.3</td>
</tr>
<tr>
<td>Agar gel immunodiffusion test (n = 1)</td>
<td>156</td>
<td>87.5</td>
<td>100.0</td>
<td>187.5</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Milk fluorescence polarisation assay (n = 3)</td>
<td>5,580</td>
<td>88.8</td>
<td>99.2</td>
<td>188.0</td>
<td>11.2</td>
<td>6.0</td>
</tr>
<tr>
<td>Competitive enzyme-linked immunosorbent assay (n = 14)</td>
<td>15,865</td>
<td>97.7</td>
<td>90.5</td>
<td>188.2</td>
<td>16.6</td>
<td>8.8</td>
</tr>
<tr>
<td>Indirect enzyme-linked immunosorbent assay (n = 37)</td>
<td>60,985</td>
<td>96.0</td>
<td>93.8</td>
<td>189.8</td>
<td>18.1</td>
<td>9.5</td>
</tr>
<tr>
<td>Buffered antigen plate agglutination test (n = 15)</td>
<td>60,634</td>
<td>95.4</td>
<td>97.7</td>
<td>193.1</td>
<td>6.3</td>
<td>3.3</td>
</tr>
<tr>
<td>Milk indirect enzyme-linked immunosorbent assay (n = 7)</td>
<td>16,921</td>
<td>97.9</td>
<td>96.1</td>
<td>194.0</td>
<td>5.6</td>
<td>2.9</td>
</tr>
<tr>
<td>Blood fluorescence polarisation assay (n = 2)</td>
<td>1,385</td>
<td>97.2</td>
<td>98.4</td>
<td>195.6</td>
<td>4.3</td>
<td>2.2</td>
</tr>
<tr>
<td>Bulk milk tank fluorescence polarisation assay (n = 1)</td>
<td>258</td>
<td>100.0</td>
<td>95.9</td>
<td>195.9</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Fluorescence polarisation assay (n = 7)</td>
<td>39,934</td>
<td>97.5</td>
<td>98.9</td>
<td>196.4</td>
<td>4.4</td>
<td>2.2</td>
</tr>
</tbody>
</table>

### Fig. 2

Performance indices (the average of the sum of the sensitivity and specificity values) of 11 Rose Bengal tests reported in the literature between 1972 and 1999.
Conventional tests tended to cluster at the lower end of the PI scale while primary binding assays tended to cluster at the upper end of the scale. The exception was the BPAT, which was consistently accurate at a high level (Fig. 5) with little variation (Table II) when compared to the Card test, the Rivanol test, the RBT and the TAT. The CFT was found to be the least accurate test with the most variation throughout the early years; it was more accurate, although less variable, in the last ten years (Fig. 6). When compared to the CFT, the primary binding assays (IELISA, CELISA and FPA) were significantly more accurate and stable (Fig. 6, Table II).

Figure 2 demonstrates the variation inherent in the RBT due to the subjectivity in the interpretation of individual tests and the probable lack of standardisation in antigen production. The IELISA (Fig. 4) was less variable than the CFT (Fig. 3) and the RBT (Fig. 2). The %CV for the RBT,
CFT and IELISA were 14.8%, 14.1% and 9.5%, respectively (Table II).

Table II lists the mean PI (in descending order), SD and %CV for each test. The majority of the conventional tests have higher %CVs than the primary binding assays, which is probably an indication of the subjectivity versus the objectivity of the instrumentation. The %CVs and the associated SD represent measures of uncertainty about the mean PI indicating the amount of variation associated with each mean PI.

Although the PCFIA was the least accurate test, it was the third most expensive serological test after bacteriological isolation and PCR. The least expensive cost per unit sample was for the FPA. The conventional tests were almost as costly, and generally less accurate, to perform as the primary binding tests (Fig 7).

Discussion

The BPAT was found to be the most accurate of the conventional tests. Originally, Angus and Barton (4) developed this protocol to replace the acidified plate antigen test, the buffered Brucella antigen test (also known as the Card) and the RBT, to reduce the non-specific reactions. Their study showed that the BPAT was more sensitive than the other conventional tests for the detection of B. abortus in bovine serum. This was attributed in part to the instability of some of the antigen preparations used in the other tests. In addition, MacMillan (26) reported that the RBT antigen could deteriorate when repeatedly cycled between refrigerator and room temperature during use. The data presented in Figure 5 support the original conclusions made by Angus and Barton (4) and provide evidence that the preparation of the BPAT antigen, and thus the performance, has remained unchanged since its development. By comparison, the concentration of whole cells and the pH of the antigen preparations used in other agglutination tests, such as the RBT, vary within and between countries (31) and result in variations in the accuracy of these tests (Fig. 2). Consistent antigen production may explain why the BPAT has remained more repeatable and reproducible than the other conventional tests.

The PI of the CFT was lower than the PI of the CELISA, IELISA and FPA, and was the most variable (Table II). This was probably due to a combination of factors including inconsistent results and studies, the complexity and labour intensity of the test, lack of standardisation, inability to test haemolysed samples, anti-complementary results, prozoning (a secondary phenomenon where the concentration of antigen or antibody, or both, are outside the zone of equivalence), false positive reactions in samples...
collected from B. abortus strain 19 vaccinated animals, and detection of cross-reacting antibody resulting from exposure to other micro-organisms such as Salmonella urbana O:30, Escherichia coli O:116 and O:157, and Yersinia enterocolitica serotype 9 (26). Previous studies suggested that the CFT is a highly sensitive and specific test when performed correctly, but the data reviewed in this study were ambivalent. Results reported in early studies, and occasionally in more recent investigations as well (25), contradict the widely held consensus that the CFT is a highly sensitive and specific test. However, beginning in 1995 the CFT showed improved performance, as depicted in Figure 3. This may be associated with the recent recognition of and improvements in quality assurance (QA). As an OIE prescribed test for international trade the inconsistent performance by the CFT must be troublesome to reference laboratories. Exchange of proficiency panels between laboratories would assist in harmonising results and improve diagnostic performance nationally and internationally.

The primary binding assays (i.e. CELISA, IELISA and the FPA) were developed as more sensitive and specific alternatives to conventional tests. Conventional tests and the IELISA were unable to distinguish between B. abortus strain 19 vaccinated animals and naturally infected animals. Vaccination induces antibody thought to be of lower affinity due to a short exposure time to the antigen because it is eliminated by the immune system. Alternatively, antibody produced in response to natural infection is of higher affinity because the antigen is not removed as quickly by the immune system and, therefore, persists for a much longer period (27, 38). Thus, the CELISA and the FPA were developed to overcome this problem. These tests are capable of distinguishing vaccinated animals or animals infected with cross-reacting organisms from naturally infected animals, thereby reducing the number of false positive reactions and subsequent trace backs or slaughter of animals in an otherwise negative or healthy population. The use of these tests together with the development of the B. abortus RB51 vaccine, which contains no O-polysaccharide on its cell surface (67), can reduce the number of false positive reactions in healthy populations, thereby reducing laboratory and producer costs. The primary binding assays have been shown to have better sensitivity (defined by the ratio of the number of animals positively identified by a test to the number of truly infected animals [18]) than conventional tests and other tests, including the AGID test, AST, FIA, HIGT, IHLT, PCFIA and RID assay. This observation is supported by PIs calculated in the present study from the sum of sensitivities and specificities reported in peer-reviewed publications. As previously mentioned, tests that are highly sensitive, inexpensive and rapid are ideal as screening tests. Confirmatory tests should ideally be both highly sensitive and specific. The primary binding tests fit the criteria for both screening and confirmatory tests. Another protocol used by some laboratories combines tests of lesser sensitivity and specificity by testing samples in series or parallel. Series testing denotation means that all of the samples tested are positive on all of the tests whereas parallel testing denotation means that the samples are positive on one or more of the tests or all of the tests (29). Parallel testing increases the sensitivity of combined tests while series testing increases the specificity of combined tests. Both series and parallel testing may be useful protocols for providing an orderly transition from the use of conventional tests to primary binding tests by diagnostic laboratories, however, the use of a panel of tests increases costs and delays the output of results.

Although the AGID test, AST, FIA, HIGT, IHLT, PCFIA and RID were performed in a number of different laboratories, fewer data were available in the literature for these tests compared to the BPAT, CFT and the IELISA. This may be due, in part, to labour requirements, reagent costs, and longer test performance times making these tests less suitable for use in eradication programmes or in laboratories with large diagnostic workloads. More recent tests, such as the FPA, are more cost-effective due to lower reagent costs, automation, and the objective means of assessment (versus subjective assessment), which in turn makes them more suitable for modern ISO accredited diagnostic laboratories. While not absolute, the data in Figure 7 show that when compared to the conventional tests the CELISA and FPA are generally less expensive (with the exception of the BPAT) and the IELISA is slightly more expensive. Thus, the primary binding tests are ideal screening or confirmatory tests because of their low unit cost and higher sensitivity and specificity estimates. Combining the primary binding tests in either series or parallel only marginally increases either sensitivity or specificity, however, combining these tests with other conventional tests may potentially increase the overall sensitivity or specificity for a diagnostic laboratory. In addition, combining a primary binding test and a conventional test has the added advantages of reducing false positive reactions (due to vaccination with strain 19 or due to cross-reacting organisms) and reducing testing and programme costs.

The PIs for both PCR and bacterial isolation of B. abortus, based on results in a limited number of publications, were generally poor. Factors affecting the performance of PCR, such as sample storage, sample handling (58) and sample contamination, are also applicable to culture tests. In addition, bacterial isolation depends on the sample site, sample numbers, volume of inoculum, culture technique and expertise. Properly conducted bacterial isolation and biotyping can be more sensitive than serology in identifying infected cattle in herds, however, the amount of effort required and the expense incurred makes this test
impractical for use in large scale surveillance programmes (Fig. 7). Nevertheless, bacteriological isolation is still the only absolute method for establishing infection status. Bacteriological isolation from a single animal is sufficient evidence to establish the infection status of a herd and is considered the Gold standard test, and therefore the preferred method, for defining sera to be used for estimates of sensitivity. This test is particularly useful in dairy herds where a sample of milk may be collected via non-invasive procedures and tested for evidence of brucellosis. Hamdy and Amin (17) reported that the sensitivity of PCR could be affected by the deoxyribonucleic acid (DNA) extraction procedure. The sensitivity of the PCR was generally considered to be higher than bacterial isolation mainly due to the ability of the test to detect Brucella DNA in both living and dead organisms. More studies comparing the PCR and bacterial isolation would be desirable since the data presented to date suggests that the PCR alone is of limited value as a diagnostic screening test because it requires specialised skills not usually found in routine diagnostic laboratories; because the samples can take a long time to process (between 24 and 48 h) and quality control is still difficult to implement for PCR. Combining the PCR with a primary binding test may be an effective way to improve diagnostic accuracy (58).

The mean PI was derived from individual performance indices calculated from the sensitivity and specificity values from each study. The %CVs in Table II are a measure of uncertainty and variation in the mean PIs. Some of the higher variations shown in Table II (i.e. IHLT) may be directly attributed to the limited number of publications available and indirectly attributed to a variety of factors associated with the individual studies including sample sizes, experimental design and classification criteria (biases).

In summary, while test procedures vary considerably in methodology, reagents and classification criteria, their sensitivity and specificity data are a common basis for comparison. Using the PI simplifies this comparison and, along with the associated specificity and sensitivity values, can be a useful statistic for deciding whether to run tests serially or in parallel when screening for or confirming Brucella infection. As previously defined (82), a new test may replace an existing test if it is of equal or superior diagnostic performance and is cost-effective with regard to the amount of reagent(s) required to support national programmes. The data reviewed in this paper suggest that the FPA has equal or greater diagnostic accuracy than the other primary binding assays (i.e. IELISA and CELISA) and the BPAT. In addition, the FPA is less costly (i.e. one reagent), easier to perform (i.e. simple and rapid) and adaptable to automation and enhanced QA, making it suitable for eradication programmes, national serological surveys, and laboratories with large diagnostic workloads. Unlike the other tests, it can be used under field conditions to detect antibody to Brucella spp. in whole blood or milk (47). Moreover, using the FPA for the detection of antibody to Brucella spp. in milk bulk tanks (13) would further reduce the laboratory and programme costs relative to the use of the FPA for antibody detection in serum and blood as well as to other primary binding tests and conventional tests. Since only a limited number of laboratories have published sensitivity and specificity data on the FPA test, further studies of the FPA with defined Brucella positive and negative cattle populations, and Brucella spp. other than B. abortus, are required to corroborate the PI value calculated in this paper.

Acknowledgements
The authors would like to thank Drs Halbert, Forbes and Parker for their review of this article.
Comparaison des méthodes sérologiques de diagnostic de la brucellose bovine en termes de performances et de coûts

D. Gall & K. Nielsen

Résumé
Une revue bibliographique sur les valeurs de sensibilité et de spécificité des épreuves de dépistage d’une exposition à Brucella abortus a amené les auteurs à consulter plus de 50 publications. Ils ont calculé la moyenne des valeurs de sensibilité et de spécificité de chaque épreuve, après les avoir additionnés, afin de déterminer l’indice de performance (IP) des différentes méthodes et permettre ainsi leur comparaison. Un indice de 200 correspondait à une valeur idéale. L’indice de performance le plus élevé (IP = 193,1) a été obtenu par l’épreuve d’agglutination sur lame à l’antigène tamponné, traduisant sa plus grande précision par rapport aux autres méthodes classiques, y compris par rapport à l’épreuve au rose Bengale (IP = 167,6) et à l’épreuve de fixation du complément (IP = 172,5). Les épreuves par fixation primaire comme l’épreuve de polarisation par fluorescence (IP = 196,4), l’épreuve immuno-enzymatique indirecte (IP = 189,8) et l’épreuve immuno-enzymatique de compétition (IP = 188,2) étaient généralement plus précises que les épreuves classiques, à l’exception de l’épreuve d’agglutination sur lame à l’antigène tamponné. En outre, la comparaison des coûts a révélé que les épreuves de diagnostic de la brucellose par fixation primaire étaient plus intéressantes sur le plan économique que les épreuves classiques et qu’elles présentaient, par conséquent, un meilleur rapport coût/performance.

Mots-clés

Estudio comparativo del rendimiento y el coste de las pruebas de diagnóstico serológico de la brucelosis bovina

D. Gall & K. Nielsen

Resumen
Los autores pasaron revista a la bibliografía en que se tratan los valores de sensibilidad y especificidad de ensayos utilizados para detectar la exposición a Brucella abortus, estudio que abarcó más de 50 publicaciones. Tras sumar los valores de sensibilidad y especificidad de cada prueba, calcularon las medias para obtener un índice de rendimiento (IR) y poder así comparar los distintos métodos. Una puntuación de 200 equivalía a la perfección. Atendiendo al IR, la prueba de aglutinación en placa del antígeno tamponado fue, de entre las pruebas convencionales, la que arrojó una puntuación más alta (IR = 193,1), lo que pone de manifiesto su mayor exactitud en comparación con otras pruebas de esa clase, como la de rosa de Bengala (IR = 167,6) o la de fijación del complemento (IR = 172,5). En conjunto, los ensayos de fijación primaria, en particular el de fluorescencia polarizada (IR = 196,4), el ensayo inmunoenzimático (ELISA) indirecto (IR = 189,8) y el ELISA de competición (IR = 188,2), ofrecían más exactitud que las pruebas convencionales,
exceptando la de aglutinación en placa del antígeno tamponado. Por otro lado, de la comparación entre los precios de las pruebas se desprende que las de fijación primaria pueden competir en ese terreno con las pruebas convencionales de diagnóstico de la brucelosis y que por lo tanto ofrecen una mejor relación entre coste y eficacia.

**Palabras clave**

**References**


