An Evaluation of Selected Screening Tests for Bovine Paratuberculosis

W. Bruce McNab, Alan H. Meek, J. Robert Duncan, Brian W. Brooks, Anthony A. Van Dreumel, S. Wayne Martin, Klaus H. Nielsen, Edward A. Sugden and Claude Turcotte

ABSTRACT

The objective of this study was to evaluate the performance of the lipoarabinomannan antigen enzyme-linked immunosorbent assay (LAM-ELISA), carbohydrate antigen complement fixation (CH-CFT), and protein D antigen agar gel immunodiffusion (D-AGID) tests for bovine paratuberculosis, relative to histopathology, and to culture and isolation of *Mycobacterium paratuberculosis* from tissues and feces. Samples for test evaluation were collected from four sources including blood and tissues from 400 cull cows at three abattoirs in Ontario, blood and feces from a paratuberculosis survey of cattle from 120 dairy farms in Ontario, a serum bank containing samples from cattle from Ontario and Québec, and a bank of sera from cattle from Pennsylvania and the northeastern United States. The data were analyzed using receiver operator characteristic curves, estimates of relative sensitivity and specificity, and kappa statistics of agreement between tests. The LAM-ELISA performed significantly better than both the CH-CFT and the D-AGID tests. The LAM-ELISA was better at predicting fecal shedding status than tissue infection. However, the LAM-ELISA also had limitations. When interpreted as positive or negative (+/−), at a critical optical density of 0.675, its sensitivity and specificity relative to bacteriology were 49% and 87% respectively. Although the serological tests examined in this study provided some information, they did not predict well the infection status of individual animals.

RÉSUMÉ

Cette étude avait pour but de comparer :

1) **l'essai d'enzymo-immunocaptation de l'antigène lipoarabinomannan (ELISA-LAM),**
2) **du complément de fixation de l'antigène aux hydrates de carbone (CH-CFT)**
3) **des tests d'immunodiffusion en gélose utilisant l'antigène de la protéine D (D-AGID), à l'histopathologie ainsi qu'à la culture bactérienne et à l'isolement du *Mycobacterium paratuberculosis* à partir des tissus et des matières fécales de bovins.**

Les échantillons pour ces tests provenaient de quatre sources : 1 — du sang et des tissus de 400 vaches de réforme envoyées à trois différents abattoirs de l'Ontario; 2 — du sang et des matières fécales de 120 fermes laitières ontariennes ayant préalablement fait l'objet d'une enquête; 3 — d'une banque de sérums de troupeaux ontariens et québécois et 4 — d'une autre banque de sérums de troupeaux de l'état de Pennsylvanie et des états du nord-est des États-Unis.

Les données furent analysées au moyen de courbes spéciales d'évaluation; de tests de sensibilité et de spécificité relative et comparées entre elles par des méthodes statistiques de type Kappa. L'ELISA-LAM donna de meilleurs résultats que l'un (CH-CFT) ou l'autre (D-AGID) test.

Le test ELISA-LAM s'avéra meilleur pour prédire le degré d'élimination dans les matières fécales que l'infection tissulaire.

Quoiqu'il en soit, l'ELISA-LAM présenta lui aussi certaines limites. Lorsqu'interprété comme positif ou négatif (+/−) à une densité optique critique de 0.675, sa sensibilité et sa spécificité comparées à la méthode bactériologique étaient de l'ordre de 49 et 87 %, respectivement.

Bien que les tests sérologiques testés par cette étude nous fournissent un certain degré d'information, ils ne peuvent prédire le degré d'infection des animaux sur une base individuelle. (*Traduit par Dr André Cécyre*)

INTRODUCTION

Accurate and precise estimates of the prevalence and impact of bovine paratuberculosis, and the success of Johne's disease control efforts, have been limited by the lack of a reliable cost effective screening test to detect subclinical infection with *Mycobacterium paratuberculosis* (1). To this end, several tests have been investigated.

The most commonly used test for direct detection of *M. paratuberculosis* is isolation of the organism from tissues or feces, but growth of the organism is slow, resulting in a test turn around time of three to five months (2). Isolation of the bacillus from tissues or feces is considered to have a specificity of 100% relative to infection (3). The sensitivity of isolation from tissues, relative to infection, is unknown. The sensitivity of isolation from feces, relative to infection, is poor on an individual animal basis (30% to 50%), because shedding of the bacillus (at levels detectable by

Animal Diseases Research Institute, Agriculture Canada, 3851 Fallowfield Road, Nepean, Ontario K2H 8P9 (McNab, Duncan, Brooks, Nielsen, Sugden, Turcotte), Department of Population Medicine, Ontario Veterinary College, University of Guelph, Guelph, Ontario N1G 2W1 (Meek, Martin) and Veterinary Laboratory Services, Ontario Ministry of Agriculture and Food, Guelph, Ontario (Van Dreumel).

Submitted September 19, 1990.

fecal culture) is irregular and does not occur during early stages of infection (1,3).

Various antigen preparations have been applied in different ways to detect products of the host's immune response to *M. paratuberculosis* infection. Examples of test protocols include intradermal or intravenous antigen injection (4,5), lymphocyte transformation (6), leukocyte migration inhibition (7), complement fixation (8), agar gel immunodiffusion (9), enzyme-linked immunosorbent assay (ELISA) (10,11), dot immunobinding assay (12), radioimmunoassay (13), hemmagglutination (14), indirect fluorescent antibody (15) and indirect immunoperoxidase tests (16). Although these tests appear to provide some information about infection status, most immunological tests investigated to date lack sufficient sensitivity or specificity relative to subclinical infection to be used for individual culling or replacement decisions (1,17).

Recently, a lipoarabinomannan antigen enzyme-linked immunosorbent assay (LAM-ELISA) has been described by Sugden *et al* (11,18), but its field performance has not yet been investigated. This paper addresses one of the objectives of a recent epidemiological study of paratuberculosis in Ontario dairy cattle (19,20), namely to evaluate the performance of selected tests for bovine paratuberculosis, with emphasis on the LAM-ELISA. The tests investigated included (a) the LAM-ELISA, (b) a carbohydrate antigen complement fixation test (CH-CFT), and (c) a protein D antigen agar gel immunodiffusion (D-AGID) test. Their performance was investigated relative to one another, relative to histopathology, and relative to culture and isolation of *M. paratuberculosis* from tissues and feces.

TABLE I. Summary of analyses and sample sources used to evaluate the performance of selected tests for bovine paratuberculosis

Trial[a] test	Trial test outcome	Analytic technique	Ref[b] test	Ref test outcome	Sample source
LAM-ELISA[c]	OD[d]	ROC[e]	TC[f]	+/−[g]	Ont AS[h]
LAM-ELISA	OD	ROC	TC	+/−	ADRI SB[i]
LAM-ELISA	OD	ROC	TC	+/−	Pa SB[j]
LAM-ELISA	OD	ROC	FC[k]	+/−	ADRI SB
LAM-ELISA	OD	ROC	FC	+/−	Pa SB
LAM-ELISA	+/−	Sens/Spec[l]	TC	+/−	Ont AS
LAM-ELISA	+/−	Sens/Spec	TC	+/−	ADRI SB
LAM-ELISA	+/−	Sens/Spec	TC	+/−	Pa SB
LAM-ELISA	+/−	Sens/Spec	FC	+/−	ADRI SB
LAM-ELISA	+/−	Sens/Spec	FC	+/−	Pa SB
LAM-ELISA	+/−	Sens/Spec	TC/FC	+/−	Pld All[m]
LAM-ELISA	+/−	Kappa[n]	TC	+/−	Pld TC[o]
LAM-ELISA	+/−	Kappa	FC	+/−	Pld FC[p]
LAM-ELISA	+/−	Kappa	CH-CFT[q]	+/−	Ont Frm[r]
LAM-ELISA	+/−	Kappa	D-AGID[s]	+/−	Ont Frm
CH-CFT	+/−	Sens/Spec	TC	+/−	Ont AS
CH-CFT	+/−	Kappa	TC	+/−	Ont AS
CH-CFT	+/−	Kappa	D-AGID	+/−	Ont Frm
D-AGID	+/−	Sens/Spec	TC	+/−	Ont AS
D-AGID	+/−	Kappa	TC	+/−	Ont AS
Histology	+/−	Kappa	TC	+/−	Ont AS

[a]Test under investigation
[b]Reference test
[c]Lipoarabinomannan antigen ELISA
[d]Optical density (continuous)
[e]Receiver operator characteristic curve analysis
[f]Culture and isolation from tissues
[g]Positive/negative dichotomous outcome
[h]Ontario abattoir survey
[i]ADRI serum bank
[j]Pennsylvania serum bank
[k]Culture and isolation from feces
[l]Relative sensitivity and specificity
[m]Pooled TC and FC data
[n]Kappa statistic agreement between dichotomous tests
[o]Pooled TC data
[p]Pooled FC data
[q]Carbohydrate antigen complement fixation test
[r]Ontario farm survey phase 2 (case control)
[s]Protein D antigen agar gel immunodiffusion test

MATERIALS AND METHODS

SAMPLE COLLECTION

The methods used to collect samples for the present study have been previously described in detail (19,20). Briefly, samples were acquired from four sources: (a) blood and tissues (distal ileum, and ileocecal lymph node) from 400 cull cows at three Ontario abattoirs, (b) blood and feces from cattle on 120 dairy farms that participated in an epidemiological study of paratuberculosis in Ontario, (c) an Animal Diseases Research Institute (ADRI) serum bank that consisted of samples from research and privately owned Canadian cattle, and (d) a serum bank from an abattoir survey and privately owned herds in Pennsylvania, USA (21).

TEST PROTOCOLS

Six tests for paratuberculosis were investigated as follows:

(a) Histopathology

Distal ileum and the ileocecal lymph node from 200 of the 400 cull cows sampled during the Ontario abattoir survey were fixed in formalin, sectioned and stained with hematoxylin and eosin (HE) and Ziehl-Neelsen (ZN) stains. Ten different fields of each section were examined at 100× and 400× magnification by an experienced pathologist who was kept blind to the associated bacteriological and serological test results. The HE sections were examined for evidence of epithelioid type macrophages in the lymph nodes and ileal lamina propria and submucosa. The ZN sections were examined for the presence of acid fast bacilli in the same locations. Results were interpreted as positive or negative (+/−) based on the presence or absence of acid fast bacilli.

(b) Isolation from tissues (TC)

Isolation of *M. paratuberculosis* was attempted at two separate laboratories from tissues from three separate sources. Protocols varied by sample

source. For this study, all TC results were interpreted as positive or negative (+/−). Samples collected from Ontario abattoirs were shipped by courier to the Agriculture Canada, Animal Diseases Research Institute (ADRI), Nepean, Ontario for processing. The isolation protocol has been described in detail elsewhere (19). Briefly, it involved mincing a 1 g sample of tissue, decontamination with 0.7% cetylpyridinium chloride, concentration by centrifugation for 30 min at 2250 × *g*, resuspension, inoculation of five Herrold's egg yolk agar (HEYA) slants (all containing mycobactin J and four containing various combinations of inhibitors, e.g. ketoconazole and cycloheximide), and incubation for 16 wk at 37°C. Isolates were subcultured onto HEYA with and without mycobactin J. The TC protocol used on tissues from cattle represented by the ADRI serum bank was similar to that described for that of the Ontario abattoir survey, but involved concentration by sedimentation rather than centrifugation. Culture and isolation from tissues of cattle represented by the Pennsylvania serum bank was conducted at the US Department of Agriculture, National Animal Diseases Center (NADC) laboratory. The TC protocol used on those samples has been described by Merkal *et al* (22), and involved decontamination of a 1 g sample with 0.75% cetylpyridinium chloride, concentration by sedimentation, and incubation for 16 wk on egg yolk agar or Lowenstein-Jensen slants with and without sodium pyruvate and mycobactin J.

(c) Isolation from feces (FC)

Fecal samples collected on 120 farms during the second phase of the Ontario farm survey were shipped by courier to ADRI for processing. The protocol for isolation of the organism from these fecal samples has been described in detail elsewhere (19). Briefly, a 1 g fecal sample was suspended in 40 mL of distilled water and left to stand 30 min. Five mL of supernatant were transferred to a centrifuge tube containing 35 mL of a cetylpyridinium chloride solution (CPC final concentration 1.0%), left to stand for 18 to 24 h, then centrifuged for 30 min at 2250 × *g*. The pellet was resuspended in 1.2 mL saline, and used to inoculate

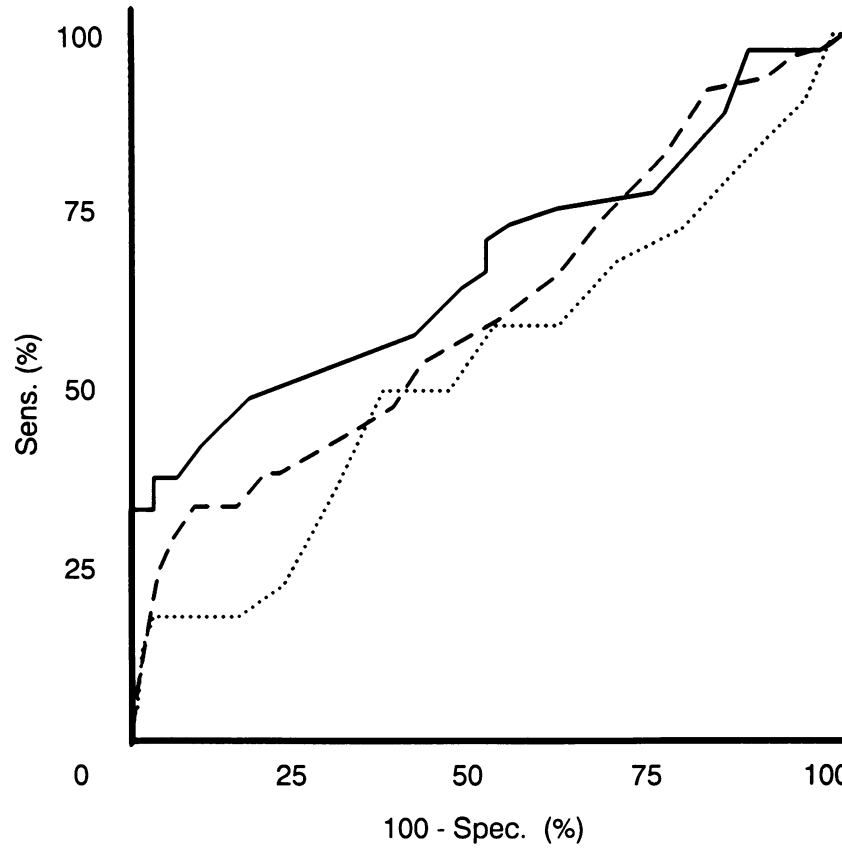

Fig. 1. Receiver operator characteristic curves for LAM-ELISA relative to culture of tissues from three sample sources, __ ADRI serum bank ROC area = 68%, — — — Pennsylvania serum bank ROC area = 61%**, ... Ontario abattoir survey ROC area = 53% (**p < 0.01).**

five HEYA slants as described above for the Ontario abattoir tissues. The NADC fecal culture results were available for animals represented by the Pennsylvania serum bank. The FC protocol used has been reported previously by Whipple and Merkal (2), and involved decontamination of a 1 g sample with cetylpyridinium chloride, concentration by overnight sedimentation and incubation for 16 wk on mycobactin J Herrold's egg yolk media. All FC results were interpreted as positive or negative.

(d) LAM-ELISA

The LAM-ELISA test protocol described by Sugden *et al* (11) was modified by Duncan and is described in detail elsewhere (19). It involves the use of a purified component antigen lipoarabinomannan (LAM) applied in an indirect ELISA format using a monoclonal mouse antibovine IgG_1 conjugate (18). Separate analyses were conducted interpreting LAM-ELISA test results both in a continuous optical density (OD) format, and in a dichotomous (+/−) format, utilizing various critical OD cut off values to dichotomize the test outcome.

(e) CH-CFT

The CH-CFT test protocol employed a relatively crude carbohydrate antigen used in a complement fixation format as described by Rice *et al* (8). It is the same test as that described by Sugden *et al* (11) who referred to it as the polysaccharide antigen complement fixation test (P-CFT). The CH-CFT was interpreted as positive or negative at a 1/10 titer.

(f) D-AGID

The D-AGID test involved a protein D antigen used in an immunodiffusion format as described by Brooks *et al* (9) and was interpreted as positive or negative.

ANALYSES

Selected serological test results were compared to bacteriological, histopathological and other serological test

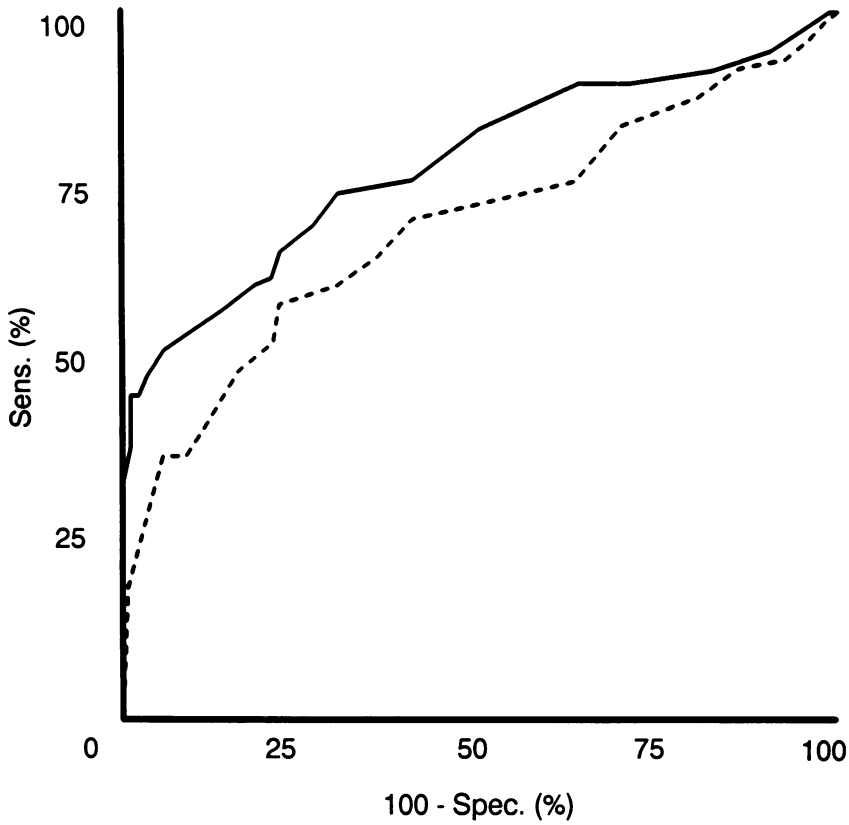

Fig. 2. Receiver operator characteristic curves for LAM-ELISA relative to culture of feces from two sample sources, ___ ADRI serum bank ROC area = 79%*, ---- Pennsylvania serum bank ROC area = 70%*** (***p < 0.001).**

results using samples from various populations and various methods of analysis (Table I).

Receiver operator characteristic (ROC) curve analysis (23), was used to evaluate the LAM-ELISA relative to TC and FC. The ROC analysis allows the evaluation of a continuous trial test outcome (e.g. LAM-ELISA OD) relative to a dichotomous reference outcome (e.g. TC or FC +/−). The area under the ROC curve provides an estimate of the probability of a randomly selected reference positive animal having a higher trial test outcome on the continuous scale than a randomly selected reference negative animal (23). The higher the probability, the better the ability of the trial test to correctly distinguish between reference positive and negative individuals. Separate analyses were conducted for each of the three major sample sources to allow comparison of test performance between different populations and relative to different reference standards. The area under each ROC curve was estimated for the LAM-ELISA test relative to TC and FC using a nonparametric method as described by DeLong *et al* (24) and the Statistical Analysis System (SAS) for personal computers (SAS Institute Inc., Cary, North Carolina, 1985).

Diagnostic sensitivity (defined as the proportion of reference test positives that are classified as positive by the trial test), and specificity (defined as the proportion of reference test negatives that are classified as negative by the trial test), were estimated for the LAM-ELISA test (relative to TC and FC for samples from the various sample sources), along the entire OD scale by dichotomizing (+/−) the LAM-ELISA result at 0.05 OD increments. Even though LAM-ELISA sensitivity and specificity varied with the critical OD cut off used to dichotomize the result, it was desirable to select one critical OD cut off to be used to estimate overall test performance. To select this cut off, all available test results were pooled from the various sample sources. Serum samples for which both TC and FC results were available, were treated as two separate observations, so that an overall estimate of performance was derived from the available data. The OD at which the sum of relative sensitivity and specificity was maximized was used as the critical OD to dichotomize LAM-ELISA into positive or negative results in subsequent analyses.

The sensitivity and specificity of the CH-CFT and the D-AGID tests were estimated relative to TC and FC.

Finally, agreement between LAM-ELISA, CH-CFT, D-AGID, TC and FC was evaluated by calculating Kappa statistics, which summarized test agreement as a proportion of that which is possible beyond chance (25).

RESULTS

Figure 1 illustrates receiver operator characteristic (ROC) curves for the LAM-ELISA relative to culture of tissues from three separate sample sources. The areas under the three ROC curves were 68%, 61% and 53% for samples from the ADRI paratuberculosis serum bank, the Pennsylvania paratuberculosis serum bank and the Ontario abattoir survey respectively. These areas correspond to the probability of a randomly selected TC positive animal having a higher optical density than a randomly selected TC negative animal. Bearing in mind that a ROC area of 50% can be attained by chance alone, the LAM-ELISA test was significantly better than chance at predicting TC status only in the ADRI and Pennsylvania serum bank samples.

Figure 2 illustrates receiver operator characteristic (ROC) curves for the LAM-ELISA relative to culture of feces (FC) from two separate sample sources. The areas under the two ROC curves were 79% and 70% for samples from the ADRI paratuberculosis serum bank, and the Pennsylvania paratuberculosis serum bank respectively. Both areas were significantly better than that expected due to chance. A comparison of Figs. 1 and 2 demonstrates the improved performance of the LAM-ELISA test relative to FC over that of its performance relative to TC.

Tables II and III demonstrate that sensitivity decreased and specificity increased as the LAM-ELISA OD cut off used to dichotomize LAM-ELISA

TABLE II. Sensitivity and specificity of LAM-ELISA relative to culture of tissues (TC)

LAM-ELISA[a] OD cut off	Ont AS[b] Sens %	Ont AS[b] Spec[e] %	ADRI SB[c] Sens %	ADRI SB[c] Spec %	Pa SB[d] Sens %	Pa SB[d] Spec %
0.00	100	0	100	0	100	0
0.05	100	2	100	0	100	0
0.10	91	5	100	0	98	2
0.15	82	14	98	3	97	7
0.20	73	22	98	10	94	11
0.25	68	31	98	13	92	19
0.30	59	39	89	17	83	25
0.35	59	48	78	27	74	34
0.40	50	55	76	40	66	40
0.45	50	64	73	47	60	48
0.50	36	71	71	50	54	59
0.55	23	78	67	50	48	63
0.60	18	84	64	53	45	68
0.65	18	89	58	60	38	79
0.70	18	93	49	83	38	82
0.75	18	95	42	90	34	85
0.80	18	96	38	93	34	91
0.85	14	98	38	97	29	94
0.90	5	99	33	97	25	96
0.95	5	99	33	100	14	98
1.00	0	100	0	100	0	100

[a]LAM-ELISA optical density cut off value used to dichotomize the test result (+/−)
[b]Sample source Ontario abattoir survey including 22 TC positive and 378 TC negative samples
[c]Sample source ADRI serum bank including 45 TC positive and 30 TC negative samples
[d]Sample source Pennsylvania serum bank including 65 TC positive and 100 TC negative samples
[e]Percent sensitivity and specificity relative to TC

TABLE III. Sensitivity and specificity of LAM-ELISA relative to fecal culture (FC)

LAM-ELISA[a] OD cut off	ADRI SB[b] Sens %	ADRI SB[b] Spec[d] %	Pa SB[c] Sens %	Pa SB[c] Spec %
0.00	100	0	100	0
0.05	100	2	100	0
0.10	91	9	99	2
0.15	88	17	96	4
0.20	84	27	93	7
0.25	84	40	92	14
0.30	80	52	88	20
0.35	75	61	84	30
0.40	73	69	76	37
0.45	67	77	73	48
0.50	63	82	71	59
0.55	59	85	65	64
0.60	59	87	61	70
0.65	58	89	59	78
0.70	55	91	53	79
0.75	53	93	49	84
0.80	52	95	41	89
0.85	52	96	37	91
0.90	42	98	37	94
0.95	33	99	28	97
1.00	0	100	0	100

[a]LAM-ELISA optical density cut off value used to dichotomize the test result (+/−)
[b]Sample source ADRI serum bank including 64 FC positive and 282 FC negative samples
[c]Sample source Pennsylvania serum bank including 75 FC positive and 123 FC negative samples
[d]Percent sensitivity and specificity relative to FC

results was increased, among samples from different sample sources and relative to different references.

Even though Figs. 1 and 2 and Tables II and III demonstrate that LAM-ELISA performance varied with: (a) the population sampled, (b) the reference test used and (c) the critical OD cutoff selected, it remained desirable to select one critical OD cutoff to be used to estimate overall test performance. Figure 3 illustrates the sensitivity and specificity and the sum of sensitivity plus specificity, of the LAM-ELISA relative to culture of tissues and/or feces, using various cutoffs along the OD scale to dichotomize the LAM-ELISA result. The sum of relative sensitivity and specificity was maximized at OD = 0.675 (Fig. 3 vertical line). This was the OD used to dichotomize LAM-ELISA results in subsequent analyses, and provided a sensitivity and specificity of 49% and 87% respectively relative to pooled bacteriological results.

Table IV illustrates the sensitivity, specificity and sum of sensitivity plus specificity for LAM-ELISA at OD = 0.675 and CH-CFT and D-AGID tests relative to TC and FC, using samples from various sources. The relative sensitivity of the serological screening tests investigated ranged from 5% to 58%, whereas, their relative specificity ranged from 77% to 99%. Since the sensitivity of FC relative to true infection status is poor, and the sensitivity of TC relative to true infection status is unknown, the specificity of serological tests relative to true infection status is likely underestimated by the estimates of specificity relative to bacteriology.

The LAM-ELISA optical density and age of cattle were positively correlated among mature cattle over two years of age ($r = 0.15$ $p < 0.001$). Also, LAM-ELISA optical densities were positively correlated ($r = 0.70$ $p < 0.001$), within animals over time, among animals that participated in both phase 1 (1987) and phase 2 (1988) of the overall epidemiological study of paratuberculosis (19,20).

Table V illustrates the kappa statistics of agreement beyond chance, observed between test results. The test pairs are listed in descending order of agreement (kappa) and the values ranged from 61% to 0%.

DISCUSSION

Sugden *et al* (11,18) described a LAM-ELISA paratuberculosis test in sheep that demonstrated potential advantages over the CH-CFT. This test was modified by Duncan (19) for use in cattle. The present study is the first to investigate the field performance of the LAM-ELISA for bovine paratuberculosis, and is significant because it investigates test performance among more populations than previous reports (5,11,16).

The ROC analyses summarized in Figs. 1 and 2 demonstrate the variability of performance of the LAM-ELISA test between reference tests and populations. Comparing the larger areas under the FC-ROC curves (Fig. 2) to the areas under the TC-ROC curves (Fig. 1) demonstrates that the LAM-ELISA was better at predicting FC status than TC status. The variability in ROC areas between populations demonstrates that test performance varies between populations. One reason for variability across populations is that the specificity of the LAM-ELISA test may be influenced by the presence of cross-reacting antibodies. For example, Sugden *et al* (11) identified cross-reactive responses to LAM-ELISA between *M. paratuberculosis* and *Corynebacterium pseudotubercu-*

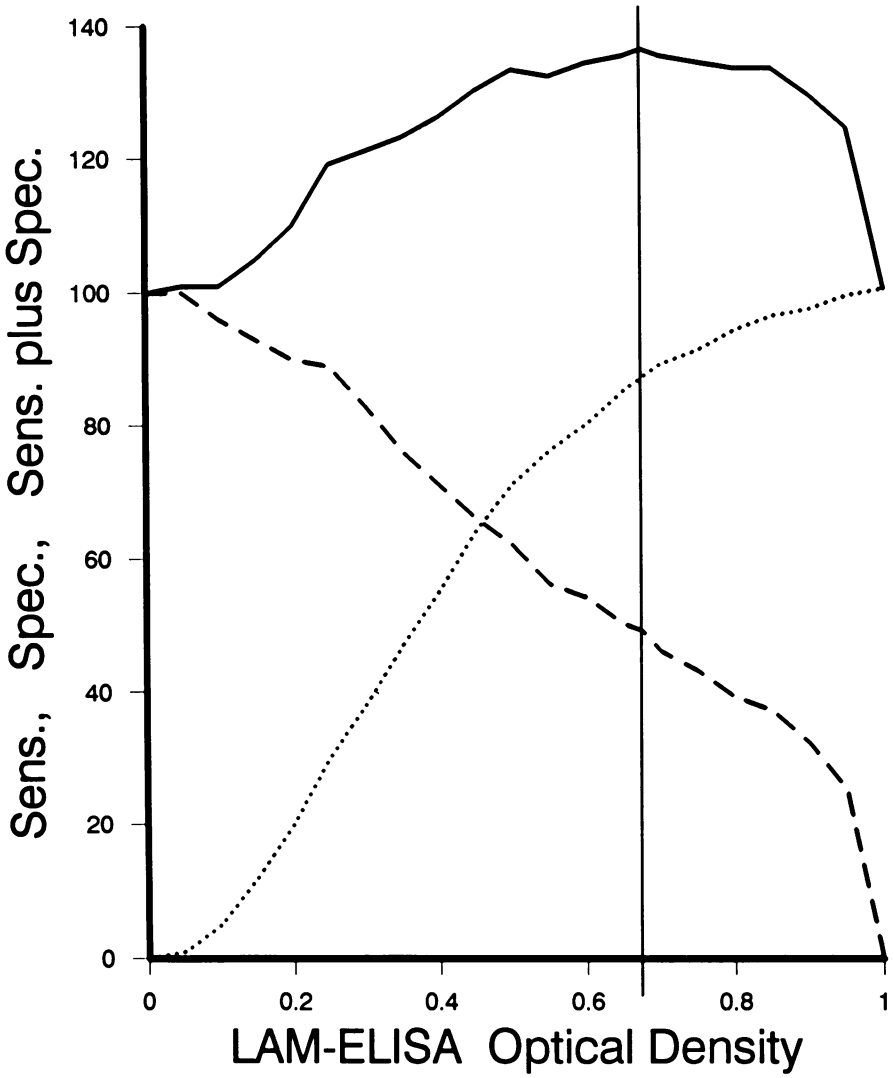

Fig. 3. LAM-ELISA sensitivity (———), specificity (.....) and sum of sensitivity plus specificity (__), relative to tissues and/or feces using pooled data from three sources. Maximum of sensitivity plus specificity (136) occurs at LAM-ELISA OD = 0.675.

losis in sheep. They noted that this was not surprising in view of the existence of arabinomannans in mycobacterial and corynebacterial species. Differences in the distributions of immune responses to true *M. paratuberculosis* infection and to cross-reacting antigens are likely the primary cause for differences in LAM-ELISA test performance between populations. Therefore, depending on the level of acceptable error, one may wish to confirm the performance of a screening test within the population of concern, before using it in a control program.

Table V summarizes agreement between various dichotomized test results using the kappa statistic. Agreement implies only that the two tests are measuring the same or closely correlated factors. Therefore, good agreement does not necessarily imply correctness of test results relative to infection. As a caution, MacLure and Willett (26) noted that the kappa statistic was originally proposed as a measure of reproducibility, and that sensitivity and specificity represent better measures of test validity than does kappa. Also, MacLure and Willett (26) challenged the use of significance testing of kappas to assess the degree of agreement. Therefore, in the present study, kappas are reported as an adjunct to sensitivity and specificity results, and should be interpreted only as indications of test performance trends.

Of the test pairs examined, agreement was best between the LAM-ELISA and the CH-CFT. This is understandable because the LAM antigen is a purified component of the carbohydrate antigen used in the CH-CFT.

Agreement between biologically independent tests, such as serology and bacteriology, provides a reasonable summary of overall test performance at determining infection (TC) and shedding (FC) status. Better agreement was achieved between LAM-ELISA and FC than between LAM-ELISA and TC. This is consistent with the ROC analyses described previously and with other studies that have noted better serological test performance in the more advanced stages of the disease (1,17).

Sugden *et al* (11) reported a threefold increase in relative sensitivity of the LAM-ELISA compared to that of the CH-CFT in sheep. However, this increase in relative sensitivity was accompanied by a decrease in relative specificity. The results of the present study concur with the trend reported by Sugden *et al* (11) in that the LAM-ELISA is better at discriminating bacteriological status than is the CH-CFT. In addition, the present results are more rigorous than those of Sugden *et al* (11) because they account for sensitivity, specificity and performance beyond chance.

Landis and Kock (27) suggested that a kappa of less than 40% demonstrates poor agreement, a kappa between 40% and 70% demonstrates fair to good agreement, but for most purposes, a kappa greater than 75% is desired. Using this rating scheme, none of the test pairs evaluated demonstrated the desired level of agreement. Only LAM-ELISA/CH-CFT and LAM-ELISA/FC test pairs demonstrated fair to good agreement. However, because the LAM-ELISA and CH-CFT tests are not biologically independent, their agreement is of little diagnostic value. Therefore, of the tests examined, only the LAM-ELISA demonstrated reasonable potential at discriminating FC status. The relatively poor agreement between histopathology and bacteriology concurs with the trend of poor agreement between pathology and clinical signs noted by previous researchers (28). The poor agreement between CH-CFT and D-AGID and between LAM-ELISA and D-AGID indicates that these tests

TABLE IV. Sensitivity and specificity estimates of dichotomized LAM-ELISA, CH-CFT, D-AGID test results relative to culture of tissues and feces using samples from three sources

Trial[a] test	Ref[b] test	No. of ref Pos	No. of ref Neg[c]	Sample source	Sens[d] %	Spec[e] %	Sens[f] + spec
LAM-ELISA[g]	TC[h]	22	378	Ont AS[i]	18	91	109
LAM-ELISA	TC	45	30	ADRI SB[j]	53	77	130
LAM-ELISA	TC	65	100	Pa SB[k]	38	81	119
LAM-ELISA	FC[l]	64	282	ADRI SB	58	89	147
LAM-ELISA	FC	75	123	Pa SB	56	79	125
CH-CFT[m]	TC	22	378	Ont AS	14	98	112
D-AGID[n]	TC	22	378	Ont AS	5	99	104

[a]Trial test under investigation
[b]Reference test
[c]Number of samples positive and negative to the reference
[d]Sensitivity (the proportion of reference test positives that are trial test positive)
[e]Specificity (the proportion of reference test negatives that are trial test negative)
[f]Sum of sensitivity plus specificity
[g]Dichotomized LAM-ELISA (+/−) at optical density 0.675
[h]Culture of tissues (pos to any, neg to all tissues)
[i]Ontario abattoir survey
[j]ADRI paratuberculosis research serum bank
[k]Pennsylvania paratuberculosis research serum bank
[l]Culture of feces (+/−)
[m]CH-CFT test (+ve complement fixation at titer > = 1/10)
[n]D-AGID (+/−)

TABLE V. Kappa statistics of test agreement beyond chance among selected pairs of dichotomized tests for bovine paratuberculosis

Test[a] 1	Test[b] 2	No. of test 2 Pos	No. of test 2 Neg[c]	Data source	Kappa[d] %
LAM-ELISA[e]	CH-CFT[f]	122	677	PLD-CFT[g]	61
LAM-ELISA	FC[h]	139	405	PLD-FC[i]	43
LAM-ELISA	TC[j]	132	508	PLD-TC[k]	30
Histology[l]	TC	22	178	ONT-AS[m]	15 !o
CH-CFT	TC	22	378	ONT-AS	14 !
D-AGID[n]	TC	22	378	ONT-AS	7 !
CH-CFT	D-AGID	4	795	PLD-AGID[p]	1 #
LAM-ELISA	D-AGID	4	796	PLD-AGID	0 #

[a]First of two tests to be compared for agreement
[b]Second of two tests to be compared for agreement
[c]Number of positive and negative results to second test (i.e. to demonstrate available base line for agreement)
[d]Kappa statistic (agreement beyond chance as a proportion of that agreement which is possible beyond chance)
[e]LAM-ELISA dichotomized (+/−) at optical density = 0.675
[f]CH-CFT test (+ve complement fixation at titer > = 1/10)
[g]Pooled CFT data involving samples from the Ontario farm survey and the Ontario abattoir survey
[h]Culture of feces (+/−)
[i]Pooled FC data involving samples from the ADRI and Pennsylvania serum banks
[j]Culture of tissues (+ve to any, −ve to all tissues)
[k]Pooled TC data involving samples from the Ontario abattoir survey, ADRI and Pennsylvania serum banks
[l]Histopathology (+/−) acid fast organisms
[m]Samples from the Ontario abattoir survey
[n]D-AGID (+/−)
[o]! and # symbols join Kappas not significantly different
[p]Pooled AGID data involving samples from the Ontario farm survey and the Ontario abattoir survey

are not likely identifying the same factor or highly correlated factors.

The ELISA results are often interpreted as being positive or negative, but information is lost when the result is dichotomized. Therefore, in situations in which the end user will not be confused by a more sophisticated method of reporting results, additional information may be conveyed by reporting results as optical densities. The end user may wish to use this additional information to rank animals or tailor test interpretation to their unique needs of sensitivity and specificity (Tables II and III). The potential advantages and disadvantages of such reporting need to be further investigated.

A protoplasmic antigen ELISA protocol that includes pre-absorption of test sera with heat-killed *M. phlei*, has been described by Yokomizo *et al* (10). When they applied the test to 156 cattle from five herds that were experiencing clinical paratuberculosis, they obtained a sensitivity and specificity of 67.9% and 99.9% respectively, relative to fecal culture, when the test was dichotomized at their recommended critical OD of 0.6 (10). Similarly, Tsai *et al* (12) reported a sensitivity and specificity of 68% and 84% respectively for a purified protoplasmic antigen dot immunobinding assay (DIA) relative to FC. In addition, the DIA has advantages over the ELISA in that it is simpler, more rapid, costs less and may be applied in the field (12).

Recently nucleic acid probes have been developed that identify *M. paratuberculosis* DNA within a sample (29,30). They offer some promise as research and diagnostic tools, especially when combined with amplification systems such as the polymerase chain reaction (29).

Nevertheless, even though the preliminary evaluation of some of these test systems offers some promise, the performance of various serological, bacteriological and DNA probe tests needs to be compared directly among the same animals, with many more animals of known infection status participating in the study than were available even in the present relatively large study.

In conclusion, the LAM-ELISA for bovine paratuberculosis performed significantly better than both the CH-CFT and D-AGID test currently used in Canada. The LAM-ELISA was significantly better at predicting fecal shedding status than tissue infection status. Identification of animals shedding the organism in their feces is of value in the control of paratuberculosis. However, even though all of these tests provided some information, they did not perform well at predicting infection status in individual animals. This finding is consistent with previous authors' conclusions concerning serological tests for paratuberculosis (1),

and likely reflects inherent limitations to the discriminatory power of serological tests for the diagnosis of paratuberculosis, given the variability in humoral immune response to infection within and between animals and populations over time. Nevertheless, given the relatively high specificity of the D-AGID, CH-CFT and the LAM-ELISA (at high OD), a positive serological test result may assist in the selection of animals to be culled from an infected herd, without waiting for bacteriological test results.

ACKNOWLEDGMENTS

We are grateful for the funding support of Agriculture Canada and the Ontario Ministry of Agriculture and Food. We thank the staff at the Animal Diseases Research Institute, Nepean Ontario, the Health of Animals Laboratory, Guelph, Ontario, and the Veterinary Services Laboratory, Guelph, Ontario, for processing the thousands of sera, fecal and tissue samples collected during this study. We thank Dr. Robert H. Whitlock and his colleagues at the University of Pennsylvania for allowing us to access their paratuberculosis serum and data banks.

REFERENCES

1. **CHIODINI RJ, van KRUININGEN HJ, MERKAL RS.** Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet 1984; 74: 218-262.
2. **WHIPPLE DL, MERKAL RS.** Procedures for the field and laboratory processing of fecal specimens for the isolation of *Mycobacterium paratuberculosis*. Proc Am Assoc Vet Lab Diagnost 1985; 28: 239-246.
3. **MERKAL RS.** Paratuberculosis: Advances in cultural, serologic, and vaccination methods. J Am Vet Med Assoc 1984; 184: 939-943.
4. **KONST H, McINTOSH CW.** Studies of Johne's disease in Canada II. Results of early intradermal testing with johnin ppd. Can J Comp Med 1958; 22: 203-214.
5. **LARSEN AB, KOPECKY KE.** Studies on the intravenous administration of johnin to diagnose Johne's disease. Am J Vet Res 1965; 26: 673-675.
6. **deLISLE GW, DUNCAN JR.** Bovine paratuberculosis III. An evaluation of a whole blood lymphocyte transformation test. Can J Comp Med 1981; 45: 304-309.
7. **BENDIXEN PH.** Application of the direct leukocyte-migration agarose test in cattle from a *Mycobacterium paratuberculosis* infected herd. Am J Vet Res 1977; 38: 2027-2028.
8. **RICE CE, KONST H, SMITH AN.** Studies of Johne's disease in Canada. III. Diagnostic complement-fixation tests. Can J Comp Med 1958; 22: 249-254.
9. **BROOKS BW, ROBERTSON RH, CORNER AH, SAMAGH BS, GARCIA MM, TURCOTTE C, DUNCAN JR.** Evaluation of the serological response of sheep in one flock to *Mycobacterium paratuberculosis* by crossed immunoelectrophoresis. Can J Vet Res 1988; 52: 199-204.
10. **YOKOMIZO Y, NISHIMORI K, KISHIMA M, YUGI H.** Evaluation of an enzyme linked immunosorbent assay (ELISA) for the diagnosis of bovine paratuberculosis. Proceedings of the Second International Colloquium on Paratuberculosis, Maison-Alfort Cedex, France, September 22-23, 1988: 206-211.
11. **SUGDEN EA, CORNER AH, SAMAGH BS, BROOKS BW, TURCOTTE C, NIELSEN KH, STEWART RB, DUNCAN JR.** Serodiagnosis of ovine paratuberculosis using lipoarabinomannan in an enzyme-linked immunosorbent assay. Am J Vet Res 1989; 50: 850-854.
12. **TSAI SJ, HUTCHINSON LJ, ZARKOWER A.** Comparison of dot immunobinding assay, enzyme-linked immunosorbent assay and immunodiffusion for serodiagnosis of paratuberculosis. Can J Vet Res 1989; 53: 405-410.
13. **WORSAAE H.** Radioimmunoassay for antibodies against *Mycobacterium paratuberculosis* using 125-I-labelled ppd. Acta Vet Scand 1978; 19: 153-155.
14. **LARSEN AB, PORTER DA, VARDAMAN TH.** A modification of the Middlebrook-Dubos hemagglutination test for use in the diagnosis of Johne's disease. Am J Vet Res 1953; 14: 362-365.
15. **BENNEDSEN J.** The specificity of circulating antibodies in experimental infections with *Mycobacterium bovis* and *Mycobacterium tuberculosis* demonstrated by immunofluorescence. Acta Pathol Microbiol Scand 1969; 76: 245-258.
16. **NGUYEN HT, BUERGELT CD.** The indirect immunoperoxidase test for the diagnosis of paratuberculosis. Am J Vet Res 1983; 44: 2173-2174.
17. **JORGENSEN JB.** Diagnosis of clinical and non-clinical paratuberculosis by immunological and bacteriological methods. Proceedings of the International Colloquium on Research in Paratuberculosis, Ames, Iowa, June 16-18, 1983: 136-141.
18. **SUGDEN EA, SAMAGH BS, BUNDLE DR, DUNCAN JR.** Lipoarabinomannan and lipid-free arabinomannan antigens of *Mycobacterium paratuberculosis*. Infect Immun 1987; 55: 762-770.
19. **McNAB WB.** An epidemiological study of paratuberculosis in Ontario dairy cattle. PhD thesis, University of Guelph, 1990.
20. **McNAB WB, MEEK AH, DUNCAN JR, MARTIN SW, van DREUMEL AA.** An epidemiological study of paratuberculosis in dairy cattle in Ontario: Study design and prevalence estimates. Can J Vet Res 1991; 55: 246-251.
21. **WHITLOCK RH, HUTCHINSON LJ, MERKAL RS, CLICKMAN LT, ROSSITER C, HARMON S, SPENCER P, FETROW J, BRUCE J, BENSON CE, DICK J.** Prevalence and economic consideration of Johne's disease in the Northeastern U.S. Proc Annu Meet US Anim Health Assoc 1985; 89: 484-490.
22. **MERKAL RS, WHIPPLE DL, SACKS JM, SNYDER GR.** Prevalence of *Mycobacterium paratuberculosis* in ileocecal lymph nodes in cattle culled in the U.S. J Am Vet Med Assoc 1987; 190: 676-680.
23. **HANLEY JA, McNEIL BJ.** The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29-36.
24. **DeLONG ER, DeLONG DM, CLARKE-PEARSON DL.** Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 1988; 44: 837-845.
25. **FLEISS JL.** Statistical Methods for Rates and Proportions. 2nd ed. Toronto: John Wiley & Sons, 1981: 212-225.
26. **MacLURE M, WILLETT WC.** Misinterpretation and misuse of the kappa statistic. Am J Epidemiol 1987; 126: 161-169.
27. **LANDIS JR, KOCH GG.** The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-174.
28. **JUBB KVF, KENNEDY PC, PALMER N.** Pathology of Domestic Animals, 3rd ed. Vol 2. Toronto: Academic Press Inc., 1985: 155-159.
29. **McFADDEN JJ, GREEN E, HERMON-TAYLOR J.** DNA probes to identify and detect *Mycobacterium paratuberculosis* in clinical and veterinary samples. Proceedings of the Second International Colloquium on Paratuberculosis, Maison-Alfort Cedex, France, September 22-23, 1988: 201-205.
30. **HURLEY SS, SPLITTER GA, WELCH RA.** Development of a diagnostic test for Johne's disease using a DNA hybridization probe. J Clin Microbiol 1989; 27: 1582-1587.