Vet. Res. 37 (2006) 553–564

DOI: 10.1051/vetres:2006018

Original article

Sensitivity and specificity of the agar-gel-immunodiffusion test, ELISA and the skin test for detection of paratuberculosis in United States Midwest sheep populations

Suelee ROBBE-AUSTERMAN[a]*, Ian A. GARDNER[b], Bruce V. THOMSEN[c], Daniel G. MORRICAL[d], Barbara M. MARTIN[c], Mitchell V. PALMER[a], Charles O. THOEN[e], Chad EWING[f]

[a] National Animal Disease Center, Agricultural Research Service, USDA, 2300 Dayton, Ames, Iowa 50010, USA
[b] Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, CA 95616, USA
[c] National Veterinary Services Laboratories, USDA, 1800 Dayton, Ames, Iowa 50010, USA
[d] Department of Animal Sciences, Iowa State University, Ames, Iowa 50011, USA
[e] Veterinary Microbiology and Preventive Medicine, Iowa State University, Ames, IA 50011, USA
[f] Groton Veterinary Clinic, 1503 N Broadway, Groton, SD 57445, USA

(Received 24 May 2005; accepted 13 January 2006)

Abstract – Our objective was to estimate the sensitivity and specificity of the agar-gel-immunodiffusion test (AGID), the ELISA, and the skin test for the detection of *Mycobacterium avium* subspecies *paratuberculosis* (*MAP*) in sheep using Bayesian methods without a gold standard. Fourteen flocks (2 465 sheep) were used. Five flocks (450 sheep) were considered *MAP* non-infected and 9 flocks (2 015 sheep) had sheep infected with *MAP*. Sheep were skin tested and blood was collected for AGID and ELISA testing. Results were analyzed using a Bayesian 3-test in 1-population model fitted in WinBUGS. The model allowed for dependence (correlation) between the two serologic tests, but these two tests were assumed to be conditionally independent of the skin test. The estimated specificity was 99.5% (95% PI of 98.9–99.9%) for the AGID; 99.3% (98.4–99.8%) for the ELISA using an optical density measured cutoff of 0.20; 99.2% (98.1–99.8%) using a cutoff of 0.15; 97.5% (95.8–98.7%) using a cutoff of 0.10; and 98.7% (97.3–99.5%) for the skin test. The estimated sensitivities were 8.3% (6.2–10.7%) for the AGID; 8.0% (6.0–10.4%), 10.6% (8.3–13.1%), and 16.3% (13.5–19.4%) for the ELISA using the cutoffs 0.20, 0.15, and 0.10 respectively; and 73.3% (62.3–85.8%) for the skin test. The skin test was specific in non-infected populations and sensitive in infected populations, although in some cases a positive skin test might represent *MAP* exposure rather than infection. The AGID and ELISA were specific but lacked sensitivity. The AGID and ELISA consistently identified two different populations of infected sheep with only moderate overlap between positive test results.

Johne's disease / sheep / ELISA / AGID / skin test

* Corresponding author: sausterm@nadc.ars.usda.gov

1. INTRODUCTION

Mycobacterium avium subspecies *paratuberculosis* (*MAP*) causes Johne's disease which is characterized by weight loss and premature culling/death in all ruminant species. Incubation period is typically years with not all animals developing clinical signs.

The two most common serologic tests available are the agar gel immunodiffusion (AGID) test and the absorbed ELISA. One ELISA (Parachek^{TM}, Biocor Animal Health, Omaha, NE, USA) is approved for use in sheep in the USA. Several studies have evaluated the sensitivity and specificity of the AGID and ELISA in sheep [5, 14, 29]. The AGID usually has slightly higher specificity (99–100%) than the ELISA (95–100%) but this depends on the gold standard used and the ELISA cutoff chosen. The sensitivity of both tests is dependent on stage of infection, with the highest sensitivity in sheep with multibacillary lesions and poor body condition and lowest in sheep in the early stages of infection [25, 29]. Overall sensitivity has been estimated to be low for both tests, from 14–62% depending on the gold standard used and the population tested. The majority of reports suggest a sensitivity between 20–30% if no consideration is given to body condition [5, 14, 24, 29, 30].

There are no published studies that have evaluated the ELISA in USA sheep populations. Most of the studies evaluating diagnostic tests in sheep cited above used populations consisting mainly of one breed and similar production systems, which is reflective of the sheep in those environments. In the USA, the sheep industry is highly diverse with many different and often unusual breeds, as well as very different production systems.

Because antibody tests lack sensitivity, especially in early preclinical stages of infection, and the cell-mediated immune (CMI) response is thought to dominate in such animals, investigations of CMI-based diagnostic tests are warranted [33]. The skin test and the interferon gamma (IFN-γ) ELISA are two such tests that may have application to the diagnosis of *MAP* infection. Several investigators have used the skin test [12, 16, 19–21, 26, 27, 36, 38] and the IFN-γ ELISA [1, 3, 13, 26, 27, 32] as research tools to characterize immune response and some have promoted these tests as potential diagnostic tools especially in the early phases of infection.

Validation of CMI-based diagnostic tests for paratuberculosis infection can be particularly difficult because a highly accurate reference test for animals in the early stages of infection is lacking. The sensitivity of bacteriologic culture of tissues and/or histology, the reference test which CMI tests have been evaluated against in the past, is affected by sample location and number of tissues taken [6, 22]. Some authors suggest up to 100 different tissue samples may be required to evaluate true infection status [39]. This could partially explain why previous investigators have concluded the skin test lacks specificity when animals in infected populations have been used for specificity determination [6, 22].

Alternative methods for evaluation of diagnostic tests must be employed when a reasonable reference test or gold standard is not readily available. Recent developments [15] have been made to techniques first suggested in the early 1980's by Hui and Walter that allow for the evaluation of more than two tests without assuming conditional independence. Estimation of the accuracy of conditionally dependent tests also can be done using Bayesian modeling, which is reviewed in detail elsewhere [2].

The objective of this study was to estimate sensitivity and specificity of the AGID, ELISA and skin test using Bayesian methods in the absence of a gold standard.

2. MATERIALS AND METHODS

2.1. Non-infected sheep population

Flocks owned by Universities or veterinarians located in the Midwest USA were

Table I. The number of sheep with positive and negative test results for various combinations of skin test, AGID and ELISA in 9 flocks with *M. avium* subsp. *paratuberculosis* (*MAP*) infected sheep and 5 non-infected flocks as well as overall test prevalence for the AGID, ELISA, and skin test. The cutoff used for the ELISA was the mean value of negative controls + 0.20.

	MAP infected flocks									*MAP* non-infected flocks				
	1P	2P	3P	4P	5P	6P	7P	8P	9P	1N	2N	3N	4N	5N
No. sheep	173	502	27	381	342	50	130	260	150	28	23	119	174	106
S+, A+, E+[a]	3	1	0	1	1	0	1	0	0	0	0	0	0	0
S+, A+, E–	6	15	0	5	8	0	2	0	0	0	0	0	0	0
S+, A–, E+	14	11	0	2	4	0	0	1	0	0	0	0	0	0
S+, A–, E–	81	183	12	99	80	8	16	18	8	0	0	3	1	1
S–, A+, E+	1	7	1	2	1	1	0	1	0	0	0	0	0	0
S–, A+, E–	3	2	0	2	4	0	0	0	0	0	0	0	0	0
S–, A–, E+	7	2	0	1	0	2	2	3	0	0	0	0	2	0
S–, A–, E–	58	281	14	269	244	39	109	237	142	28	23	116	171	105
Test prevalence %														
AGID	7.5	5.0	3.7	2.6	4.1	2.0	2.3	0.4	0.0	0.0	0.0	0.0	0.0	0.0
ELISA	14.5	4.2	3.7	1.6	1.8	6.0	2.3	1.9	0.0	0.0	0.0	0.0	1.2	0.0
Skin Test	60.1	41.8	44.4	28.3	27.1	16.0	14.6	7.3	5.3	0.0	0.0	2.5	0.6	0.9

[a] S = Skin test, A = AGID, E = ELISA.

contacted by the senior author (SRA) as they were likely to have detailed information on health events. Flocks were selected for consideration as potentially free of *MAP* infection if they had no history of Johne's disease, conducted routine necropsies on dead and debilitated animals, had not recently purchased ewes (5 years or longer) and had no purchased ewes in the flock, and followed management practices (i.e. no contact with cattle, purchased colostrum, etc.) that minimized the risk of introduction or transmission of *MAP*, if present. Owners of these flocks also agreed to submit all sheep testing positive on AGID, ELISA, or skin test in the study (or a subset of at least 5 if more than 5 tested positive) for necropsy and further testing as described in Section 2.6.

Five flocks (450 total sheep) were classified as potentially *MAP* non-infected based on the described criteria. They were located in Iowa (3), Nebraska (1) and South Dakota (1). Two (1N, 2N, see Tab. I) were small flocks with 23 and 28 animals, respectively. These animals were not regularly bred and all offspring produced stayed on the farm; one flock consisted of Jacob and Icelandic sheep, the other Finnish Landrace/Dorset cross. The other three flocks were commercial flocks, one (3N) a traditional Midwest winter lambing flock consisting of 119 Columbia/Hampshire cross ewes, another (4N) an intensively-managed accelerated (rams exposed to the ewes every 8 months) lambing flock of 174 Polypay/Dorset/Romanov cross ewes and lastly (5N) a 106 Dorset/Polypay flock that pasture lambed in May. Two of these flocks (4N, 5N) had high prevalences of caseous lymphadenitis (CLA).

2.2. Infected sheep population

MAP-infected flocks were identified by diagnostic pathologists from Iowa, South Dakota, and North Dakota based on animal or tissue samples submitted to state diagnostic laboratories. Pathologists contacted owners of these flocks and requested their participation in the study. For those producers who agreed to participate and before they were enrolled in the study, archived formalin-fixed tissues submitted previously to the diagnostic lab were confirmed infected with *MAP* by using a polymerase chain reaction (PCR) test with primers for the IS*900* sequence on the formalin-fixed tissues [23]. Flocks that had at least one *MAP* PCR-confirmed sheep fit the criteria to be included in the study as an infected flock. Detailed histories including production records, culling rate and death loss were collected from each flock to estimate the annual incidence of clinical disease. This information was then used to estimate the likely prevalence of infection for the Bayesian model, described in Section 2.7.

Nine sheep flocks, ranging in size from 27 to 502 animals, were included in the study. The annual incidence of clinical disease ranged from 0 to 11% (of total flock size). Two flocks consisting of 173 and 27 sheep were Suffolk cross club lamb (sheep raised for exhibition) flocks (1P, 3P). One was a 502 sheep whiteface crossbred (Rambouillet, Dorset, Finn, Targhee, Romanov) accelerated lambing flock (2P). Three were seedstock, two Suffolk, one with 130 sheep, the other 150 sheep (7P, 9P), and one a 50 ewe Romanov Flock (6P). Three were Rambouillet cross range flocks consisting of 381, 342, and 260 sheep (4P, 5P, 8P). Flocks were located in Kansas (2), Iowa (4), South Dakota (2), and North Dakota (1).

2.3. Skin testing and serum collection procedure

All sheep aged ≥ 12 months were tested. On day one, animals were injected intradermally with 0.10 mL of Johnin purified protein derivative (PPD) (lot 9801, NVSL/VS/APHIS/USDA, Ames, IA, USA) in the woolless area of the axillary region. This PPD was produced and evaluated in guinea pigs as previously described [34]. Ten millilitres of blood were also collected for AGID and ELISA testing. After 72 ± 4 h, the skin test site was palpated and measured for induration. Because there was minimal variation in skin thickness (mean ± SD = 4 mm ± 0.5) in the axillary region from sheep to sheep, pre-measurements were not taken. If there was a palpable swelling at the injection site when reading the skin test, the animal was considered positive and a post injection measurement was taken. Four millimetres were subtracted from the overall measurement to get the final induration measurement. All animals with palpable indurations were considered positive.

2.4. AGID testing

AGID testing was performed as previously described [30]. Briefly, agar plates were prepared with 0.7% agar (Agarose 1, biotechnology grade, Amresco, Solon, OH, USA) dissolved in a borate buffer solution (0.2% NaOH, 0.9% H3BO3, pH 8.6) containing 7.0% NaCl. Thirteen millilitres of agar were placed in 100 mm diameter petri dishes and allowed to cool. Five millimetres wells were punched in the cooled agar with one center well and 6 surrounding wells placed 3 mm from the center well. Paratuberculosis protoplasmic antigen (Allied Monitor, Fayette, MO, USA) was obtained and diluted adding 2 mL of sterile saline to 10 mg of antigen. Forty microlitres were pipetted into the center well. A positive control serum was placed in every other well alternating with the test sera. Plates were read at 24 h and 48 h. Tests were considered positive if a line of precipitation was fully formed between the test well and the antigen well and was continuous with the line formed by the positive control wells.

2.5. ELISA Testing

The ELISA (Parachek^{TM}, Biocor Animal Health, Omaha, NE, USA) was used according to manufacturer's directions. Briefly, 25 μL of test sera were added to 475 μL of serum diluent buffer for a 30 min absorption incubation at room temperature, then 100 μL of this mixture were added to the coated microtiter plate and incubated another 30 min. Plates were washed 6 times with supplied wash buffer. One hundred microlitres of diluted conjugant were added to the plate and incubated for 30 min, then the washing step was repeated. One hundred microlitres of freshly prepared enzyme substrate solution were added and plates were read with a 650 nm filter. Stopping solution was added when the OD value of the positive controls read between 0.35–0.40. The final OD values were read at 450 nm. The test was evaluated against three different cutoffs, the mean of the negative controls plus 0.1 (the USDA approved cattle cutoff), 0.15, and 0.2 (the USDA approved sheep cutoff).

2.6. Necropsy

Seven sheep from non-infected flocks and 32 sheep from infected flocks (not including the initial necropsy submitted to diagnostic laboratories that allowed flocks to be considered for inclusion in the study) were euthanized with intravenous sodium pentobarbital, 90 mg/kg. Animals were examined for gross lesions and the following tissues were taken for histologic examination: ileo-cecal valve, and associated mesenteric lymph node, distal ileum, proximal ileum, distal jejunum, and associated mesenteric lymph node, mid jejunum, and associated mesenteric lymph node, proximal jejunum, and associated lymph node, duodenum, and hepatic lymph node. Tissues were routinely processed to paraffin blocks. Sections of 5 μm thickness were stained with hematoxylin and eosin and examined by light microscopy. Adjacent sections were stained by the Ziehl-Neelsen technique to visualize acid-fast bacteria. Animals were considered infected with *MAP* when a granulomatous enteritis and or granulomatous lymphadenitis with acid-fast bacilli were present. Pathologists were blinded to the antemortem test status of the sheep.

2.7. Statistical analysis

Sensitivity and specificity of the AGID, ELISA, and skin tests were estimated by Bayesian methods using a 3-test 1-population model. The model allowed for dependence (correlation) between the two serologic tests, but these two tests were assumed to be conditionally independent of the third test (skin test) because they measured different biological responses. First, data from the five non-infected flocks were pooled into one superflock of 450 sheep. Specificity values were estimated directly from these data and beta (α,β) distributions for the specificity of each test were derived [37] for use in the 3-test 1-population model. Beta distributions provide a flexible means for modeling binomial probabilities in a Bayesian analysis because they are constrained between 0 and 1. The shape of the beta distribution is determined by the relative magnitudes of the values for α and β. The mean and variance of the beta distribution are $\alpha / (\alpha+\beta)$ and $\alpha\beta / ((\alpha+\beta)^2(\alpha+\beta+1))$, respectively. Because the estimation problem is non-identifiable without additional information, one of the authors (SRA) provided expert opinion about the likely sensitivity values for the 3 tests, following recommended guidelines [35]. The expert-elicited most likely (modal) value and corresponding beta distributions for each parameter are shown in Table II. Second, data for the nine infected flocks were pooled into a single flock of 2015 sheep. The prevalence of *MAP* infection in this hypothetical flock was uncertain but the senior author (SRA) considered the most likely value to be 30% and she was 95% sure that prevalence was < 70%. This information equated to a beta (2.13, 3.64) distribution for prevalence.

Table II. Expert-elicited values and corresponding beta (α,β) distributions for sensitivity of three diagnostic tests for ovine Johne's disease.

Test	Sensitivity (modal value)	Sensitivity (upper or lower limit)	Beta (α,β) prior distribution for sensitivity	Beta (α,β) prior distribution for specificity‡
AGID	0.2	0.4*	(4.46, 14.84)	451, 1
ELISA	0.25	0.5*	(3.88, 9.63)	449, 3
Skin test	0.7	0.2†	(2.25, 1.52)	446, 6

* Expert was 95% sure that the sensitivity was less than this value.
† Expert was 95% sure that the sensitivity was greater than this value.
Note only either the upper or lower limit is needed for estimation of beta parameters. ‡ Beta distributions for specificity for each test were constructed using the results from non-infected flocks, α = number of test-negative results + 1; β = number of test-positive results + 1.

Dependence between the AGID and ELISA was modeled with a parameterization [7] that specified uniform prior distributions for pair-wise sensitivity and specificity covariances, which quantify the magnitude of dependence between the tests. The covariances have upper and lower limits that are determined by the numeric values of the sensitivities and specificities of the tests [9]. The model was fitted in WinBUGS (MRC Biostatistics Unit, Cambridge, UK) using Gibbs sampling and code adapted from elsewhere [2, 31]. Posterior inferences were based on 50 000 iterations after discarding an initial "burn-in" of 5 000 iterations. Outputs from the model were the median estimates and 95% probability intervals (PI) for the sensitivity and specificity of each test, prevalence, and the covariances between the AGID and ELISA. Model convergence was checked by running multiple chains from different starting values [10].

Sensitivity analyses were done using non-informative beta (1,1) priors for prevalence and sensitivity of the skin test. To examine the effects of change in ELISA cutoff values, the frequencies of various combinations of test results were recalculated at cutoff values of mean + 0.10 and mean + 0.15 instead of the default value of mean + 0.20 for the ELISA and the model was rerun for these new values.

3. RESULTS

3.1. Test results

The test prevalence in each flock infected with *MAP* correlated with the level of clinical disease observed (Tab. I). Flock 1P had the highest annual incidence of clinical Johne's disease (estimated 11%) and flock 9P had the lowest incidence (0%) of clinical disease. Test prevalence increased in all three tests from 1-year-old sheep to 2-year-old sheep. After sheep reached 2 years, test prevalence did not increase (Tab. III). There seemed to be a sight decrease in sheep older than four years of age, however high prevalence flocks had fewer older sheep than the low prevalence flocks and this likely caused the apparent decrease as this pattern did not hold true when test prevalence was compared against age within flock (data not shown).

3.2. Necropsy

All animals in the non-infected flocks testing positive on the ELISA, the AGID, or the skin test for paratuberculosis (7 sheep total, Tab. I) were necropsied and no lesions suggestive of *MAP* infection were found either grossly or histologically. Of the 32 sheep necropsied in infected flocks, 10 were clinical animals euthanized while at the farm collecting samples for the study.

Table III. Skin test, AGID, and ELISA test positive results from 9 flocks (2 015 sheep) that had sheep infected with *MAP* stratified against sheep age.

Age (years)	Skin test No. positive (%)	AGID No. positive (%)	ELISA No. positive (%)	No. sheep tested
1	99 (20.7)	7 (1.5)	8 (1.8)	478
2	116 (30.9)	8 (2.1)	16 (4.3)	375
3	83 (30.4)	15 (5.5)	12 (4.4)	273
4+	282 (31.7)	38 (4.3)	34 (3.8)	889
Total	572 (28.4)	68 (3.8)	70 (3.5)	2015

Table IV. Prior and posterior median and 95% probability intervals (PI) for estimates of sensitivity and specificity of AGID, ELISA for 3 cutoffs (0.20, 0.15, 0.10) and skin tests for ovine Johne's disease.

	Sensitivity				Specificity			
Test	Prior median	Prior 95% PI	Posterior median	Posterior 95% PI	Prior median	Prior 95% PI	Posterior median	Posterior 95% PI
AGID	0.222	0.077, 0.437	0.083	0.062, 0.107	0.998	0.992, 1	0.995	0.989, 0.999
ELISA-0.20	0.276	0.089, 0.544	0.080	0.060, 0.104	0.994	0.984, 0.999	0.993	0.984, 0.998
ELISA-0.15	0.276	0.089, 0.544	0.106	0.083, 0.131	0.994	0.984, 0.999	0.992	0.981, 0.998
ELISA-0.10	0.276	0.089, 0.544	0.163	0.135, 0.194	0.994	0.984, 0.999	0.975	0.958, 0.987
Skin test	0.613	0.146, 0.955	0.733	0.623, 0.858	0.987	0.974, 0.995	0.987	0.973, 0.995

All but two had lesions consistent with paratuberculosis, one had severe disseminated caseous lymphadenitis abscesses, and the other had histologic lung lesions consistent with ovine progressive pneumonia. All but one infected flock had animals with clinical Johne's disease. In flock 9P, the only infected flock without a history of clinical Johne's disease, 22 cull sheep were necropsied over the next 3 years after the study, 3 were skin test positive and 19 skin test negative. Two of the 3 skin test positive sheep were confirmed to have paratuberculosis on necropsy (tissue PCR) and all 19 skin test negative and one skin test positive sheep did not have lesions either grossly or histologically suggestive of paratuberculosis.

3.3. Sensitivity and specificity

The Bayesian analysis indicated that the sensitivity of the skin test (median = 0.73) was substantially greater than that of either of the two serologic tests (median for both tests = 0.08) (Tab. IV). The specificities of all three tests were high (≥ 0.98) unless the ELISA cutoff was lowered to the approved cattle cutoff of mean + 0.10. The ELISA sensitivity improved (0.08 to 0.16) by lowering the cutoff, but specificity was also affected (0.99 to 0.975). There was evidence of a positive dependence in the sensitivities and specificities of the AGID and ELISA (data not shown). The posterior 95% intervals for sensitivity estimates were much narrower than the prior 95% intervals. A sensitivity analysis using non-informative priors for prevalence and sensitivity of the skin test did not change test accuracy estimates markedly (data not shown).

3.4. Agreement between AGID and ELISA

Table V compares AGID and ELISA results using different cutoffs for the ELISA

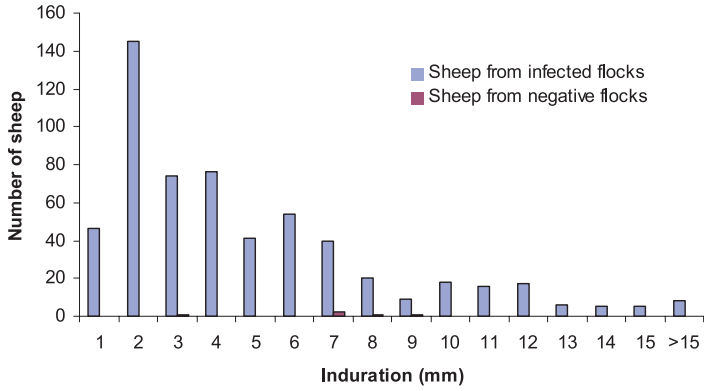

Figure 1. Induration response (mm) of the 572 of 2 015 sheep testing positive on the skin test from 9 flocks infected with *MAP* and the 5 sheep testing positive of 450 sheep from 5 non-infected flocks.

Table V. Paired comparisons of AGID and ELISA results at three different ELISA cutoffs, 0.20 (approved sheep cutoff), 0.15, 0.10 (approved cattle cutoff) in 2015 sheep from 9 USA Midwest sheep flocks. These were *MAP* infected flocks but the infection status of tested sheep was not known.

	Cutoff 0.20 + Neg. control		Cutoff 0.15 + Neg. control		Cutoff 0.10 + Neg. control		Total
	ELISA+	ELISA–	ELISA+	ELISA–	ELISA+	ELISA–	
AGID+	21	47	22	46	26	42	68
AGID–	49	1898	75	1872	146	1801	1947
Total	70	1945	97	1918	172	1843	2015

from all sheep in the flocks with Johne's disease. Decreasing the cutoff did not improve agreement between the AGID and ELISA. Only five more sheep were identified as both AGID and ELISA positive when the ELISA cutoff was lowered from 0.20 to 0.10, however the number of sheep positive on the ELISA, but not the AGID also increased from 49 to 146, an increase of 97 sheep.

3.5. Induration size of skin test responses

The skin test induration response is shown in Figure 1. Any palpable response was considered positive and measured. From previous experiments, (data not shown) indurations of ≥ 1 mm were routinely palpable. Forty-one percent (235/572) of animals that were skin test positive in infected populations had induration sizes ≤ 3 mm. These small indurations were not detected in the non-infected flocks; the false-positive induration responses in non-infected flocks ranged from 3 to 9 mm.

4. DISCUSSION

The diversity of breeds and production systems represented in this study are reflective of USA sheep industry. The location however was restricted to the Midwest, and sheep flocks in other areas of the USA should be evaluated in future studies. For

test evaluation, we used a Bayesian model that assumed that none of the tests was a gold standard. The model allowed for dependence between the two serologic tests but assumed that both tests were independent of the skin test. We had precise information on specificity of the assays using test data from the non-infected population and this allowed us to model specificity with a highly informative prior distribution. This indirectly allowed estimation of sensitivity of the three tests with reasonable precision.

A problem encountered when evaluating tests in general is the size of the populations needed [8, 11]. The probability intervals around the point estimates can be quite large, considering the number of animals tested and the weight given to the prior information. This is not such a problem if the results are near 0 or 1, such as the specificity estimates, or even the AGID and ELISA sensitivity estimates in this study; however, the probability interval around the point estimate of 0.73 for skin test sensitivity was moderately large (0.623–0.858).

The skin test in sheep was surprisingly specific (> 98%) even though no skin thickness/induration cutoff was used in the present study. This finding is contrary to previous results in cattle, which vary as to what induration size must be present in order to consider an animal positive. Kalis et al. [18] used an induration of ≥ 4 mm. If they used an induration size of ≥ 2 mm the specificity of the skin test decreased from 93 to 88%. In their study, no mention was made as to what amount of induration was routinely palpable; they measured skin thickness 72 h post injection at the injection site and 10 cm behind the injection site. Other cattle studies evaluating skin testing used 3 mm as a cutoff between positive and negative using the difference in pre- and post-testing measurements [6, 22].

Our data clearly show that many sheep in flocks with *MAP* infection tested positive at < 3 mm of induration. Test results of several of these animals were correspondingly positive with the AGID or ELISA test, making one assume that we would likely miss some infected sheep if we required an induration of ≥ 3 mm. In our study, animals in non-infected populations did not have small palpable indurations; consequently, we felt that any induration in a sheep in an infected flock is most likely to indicate *MAP* exposure or infection. It is important to note that sheep have much thinner skin than cattle, and one skin testing technique may not be ideal in both species.

Another, possibly more likely, reason for differences in induration in the non-infected populations between other studies and this one could be the antigen used. In our hands, the 9801 Johnin PPD has been more specific than any other PPD we have tried. Specificities have varied over 10% between lots, even with lots made from the same seed culture (data not shown). Our experience is consistent with previous studies [17, 18]. When and if an antigen for skin testing becomes standardized, it may be possible to assign likelihood ratios and recommendations based on induration size, so a measurement may always be desirable. However, caution must be exercised if one attempts to use a cutoff value recommended in another paper if the antigen is not the same.

The biological importance of a positive skin test result in *MAP* infected flocks is unknown. In flock 1P, 60% of the sheep tested had positive skin test results. No management changes were made and the flock continued to have a 7 to 10% annual loss attributable to clinical Johne's disease in the 3 years post testing, which was the similar to losses in the 2 years prior to testing. The future challenge will be to determine the utility of skin testing and other CMI diagnostic tests for identifying infected individuals in populations where *MAP* had been diagnosed. Obviously, culling 60% of a flock will have limited acceptance, and even in those flocks where culling is possible, the robustness and repeatability of the skin test reaction must be understood; especially knowing the probability of a subclinically infected animal testing negative.

Because the ELISA and AGID tests lack sensitivity, they usually don't significantly change negative predicative value, in other words, a negative test for a sheep in an infected flock provides little reassurance to a producer that the animal is truly negative. Consequently, these tests have been promoted as herd tests, more effective at identifying infected flocks rather than individual animals [4, 28]. In this study the AGID and ELISA were both able to identify 8 of 9 infected flocks, even with an estimated (individual animal) sensitivity of only 8%. However, infected flocks were not randomly selected without regard to previous history, so these flocks may have a higher prevalence of paratuberculosis than routinely found. With the concern of the biological relevance of a positive skin test on an individual animal basis, there is also potential for the skin test to be used as a herd test. With a much higher sensitivity, fewer animals would have to be tested. A major concern with herd level testing is specificity. Even with a specificity of 98.7%, the skin test would have a tendency to falsely identify non-infected flocks, as demonstrated by identifying 3 of the 5 non-infected flocks in this study. A small improvement in specificity would dramatically improve the potential for the skin test to be a useful herd-level test.

In the present study, the AGID and the ELISA had lower sensitivity (both at 8%) than reported in other published studies. There are several possible explanations for this finding. First, approximately 20% of animals tested in our study were between 12 and 18 months of age and many of the other studies evaluated an aged ewe population excluding ewe lambs and hoggets. Second, sensitivity in our study was evaluated using a statistical technique that was not constrained by an imperfect gold standard. Third, it is possible that the latent class for the skin test estimated in the Bayesian analysis might be a mixture of both infection and exposure whereas the AGID and ELISA latent class might only consist of infection. Since all of three tests measure immune response, it is not known if an immune response above the threshold always signifies infection, especially for the skin test and even for the ELISA at lower cutoffs. In herds or flocks with *MAP* infected animals, there is an opportunity for antigen exposure without infection but the modeling approach cannot distinguish between these two states. Because of the model structure that we used, the estimated sensitivity value of 0.73 for the skin test should be interpreted as a weighted average of the sensitivities across the 9 infected flocks rather than a flock-specific estimate.

In the United States, it is difficult to identify non-infected sheep populations, and our population of 5 flocks of 450 ewes was barely adequate to evaluate specificity across a spectrum of management practices and potential exposure to antigenically-related organisms. Our data suggest that ELISA specificity was not significantly affected if the cutoff value for sheep was decreased from 0.20 down to 0.15. This finding should be further evaluated on larger numbers of sheep before the standard cut-off is altered.

It is interesting that the ELISA and AGID identified different populations of infected animals. This was a consistent finding in the larger infected flocks which had multiple ELISA and AGID positive sheep. Several of these sheep that had discordant results were examined at necropsy and found to be infected. This finding has been reported previously [14, 29]. In this study, the agreement between these two tests was not improved by lowering the ELISA cutoff. This only slightly increased the number of AGID positive animals it identified, and dramatically increased the number of ELISA positive animals.

This study highlights the potential for skin testing to detect *MAP* infection at both the individual and herd level. However, serious obstacles need to addressed, mainly the need to understand the biological relevance of a positive skin test for proper test interpretation and the ability to consistently produce a quality antigen.

ACKNOWLEDGMENTS

The authors would like to thank Dr Janice Miller for performing the PCR testing.

Funding sources: Iowa Healthy Livestock Advisory Council, USDA/ARS CRIS Project 310-062, and the Johne's Disease Integrated Program in Research, Education, and Extension (USDA-CSREES award #2003-05165).

REFERENCES

[1] Begara-McGorum I., Wildblood L.A., Clarke C.J., Connor K.M., Stevenson K., McInnes C.J., Sharp J.M., Jones D.G., Early immunopathological events in experimental ovine paratuberculosis, Vet. Immunol. Immunopathol. 63 (1998) 265–287.

[2] Branscum A.J., Gardner I.A., Johnson W.O., Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling, Prev. Vet. Med. 68 (2005) 145–163.

[3] Burrells C., Clarke C.J., Colston A., Kay J.M., Porter J., Little D., Sharp J.M., Interferon-gamma and interleukin-2 release by lymphocytes derived from the blood, mesenteric lymph nodes and intestines of normal sheep and those affected with paratuberculosis (Johne's disease), Vet. Immunol. Immunopathol. 68 (1999) 139–148.

[4] Christensen J., Gardner I.A., Herd-level interpretation of test results for epidemiologic studies of animal diseases, Prev. Vet. Med. 45 (2000) 83–106.

[5] Clarke C.J., Patterson I.A., Armstrong K.E., Low J.C., Comparison of the absorbed ELISA and agar gel immunodiffusion test with clinicopathological findings in ovine clinical paratuberculosis, Vet. Rec. 139 (1996) 618–621.

[6] De Lisle G.W., Seguin P., Samagh B.S., Corner A.H., Duncan J.R., Bovine paratuberculosis I. A herd study using complement fixation and intradermal tests, Can. J. Comp. Med. 44 (1980) 177–182.

[7] Dendukuri N., Joseph L., Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests, Biometrics 57 (2001) 158–167.

[8] Enoe C., Georgiadis M.P., Johnson W.O., Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown, Prev. Vet. Med. 45 (2000) 61–81.

[9] Gardner I.A., Stryhn H., Lind P., Collins M.T., Conditional dependence between tests affects the diagnosis and surveillance of animal diseases, Prev. Vet. Med. 45 (2000) 107–122.

[10] Gelman A., Rubin D., Inference from iterative simulation using multiple sequences, Stat. Sci. 7 (1992) 457–511.

[11] Georgiadis M.P., Johnson W.O., Gardner I.A., Sample size determination for estimation of the accuracy of two conditionally independent tests in the absence of a gold standard, Prev. Vet. Med. 71 (2005) 1–10.

[12] Gilmour N.J., Brotherston J.G., Further studies on immunity to *Mycobacterium johnei* in sheep. Relationship between hypersensitivity and host response to infection, J. Comp. Pathol. 76 (1966) 341–349.

[13] Gwozdz J.M., Thompson K.G., Murray A., Reichel M.P., Manktelow B.W., West D.M., Comparison of three serological tests and an interferon-gamma assay for the diagnosis of paratuberculosis in experimentally infected sheep, Aust. Vet. J. 78 (2000) 779–783.

[14] Hope A.F., Kluver P.F., Jones S.L., Condron R.J., Sensitivity and specificity of two serological tests for the detection of ovine paratuberculosis, Aust. Vet. J. 78 (2000) 850–856.

[15] Hui S.L., Zhou X.H., Evaluation of diagnostic tests without gold standards, Stat. Methods Med. Res. 7 (1998) 354–370.

[16] Johnson D.W., Muscoplat C.C., Larsen A.B., Thoen C.O., Skin testing, fecal culture, and lymphocyte immunostimulation in cattle inoculated with *Mycobacterium paratuberculosis*, Am. J. Vet. Res. 38 (1977) 2023–2025.

[17] Johnson H.W., Larson A.B., Henley R.R., Groth A.H., Studies on Johnin VI. The relationship of the allergens of *Mycobacterium paratuberculosis*, *Mycobacterium tuberculosis* var. *avium, bovis*, and *hominus* and *Mycobacterium phlei*, Am. J. Vet. Res. 10 (1949) 138–141.

[18] Kalis C.H., Collins M.T., Hesselink J.W., Barkema H.W., Specificity of two tests for the early diagnosis of bovine paratuberculosis based on cell-mediated immunity: the Johnin skin test and the gamma interferon assay, Vet. Microbiol. 97 (2003) 73–86.

[19] Klawonn W., Cussler K., Drager K.G., Gyra H., Kohler H., Zimmer K., Hess R.G., The importance of allergic skin test with Johnin, antibody ELISA, cultural fecal test as well as vaccination for the sanitation of three chronically paratuberculosis-infected dairy herds

in Rhineland-Palatinate, Dtsch. Tierarztl. Wochenschr. 109 (2002) 510–516 (in German).

[20] Kluge J.P., Merkal R.S., Monlux W.S., Larsen A.B., Kopecky K.E., Ramsey F.K., Lehmann R.P., Experimental paratuberculosis in sheep after oral, intratracheal, or intravenous inoculation lesions and demonstration of etiologic agent, Am. J. Vet. Res. 29 (1968) 953–962.

[21] Kohler H., Gyra H., Zimmer K., Drager K.G., Burkert B., Lemser B., Hausleithner D., Cubler K., Klawonn W., Hess R.G., Immune reactions in cattle after immunization with a *Mycobacterium paratuberculosis* vaccine and implications for the diagnosis of *M. paratuberculosis* and *M. bovis* infections, J. Vet. Med. B Infect. Dis. Vet. Public Health 48 (2001) 185–195.

[22] Larsen A.B., Vardaman T.H., Merkal R.S., An extended study of a herd of cattle naturally infected with Johne's disease. I. The significance of the intradermic johnin test, Am. J. Vet. Res. 24 (1963) 91–93.

[23] Miller J.M., Jenny A.L., Ellingson J.L., Polymerase chain reaction identification of *Mycobacterium avium* in formalin-fixed, paraffin-embedded animal tissues, J. Vet. Diagn. Invest. 11 (1999) 436–440.

[24] Perez V., Garcia Marin J.F., Badiola J.J., Description and classification of different types of lesion associated with natural paratuberculosis infection in sheep, J. Comp. Pathol. 114 (1996) 107–122.

[25] Perez V., Tellechea J., Badiola J.J., Gutierrez M., Garcia Marin J.F., Relation between serologic response and pathologic findings in sheep with naturally acquired paratuberculosis, Am. J. Vet. Res. 58 (1997) 799–803.

[26] Reddacliff L.A., Whittington R.J., Experimental infection of weaner sheep with S strain *Mycobacterium avium* subsp. *paratuberculosis*, Vet. Microbiol. 96 (2003) 247–258.

[27] Reddacliff L.A., McGregor H., Abbott K., Whittington R.J., Field evaluation of tracer sheep for the detection of early natural infection with *Mycobacterium avium* subsp. *paratuberculosis*, Aust. Vet. J. 82 (2004) 426–433.

[28] Sergeant E.S., Whittington R.J., More S.J., Sensitivity and specificity of pooled faecal culture and serology as flock-screening tests for detection of ovine paratuberculosis in Australia, Prev. Vet. Med. 52 (2002) 199–211.

[29] Sergeant E.S., Marshall D.J., Eamens G.J., Kearns C., Whittington R.J., Evaluation of an absorbed ELISA and an agar-gel immunodiffusion test for ovine paratuberculosis in sheep in Australia, Prev. Vet. Med. 61 (2003) 235–248.

[30] Shulaw W.P., Bech-Nielsen S., Rings D.M., Getzy D.M., Woodruff T.S., Serodiagnosis of paratuberculosis in sheep by use of agar gel immunodiffusion, Am. J. Vet. Res. 54 (1993) 13–19.

[31] Spiegelhalter D., Thomas A., Best N., Gilks W., WinBUGS: Bayesian inference using Gibbs sampling, version 0.50, [on line] (1996) http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml [consulted January 27, 2004].

[32] Stabel J.R., Production of gamma-interferon by peripheral blood mononuclear cells: an important diagnostic tool for detection of subclinical paratuberculosis, J. Vet. Diagn. Invest. 8 (1996) 345–350.

[33] Stabel J.R., Transitions in immune responses to *Mycobacterium paratuberculosis*, Vet. Microbiol. 77 (2000) 465–473.

[34] Steadham E.M., Martin B.M., Thoen C.O., Production of a *Mycobacterium avium* ssp. *paratuberculosis* purified protein derivative (PPD) and evaluation of potency in guinea pigs, Biologicals 30 (2002) 93–95.

[35] Suess E.A., Gardner I.A., Johnson W.O., Hierarchical Bayesian model for prevalence inferences and determination of a country's status for an animal pathogen, Prev. Vet. Med. 55 (2002) 155–171.

[36] Vardaman T.H., Larsen A.B., Effect of intradermal johnin and tuberculin testing on the complement-fixation test for bovine tuberculosis, Am. J. Vet. Res. 27 (1966) 545–547.

[37] Vose D., Risk analysis-a quantitative guide, John Wiley and Sons, Chichester, England, 2000.

[38] Wentink G.H., Bongers J.H., Vos J.H., Zeeuwen A.A., Relationship between negative skin test with Johnin after vaccination and post mortem findings, Vet. Rec. 132 (1993) 38–39.

[39] Whitlock R.H., Rosenberger A.E., Sweeney R.W., Spencer P.A., Distribution of *M. paratuberculosis* in tissues of cattle from herds infected with Johne's disease, in: Chiodini R.J., Hines M.E., Collins M.T. (Eds.), Fifth International Colloquium on Paratuberculosis, Madison, WI, USA, 1996, pp. 168–174.