Evaluation of milk ELISA for detection of *Mycobacterium avium* subspecies *paratuberculosis* in dairy herds and association with within-herd prevalence

C. J. Lavers,^{*1} H. W. Barkema,[†] I. R. Dohoo,[‡] S. L. B. McKenna,^{*} and G. P. Keefe^{*}

^{*}Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Prince Edward Island, C1A 4P3, Canada

[†]Department of Production Animal Health, Faculty of Veterinary Medicine, University of Calgary, Calgary, Alberta, T2N 1N4, Canada

[‡]Centre for Veterinary Epidemiological Research, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Prince Edward Island, C1A 4P3, Canada

ABSTRACT

Cow-level milk ELISA results can be used to determine herd *Mycobacterium avium* ssp. *paratuberculosis* (MAP) status. Milk sample collection is minimally invasive and ELISA results can be obtained quickly and economically. The objectives were to evaluate the herd-level test characteristics of 3 commercial milk ELISA, and to determine the impact of within-herd MAP prevalence on the performance of the milk ELISA herd test. A total of 32 purposively selected herds with a median herd size of 66 milking cows were used in this 2-yr project. Fecal and milk samples were collected from all milking cows at 6-mo intervals. Fecal samples were pooled by cow age, with 5 cow samples per pool; individual fecal culture was completed on cow samples from positive pools. Herd MAP status was defined as MAP positive if, at any point during the longitudinal study, a pooled fecal culture from the herd was positive. Milk samples were analyzed using each of 3 commercial milk ELISA kits; a cow-level result from each ELISA was classified as positive following the respective manufacturer's recommended threshold for a positive result. Herd-level milk ELISA test characteristics were estimated using generalized estimating equations logistic models, which accounted for repeated measurements. Using a cutoff of 2% milk ELISA-positive cows, milk ELISA herd sensitivity relative to a herd MAP status based on all pooled fecal culture results collected during the study was as follows: ELISA A: 59% [95% confidence interval (CI): 36–78%], ELISA B: 56% (95% CI: 32–77%), and ELISA C: 63% (95% CI: 41–81%). Herd specificity for ELISA A, B, and C was 80% (95% CI: 71–88%), 96% (95% CI: 89–98%), and 92% (95% CI: 86–96%), respectively. The remainder of the analyses focused on results from ELISA B. Herd sensitivity of ELISA B increased as MAP prevalence increased. In herds with a mean MAP

prevalence $\leq 5\%$, the herd sensitivity of the milk ELISA was low, ranging from 11% when MAP prevalence was 1%, to 62% when MAP prevalence was 5%. Categorical likelihood ratios based on milk ELISA within-herd prevalence predicted that herds with milk ELISA prevalence above 0 but $< 2\%$ had a similar likelihood to be MAP positive or MAP negative, whereas herds with a milk ELISA prevalence between 2 and 4% were 3.7 times more likely to be MAP positive than MAP negative. All herds with a milk ELISA prevalence $> 4\%$ were MAP positive. Although milk ELISA B worked well to establish herd MAP status in high-prevalence herds, interpretation was unreliable in MAP-negative and low-prevalence herds.

Key words: Johne's disease, herd sensitivity and specificity, milk ELISA, prevalence

INTRODUCTION

Mycobacterium avium subspecies *paratuberculosis* (MAP) is the causative agent of Johne's disease. Although eradication of MAP is very difficult, best management practices that limit calf exposure to MAP combined with testing programs that result in management changes for, or culling of, positive cows have decreased within-herd prevalence over time (Collins et al., 2010). Although knowledge of herd MAP status is not required for best management practices to be implemented, determining that a herd is MAP positive (MAP+) provides further motivation for producers to adopt management changes to decrease both within- and between-herd transmission of the bacterium.

Testing options for determination of herd MAP status include individual or pooled fecal culture, culture of environmental samples, and ELISA on milk or serum. Performing individual cow fecal cultures on an entire herd is generally considered cost prohibitive for determining herd MAP status, as fecal culture is the most expensive cow-level test (Tiwari et al., 2006). Individual fecal culture specificity approaches 100% (Nielsen and Toft, 2008), but sensitivity is limited, with

Received June 4, 2013.

Accepted September 30, 2013.

¹Corresponding author: clavers@upepei.ca

estimates ranging from 23% (McKenna et al., 2005) to 29% (Whitlock et al., 2000) for infected and 74 to 89% for infectious animals, depending on the culture method used (Sackett et al., 1992). Pooling individual cow fecal samples by cow age, with 5 cows per pooled fecal culture (**PFC**), substantially decreases cost and is reported to be both highly sensitive and specific, relative to individual culture, for determination of herd MAP status (Kalis et al., 2004). Wells et al. (2003) reported that PFC detected 16 of the 17 herds detected with individual fecal culture, equating to a herd-level sensitivity (**HSe**), relative to individual culture, of 94%. Nine of the 17 herds had a prevalence of infection <10%, based on individual fecal culture, and the 1 herd detected positive by individual culture, but negative with PFC, had only a single light-shedding cow. In the same study, herd-level specificity (**HSp**) of 100%, relative to individual fecal culture, was reported (Wells et al., 2003). Another method of determining herd MAP status is to test individual lactating cows with a milk ELISA and compile results at the herd level. Milk ELISA herd testing is both inexpensive and rapid compared with fecal culture and is easy to implement in dairy herds that participate in DHI programs involving monthly individual cow milk collection.

The majority of MAP+ herds has a low to moderate within-herd prevalence, and although published herd-prevalence estimates vary, MAP test-negative (**MAP-**) herds are expected to represent a portion of dairy herds in most regions (Lombard, 2011). Because these low-prevalence and MAP- herds represent most of the dairy industry and HSe depends on within-herd prevalence (Dohoo et al., 2009), it is important to determine how the milk ELISA herd test performs in these herds. Previous studies have estimated an HSe for the milk ELISA, but did not take into account the within-herd prevalence levels of the herds tested (Hendrick et al., 2005; Lombard et al., 2006), whereas HSp estimates have been rare (Hendrick et al., 2005). The objectives of this study were to evaluate the herd-level test characteristics of 3 commercial milk ELISA, and to determine the impact of within-herd MAP prevalence on the performance of the milk ELISA herd test.

MATERIALS AND METHODS

Purpose for Testing and Study Terminology

The purpose for testing was to classify a dairy herd as MAP infected or not infected (Collins et al., 2006). This establishment of whether MAP infection is present in a dairy herd or not is the initial step in a herd risk management plan for MAP control within a herd (Garry, 2011). The target condition was a MAP-infected

herd, which is a herd that has at least 1 MAP-infected animal (an animal that has MAP in its tissues; Gardner et al., 2011). For the purpose of this study, the target condition was not expected to change over the period of time that herds were tested (12 to 18 mo). A case definition provides a practical definition of the target condition using a reference standard (Gardner et al., 2011). Nielsen and Toft (2008) suggested that, although not perfect, the best available method to establish cow infection status would be postmortem histopathological evaluation of up to 100 tissues per cow. This would be extremely expensive (Nielsen and Toft, 2008), and cannot be accomplished on live cows, so a case definition was used of a MAP-infected herd being a herd where MAP is shed in feces, as identified by having 1 or more positive PFC. To strengthen PFC as a herd reference standard, repeated (3 to 4) herd fecal collections were used. The case definition is referred to as a MAP+ herd, and this term, along with others defined in the manuscript, is denoted in Table 1. The case definition is similar to the use of PFC in the US Department of Agriculture (**USDA**) standards for the Voluntary Bovine Johne's Disease Control Program (USDA, 2010) in the United States, where a herd with ≥ 1 positive pooled fecal sample is considered an infected herd. The reference standard test chosen for this project was repeated PFC. It is recognized no true and perfect gold standard diagnostic test exists for MAP at either the cow or herd level (Gardner et al., 2011). With an HSe approaching that of individual fecal culture (Wells et al., 2003; Kalis et al., 2004; Collins et al., 2006), PFC was repeated over the project to minimize the misclassification of low-prevalence herds as MAP-, which is recognized as an issue when attempting to detect MAP-infected herds (Wells et al., 2002; Kalis et al., 2004).

Study Design

A total of 34 Canadian herds from the provinces of Prince Edward Island, New Brunswick, and Nova Scotia participated in this prospective study from June, 2009 to January, 2011. Details on herd selection, herd MAP-status classification, and herd demographics have been reported previously (Lavers et al., 2013). In short, the goal of the purposive herd selection was to obtain a mixture of high-prevalence, low-prevalence, and MAP-negative herds. Collection of milk and fecal samples from all lactating cows was done at 0, 6, 12, and 18 mo (4 rounds of testing). The initial round of herd testing was completed on 27 herds. Because there were too few MAP-positive herds, 7 herds were added for the remaining 3 rounds of herd testing. These herds were chosen based on the occurrence of 1 or more cows with either a positive individual fecal culture ($n = 4$) or

Table 1. Definitions and acronyms used in an evaluation of 3 commercial milk ELISA for detecting *Mycobacterium avium* ssp. *paratuberculosis* (MAP) in dairy herds

Term	Definition
PFC	A pooled fecal sample consisting of individual fecal samples from 5 cows, based on cow age
MAP+ herd	A herd was classified as MAP positive if any PFC during any test period was MAP culture positive
MAP− herd	A herd was classified as MAP negative if no PFC were MAP culture positive. It is recognized that, even with 4 rounds of whole-herd testing, some MAP+ low-prevalence herds may not have yielded any culture-positive PFC. Nevertheless, these herds were treated as MAP− for the analyses
MAP prevalence	Cow fecal samples from positive PFC were cultured individually to estimate an apparent within-herd prevalence of MAP; MAP prevalence was calculated as the number of fecal culture-positive cows divided by the total number of cows from which fecal samples were collected. In the case of a positive PFC with all negative cow results, 1 cow in the positive PFC was considered positive
Low prevalence	MAP+ herd with mean MAP prevalence $\leq 5\%$
High prevalence	MAP+ herd with mean MAP prevalence $> 5\%$
Milk ELISA positive (herd)	If the proportion of cow-level ELISA positive test results in a herd test exceeded the specified cutoff, that herd was considered milk ELISA positive for the given test period.
Milk ELISA negative (herd)	If the proportion of cow-level ELISA-positive test results in a herd test was below the specified cutoff, that herd was considered milk ELISA negative for the given test period
Milk ELISA prevalence	Within-herd prevalence of milk ELISA test results
Herd prevalence	Proportion of herds infected with MAP

clinical signs suggestive of paratuberculosis and a positive ELISA test ($n = 3$) in the previous 2 yr. Two of the 34 herds were excluded because they did not have at least 3 herd visits with collection of individual cow fecal samples. As a result, of the 32 herds included in the analyses, 25 had 4 rounds of testing and 7 herds had 3 rounds of testing. The median herd size was 66 milking cows (mean: 82; range: 28 to 220). The herd median cow age at testing ranged from 2.9 to 5.5 yr for the 32 herds. Considering all cows tested from all herds, the median age at testing was 4.0 yr (mean: 4.4 yr; range: 1.8 – 17.3 yr). One herd expanded from 190 to 220 lactating cows during the project. This herd had a PFC-positive test before the introduction of new animals. The remaining study herds were not necessarily closed, but they did not have substantial introductions of new animals during the project.

Sample Collection

The sample collection was described by Lavers et al. (2013). Briefly, individual fecal samples were collected at 6-mo intervals from all lactating cows, using a clean full-length plastic glove lubricated with sterile water. If samples could not be processed immediately, they were frozen at -20 or -80°C , if processing would occur within either 2 wk or between 2 and 6 wk after collection, respectively. Bronopol-preserved individual milk samples were collected at 6-mo intervals from all lactating cows as part of routine herd milk testing by the regional DHI organization (Valacta, Montreal, QC, Canada). The aim was to collect fecal and milk samples from each round as close as possible to each other in time, with no more than 3 mo between a herd's milk and fecal sample collection.

Laboratory Testing

Fecal Culture. The Maritime Quality Milk Laboratory (Atlantic Veterinary College, Charlottetown, PEI, Canada) processed all fecal samples. This laboratory was approved by USDA proficiency testing for this technique. Pooled fecal samples were created by pooling individual samples by cow age, with 5 cows in each PFC (Kalis et al., 2000). To estimate cow MAP status, individual cow samples from positive PFC were thawed and cultured individually. Fecal cultures were performed using ESP *para*-JEM broth (Nova Century Scientific Inc., Burlington, ON, Canada) according to the manufacturer's protocol, with the exception that samples were incubated for 49 d, rather than 42 d. Briefly, 2 g of feces from each individual cow sample was put into a weigh boat and the resulting 10 g of pooled feces was mixed thoroughly with a wooden tongue depressor; then, 2 g of the mixed sample was added to 35 mL of sterile water in a 50-mL centrifuge tube. For individual cow samples, 2 g of manure was measured and placed into 35 mL of sterile water in a 50-mL centrifuge tube. Samples were agitated vigorously for 10 min, and then placed upright for 30 min. A 5-mL sample was drawn from the feces and water mixture and placed into a centrifuge tube containing 25 mL of 0.9% 1-hexadecylpyridinium chloride in half-strength brain heart infusion broth. The tubes were incubated for 18 to 24 h at 35 to 37°C and then centrifuged at $1,500 \times g$ for 20 min at 22°C. The supernatant was discarded and the pellet was resuspended in 1 mL of brain heart infusion broth and antibiotic mixture (50 μg of amphotericin B, 100 μg of vancomycin, and 100 μg of naladixic acid). The tubes were incubated for 24 h at 35 to 37°C. On d 3, 2.5 mL of a premixed *para*-JEM reagent (1 mL of egg yolk

supplement, 1 mL of growth supplement, and 0.5 mL of antibiotic supplement) and 1 mL of the decontaminated sample were added to a *para*-JEM broth bottle. The broth bottle was agitated and then placed in the ESP Culture System II (Trek Diagnostic Systems Inc., Cleveland, OH) for incubation up to a maximum of 49 d. The culture system's computer output was examined daily for indications of positive curves. Confirmatory tests were performed as per the protocol outlined in Lavers et al. (2013). Briefly, when headspace pressure indicated growth, or at the end of the 49-d incubation period, acid-fast staining was conducted on the culture broth. Presumptive positive samples (acid-fast positive or positive growth curves) were selected for confirmation with the VetAlert Johne's Real-Time PCR kit (Tetracore Inc., Rockville, MD), which targeted the heat shock protein X (*hspX*) gene. A culture sample was considered positive if the presumptive positive sample was positive based on PCR testing.

Milk ELISA. Three indirect ELISA were used: Parachek2 *Mycobacterium paratuberculosis* test kit (ELISA A; Prionics AG, Schlieren-Zürich, Switzerland), *Mycobacterium paratuberculosis* antibody test kit (ELISA B; Idexx Laboratories Inc., Westbrook, ME), and Paratuberculosis Indirect (ELISA C; IDvet Innovative Diagnostics, Montpellier, France). All 3 commercial milk ELISA were performed by the Maritime Quality Milk laboratory, which was USDA proficiency tested for ELISA A. All 3 kits used an absorption step to remove antibodies that cross-react with *Mycobacterium phlei* (Yokomizo et al., 1985). Milk ELISA were performed following the manufacturers' instructions; a cow milk sample was defined as ELISA positive based on the criteria provided in the specific kit.

Statistical Analyses

The statistical software STATA version 11 (StataCorp LP, College Station, TX) was used for all statistical analyses. A P -value ≤ 0.05 was considered significant. For evaluation of factors influencing herd test characteristics, predictive values and likelihood ratio (LR) results for ELISA B only are presented. This ELISA had a numerically superior HSp compared with ELISA A, and is more commonly used in North America than ELISA C. Trends and models generated from all 3 ELISA were relatively similar. Unless otherwise specified, for the test under evaluation, a milk ELISA-positive herd was a herd with $\geq 2\%$ of cows milk ELISA positive within the herd. This milk ELISA cutoff was previously used by Lombard et al. (2006), and receiver operating characteristic evaluations indicated a balance between HSe and HSp with this value.

Generalized estimating equations (GEE) with an exchangeable correlation structure were used to account for repeated measures. Milk ELISA HSe was estimated from a null logistic GEE model with milk ELISA herd test results from MAP+ herds as the outcome, using the equation $HSe = e^{\beta_0} / (1 + e^{\beta_0})$, where β_0 was the intercept coefficient of a null logistic GEE model based on MAP+ herds only (Dohoo et al., 2009). Similarly, HSp was estimated from the intercept coefficient of a null logistic GEE model based on MAP- herds, using the equation $HSp = 1 - [e^{\beta_0} / (1 + e^{\beta_0})]$. In addition to the cutoff of 2% milk ELISA-positive cows, HSe and HSp were also calculated using cutoffs of 1 and 3% milk ELISA-positive cows. To evaluate predictors influencing HSe and HSp of milk ELISA B, logistic GEE models were created. The initial model for HSe, for example, was as follows:

$$\text{logit}(p) = \beta_0 + \text{SEAS} + \text{HSIZE} \\ + \text{AGE} + \text{MAP prevalence},$$

where p is the probability of a MAP+ herd having a positive milk ELISA herd test result, β_0 is the common intercept, SEAS is a categorical variable indicating the season of milk sample collection, HSIZE is the herd size (continuous), AGE is the median herd cow age (continuous), and MAP prevalence is the mean fecal culture within-herd MAP prevalence (continuous). As defined in Table 1, MAP prevalence was calculated as the number of fecal culture-positive cows divided by the total number of cows from which fecal samples were collected. For descriptive purposes, a herd's mean MAP prevalence was the mean of the herd's MAP-prevalence estimates calculated from each herd collection. Predictors with $P \leq 0.15$ in univariable analyses were entered into a multivariable logistic regression GEE model. Furthermore, $P \leq 0.05$ was considered significant for inclusion in the final model. Lowess smoothers and fractional polynomials were produced to investigate linearity and to explore power transformations of significant predictors in an effort to optimize linearity, respectively. Based on these investigations, no transformation of predictors was required.

The probability for a herd to be MAP+, given a milk ELISA-positive or milk ELISA-negative herd test (Dohoo et al., 2009), was calculated for a range of within- and between-herd prevalence scenarios for ELISA B at a 2% cutoff. The HSp estimate used in the calculation of these probabilities was generated in the null GEE model. The multivariable GEE model used to estimate HSe was influenced by the MAP prevalence; therefore, probabilities were calculated using HSe

Table 2. Frequency of positive *Mycobacterium avium* ssp. *paratuberculosis* (MAP) cultures in 5 MAP-positive herds having both positive and negative pooled fecal culture (PFC) herd test results during a 2-yr longitudinal study¹

Herd	No. of cows fecal culture positive/no. of cows tested ² (% positive)			
	Round 1	Round 2	Round 3	Round 4
1	0/66	1/64 (1.6)	1/64 (1.6)	0/58
2	1/46 (2.2)	1/48 (2.1)	1/47 (2.1)	0/47
3	Not tested	0/55	1/49 (2.0)	0/54
4	1/163 (0.6)	0/154	1/134 (0.7)	2/126 (1.6)
5	0/117	1/120 (0.8)	0/117	0/121

¹The remaining 9 MAP-positive herds had positive PFC each time the herd was sampled.

²Fecal samples from all cows tested were cultured in pools of 5; cow samples from positive pools were cultured individually.

estimates from herds with MAP prevalence of 2 and 10%, to represent low and high MAP-prevalence levels, respectively. Probabilities were calculated over a range of test population herd-prevalence estimates.

Categorical LR were determined for 4 categories of milk ELISA B prevalence (0%, >0 to 2%, >2 to 4%, and >4%). Within each milk ELISA prevalence category, the LR was calculated as the proportion of MAP+ herds within the category divided by the proportion of MAP− herds within the category.

RESULTS

Herd Prevalence

Based on PFC results collected over the 2-yr study period, 14 (44%) herds were MAP+, whereas 18 (56%) herds were MAP−. Nine of the 14 MAP+ herds had at least 1 PFC-positive result at each round of sampling, whereas 5 MAP+ herds fluctuated between PFC-negative results and at least 1 PFC-positive result (Table 2). In the first year of testing (rounds 1 and 2), 5 of 9 herd tests in these MAP+ herds had at least 1 cow

fecal sample that was culture positive. Similarly, in the second year of testing (rounds 3 and 4), 5 of 10 herd tests in these MAP+ herds had at least 1 cow fecal sample that was culture positive.

In the MAP− herds, 29, 13, and 18% of the milk ELISA herd tests had at least 1 milk ELISA-positive cow for ELISA A, B, and C, respectively (Table 3). In the MAP+ herds, 71, 67, and 78% of the milk ELISA herd tests had at least 1 cow milk ELISA positive for ELISA A, B, and C, respectively.

Within-Herd Prevalence

Mean within-herd MAP prevalence ranged from 0 to 15.6% (Figure 1). Seven of the 18 MAP− herds had at least 1 milk ELISA B-positive cow sample, whereas in the MAP+ herds, the MAP prevalence was generally higher than milk ELISA prevalence estimates (Figure 1). In the 14 MAP+ herds, mean fecal culture MAP prevalence was 6.2%, whereas the mean milk ELISA prevalence in these herds was lower, being 2.9, 3.0, and 4.0% for ELISA A, B, and C, respectively.

Table 3. Distribution of *Mycobacterium avium* ssp. *paratuberculosis* (MAP) milk ELISA test results in 14 MAP-positive and 18 MAP-negative herds, during a 2-yr longitudinal study using 3 commercial milk ELISA and 4 herd tests per herd

No. of ELISA-positive cows per herd test	No. (%) of milk ELISA herd tests					
	MAP negative (18 herds)			MAP positive (14 herds) ¹		
	ELISA A ²	ELISA B ³	ELISA C ⁴	ELISA A	ELISA B	ELISA C
0	51 (70.8)	63 (87.5)	59 (81.9)	14 (28.6)	16 (32.7)	11 (22.4)
1	11 (15.3)	6 (8.3)	9 (12.5)	10 (20.4)	9 (18.4)	9 (18.4)
2	7 (9.7)	3 (4.2)	2 (2.8)	8 (16.3)	8 (16.3)	10 (20.4)
3–4	2 (2.3)	0 (0)	2 (2.8)	8 (16.3)	5 (10.2)	6 (12.2)
5–9	0 (0)	0 (0)	0 (0)	7 (14.3)	9 (18.4)	10 (20.4)
≥10	1 (1.4)	0 (0)	0 (0)	2 (4.1)	2 (4.1)	3 (6.1)

¹Seven herds were added after the first round and were therefore only tested 3 times.

²Parachek2 *Mycobacterium paratuberculosis* test kit (Prionics AG, Schlieren-Zürich, Switzerland).

³*Mycobacterium paratuberculosis* antibody test kit (IDEXX Laboratories Inc., Westbrook, ME).

⁴Paratuberculosis Indirect (IDvet Innovative Diagnostics, Montpellier, France).

Figure 1. Mean, minimum (Min), and maximum (Max) within-herd *Mycobacterium avium* ssp. *paratuberculosis* (MAP) prevalence in 32 Maritime Canadian dairy herds using fecal culture (MAP prevalence) and milk ELISA B (*Mycobacterium paratuberculosis* antibody test kit; IDEXX Laboratories Inc., Westbrook, ME).

Milk ELISA Test Characteristics

Based on a null logistic GEE model of MAP+ herds, milk ELISA HSe ranged from 58 to 63% for the 3 milk ELISA (Table 4). Based on a null logistic GEE model of MAP– herds, HSp ranged from 80 to 96%. Results beyond this point (modeling the impact of predictors and LR) refer to ELISA B only. In the multivariable logistic GEE model evaluating predictors influencing HSe and HSp, only MAP prevalence was a significant predictor of milk HSe (Figure 2), with HSe increasing as MAP prevalence increased. The intercept coefficient was -2.68 , and the coefficient for MAP prevalence (%) was 0.63 ($P = 0.001$). As examples, milk ELISA HSe

was 19% when MAP prevalence was 2%, and increased to 97% when MAP prevalence was 10%. The probability of a herd being MAP+, given a milk ELISA-positive herd test, increased as both MAP prevalence and herd prevalence in the test population increased (Figure 3). For example, in a test population with 20% herd prevalence, the probability of being MAP+, given a milk ELISA-positive herd test was 54% in herds with a 19% HSe (the expected HSe when MAP prevalence within the herd is 2%). Conversely, in the same population of herds, the probability of being MAP+ was 85% when the HSe was 97% (the expected HSe when MAP prevalence within the herd is 10%). In contrast, the probability of a herd being MAP+, given

Table 4. Herd-level sensitivity (HSe) and specificity (HSp) for 3 commercial *Mycobacterium avium* ssp. *paratuberculosis* milk ELISA kits, using 3 cutoffs (1, 2, and 3%) based on proportion of milk ELISA-positive cows¹

ELISA	HSe [% (95% CI)]			HSp [% (95% CI)]		
	1%	2%	3%	1%	2%	3%
A ²	64 (41–83)	59 (36–78)	45 (27–65)	79 (70–86)	80 (71–88)	93 (86–97)
B ³	62 (36–83)	56 (32–77)	47 (27–68)	92 (84–96)	96 (89–98)	97 (90–99)
C ⁴	72 (48–87)	63 (41–81)	53 (32–73)	88 (79–93)	92 (86–96)	96 (89–98)

¹Results obtained from pooled fecal culture during a 2-yr longitudinal study in 32 Canadian dairy herds were used as the herd reference standard.
²Parachek2 *Mycobacterium paratuberculosis* test kit (Prionics AG, Schlieren-Zürich, Switzerland).
³*Mycobacterium paratuberculosis* antibody test kit (IDEXX Laboratories Inc., Westbrook, ME).
⁴Paratuberculosis Indirect (IDvet Innovative Diagnostics, Montpellier, France).

Figure 2. Estimated herd sensitivity of a commercial *Mycobacterium avium* ssp. *paratuberculosis* (MAP) milk ELISA (milk ELISA B; *Mycobacterium paratuberculosis* antibody test kit; IDEXX Laboratories Inc., Westbrook, ME), using 2% milk ELISA-positive cows as the cutoff, with increasing fecal culture MAP prevalence. Error bars represent the 95% CI.

a milk ELISA-negative herd test, decreased as HSe decreased. For example, in the same population of herds with a herd prevalence of 20%, at an HSe of 19% (the expected HSe when MAP prevalence within the herd is 2%), the probability of a herd being MAP+, given a milk ELISA-negative herd test, was 17%. Conversely, in the same population of herds, the probability of being MAP+ was only 1% when the HSe was 97% (the expected HSe when MAP prevalence within the herd is 10%). Likelihood ratios for 4 categories of milk ELISA prevalence ranged from 0.3 to ∞ (Table 5).

DISCUSSION

Within-herd prevalence can have a substantial effect on test characteristics at the herd level (Dohoo et al., 2009) and has apparently not been evaluated previously in a milk ELISA study. The objective of this study was to both evaluate the herd-level test characteristics of milk ELISA and determine the impact of within-herd MAP prevalence on the performance of the milk ELISA herd test. Herd-level sensitivity and HSp were estimated for 3 commercial milk ELISA, and MAP prevalence had a significant effect on HSe.

In the present study, HSe of the 3 milk ELISA ranged from 56 to 63%. Klausen et al. (2003) predicted a milk ELISA HSe of 83%, based on a single positive cow cut-

off and data collected from 6 high-prevalence and 2 negative herds. However, they suggested that sampling from a more representative set of herds was required for a less-biased estimate of test characteristics. Comparison of current study results with those of Klausen et al. (2003) is limited because of the differing within-herd prevalence estimates between study herds of the 2 projects. Hendrick et al. (2005) estimated that milk ELISA HSe ranged from 87 to 92%, depending on whether a 2- or 1-positive-cow cutoff, respectively, was used. These HSe estimates seemed higher than those in the current study. However, comparison with the results of the current study is difficult because in the previous research, estimates were used that were derived from epidemiologic formulas, rather than comparison to a reference standard. In contrast, Lombard et al. (2006) used a field application of the milk ELISA herd test compared with individual fecal culture and reported HSe values for the milk ELISA of 61% for a 2% cutoff and 83% for a 1-positive-cow cutoff. These HSe estimates seem similar to those in the current study. However, Lombard et al. (2006) did not evaluate the effect of MAP prevalence on milk ELISA HSe.

Based on modeling milk ELISA HSe over a range of within-herd MAP-prevalence values, HSe increased as within-herd MAP prevalence increased. The contrast in HSe (Figure 2), from less than 20% in herds with 2%

Table 5. Likelihood of *Mycobacterium avium* ssp. *paratuberculosis* (MAP) herd positivity for 4 categories of milk ELISA B¹ within-herd prevalence

Milk ELISA prevalence ² (%)	No. of herd tests		Likelihood ratio
	MAP-negative herds	MAP-positive herds ³	
0	63	16	0.3
>0 to ≤2	6	7	1.2
>2 to ≤4	3	11	3.7
>4	0	15	∞

¹*Mycobacterium paratuberculosis* antibody test kit (IDEXX Laboratories Inc., Westbrook, ME).

²Herd milk ELISA prevalence measurements were repeated every 6 mo during a 2-yr study period.

³A herd was considered MAP positive if at least 1 pooled fecal culture was positive during the study period.

MAP prevalence to more than 95% in herds with 10% MAP prevalence indicates that this is a critical factor. Based on the low HSe in low-prevalence herds, we inferred that whereas the milk ELISA herd test may be useful for a control program in a high-prevalence herd, as suggested by consensus recommendations (Collins et al., 2006), HSe is too low to detect a low-prevalence herd. Wells et al. (2002) considered the effect of within-herd prevalence in their serum ELISA HSe estimate, using a single-positive-cow cutoff, and also noted an

increase in HSe from 91 to 100% across 3 categories of increasing within-herd fecal culture prevalence. Although the trend is the same, the HSe estimates were higher than those from the current study. In the Wells et al. (2002) study, 41% of the herds had a within-herd prevalence of ≥10%, based on individual fecal culture, whereas in the current study, 21% of herds had a MAP prevalence ≥10%. The overall lower mean HSe estimate in the current study compared with that of Wells et al. (2002), as well as milk ELISA test characteristics estimated by Hendrick et al. (2005), was attributed to inclusion of low-prevalence herds in the current study population. Lombard et al. (2006) estimated a similar overall HSe, reflecting the fact that, similar to the current study, a greater proportion of study herds were low prevalence. Previous publications have reported that a solitary, low-shedding cow in a pool may not be detected (Wells et al., 2003). If this occurred in the current study, MAP-prevalence estimates may have been an underestimation of true within-herd MAP prevalence.

In light of the potential effect of herd characteristics (within-herd prevalence and size) on the validity of test characteristic estimates, representativeness of study herds has been previously called into question (Klausen et al., 2003). To address the issue of variable HSe, herds used in the current study were chosen to represent a

Figure 3. Influence of variable herd-level sensitivity (HSe) and herd-level prevalence on the probability for a herd to be *Mycobacterium avium* ssp. *paratuberculosis* (MAP) positive [MAP(+)], given a positive milk ELISA herd test result or a negative milk ELISA herd test result, using ELISA B (*Mycobacterium paratuberculosis* antibody test kit; IDEXX Laboratories Inc., Westbrook, ME) and 2% milk ELISA-positive cows as the cutoff.

range of within-herd prevalence values (range from 0 to 15.6%). Additionally, herd size is another important aspect of representativeness. Study herds were similar in size to Canadian (CDIC, 2013) and European Union (DairyCo, 2012) norms. Herd size in the United States is highly variable. Whereas 68% of cows are housed on farms with >200 cows, 82% of the dairy herds with ≥ 30 cows in the United States have less than 200 cows (USDA, 2012). The variability in HSe (Figure 2) indicates that concern over test characteristic variability across herds was warranted and our study design allowed for generalization of results across a large proportion of herds in the North American and European dairy industries.

Previous studies have used herds with a history of clinical disease or high prevalence of disease (Wells et al., 2002; Klausen et al., 2003). In the current study, we attempted to select herds with a wider range of within-herd prevalence, although they were not randomly selected. The magnitude of this selection bias is expected to be less than it would be in studies using only high-prevalence herds. However, additional research using a larger, randomly selected, sample of herds would be useful.

In the present study, HSp of the 3 milk ELISA ranged from 80 to 96%. In other studies, HSp could not reliably be estimated, because either few or no MAP— herds were included, or the herd MAP status was unknown. Klausen et al. (2003) estimated an HSp of 100%, but had only 2 negative herds. Using calculated estimates, without a known herd MAP status based on fecal culture, Hendrick et al. (2005) estimated an HSp of 66 and 93% for a cutoff of 1 or 2 positive cows, respectively. Lombard et al. (2006) had only 3 fecal culture test-negative herds with milk ELISA data and did not estimate an HSp. Wells et al. (2002) included 8 herds that had no prior clinical or laboratory evidence of paratuberculosis; although they did not formally calculate an HSp, all 8 herds were classified as MAP+ with the serum ELISA. These studies did not have a large number of herds that were MAP— at several consecutive tests from which to calculate HSp. With 18 herds that were repeatedly PFC negative, the present study provided a strong reference group from which to estimate HSp.

Compared with a single cross-sectional herd PFC, repeated testing with PFC allowed for the case definition to more closely reflect the target condition of a MAP-infected herd, because repeated sampling maximizes identification of herds infected at low prevalence (Wells et al., 2002; Kalis et al., 2004) and is recommended for establishment of a MAP— herd status in control programs (Collins et al., 2005). In the current study, test characteristics were estimated from a well-

characterized group of 32 herds over 2 yr. Nine of these herds were consistently positive on PFC (100% of the 30 PFC herd tests collected from these 9 herds were positive), 5 were intermittently positive on PFC (low prevalence), and 18 herds were repeatedly PFC negative and were considered MAP—. These low-prevalence herds were positive on PFC for 53% of the herd PFC tests (from Table 2, 10 out of 19 herd PFC tests were positive). If only a single cross-sectional sampling was used as a case definition, these low-prevalence herds that were intermittently positive on PFC could have been misclassified as a MAP— herd, potentially resulting in a lower reported milk ELISA HSp if misclassified herds also had positive milk ELISA herd tests.

Regarding misclassification of a MAP noninfected herd, fecal culture specificity is expected to approach 100%, so a false-positive PFC, resulting in a false-positive herd diagnosis, is unlikely. Two herds that were classified MAP+ had a single positive PFC throughout the study. Although improbable, it is possible that the shedding cow was only transiently in the herd and that no other cows in the herd were MAP infected. Given the low expected probability for this scenario, the potential impact on test evaluation outcomes is expected to be minimal. The milk ELISA herd test could be evaluated against other diagnostic herd test outcomes to ensure similar results. An alternative to the current PFC-based case definition model would be individual culture of fecal samples. Although PFC has reported high HSe relative to individual cow fecal culture (Wells et al., 2003), repeated individual cow fecal cultures may have allowed for an improved case definition, but was cost prohibitive. A Bayesian analysis, which allows for analysis of test characteristics without requiring a comparative gold standard, would also have circumvented the challenge of an imperfect gold-standard herd test. However, a sample size beyond the financial scope of this project would be required to elicit reasonable estimates and credibility intervals from a Bayesian analysis.

Both herd prevalence and herd-level test characteristics can profoundly influence interpretation of milk ELISA herd test results in low-prevalence herds (Figure 3). For practical field applications, it is important to understand the probability that a herd does or does not actually have MAP based on test results. At the cow level, these probabilities are influenced by the prevalence of disease in the study population and the cow-level test characteristics (Dohoo et al., 2009). At the herd level, the population herd prevalence and the herd-level test characteristics are important (Dohoo et al., 2009). However, as evidenced by our data, the herd-level test characteristics vary based on within-herd prevalence and, therefore, to use only 1 estimate of HSe for all herds would be inaccurate.

Another way to examine probability is using LR. This technique has been successfully used to predict the odds of MAP shedding with categorical individual cow ELISA data (Collins et al., 2005). In the context of these herd results, using LR allows for estimation of the likelihood of a herd being MAP+, given milk ELISA prevalence categories (Table 5). In herds with a milk ELISA prevalence between 0 and 2%, the test result was noninformative. However, when the milk ELISA prevalence was >4%, within the limits of the available data, test results in this category were almost certainly associated with a MAP+ herd.

CONCLUSIONS

No single HSe or HSp estimate for milk ELISA can be applied to all herds. In this study, with a range of within-herd prevalence values, HSe decreased significantly with declining MAP prevalence. Conversely, whereas HSp was high overall, LR predicted that herds with less than 2% milk ELISA prevalence were likely to be false positives. Finally, in addition to within-herd prevalence, the herd prevalence in the target population influenced interpretation of milk ELISA test results at the herd level. Consistent with previous consensus expert opinion suggesting that milk ELISA testing was most appropriate for control programs in high-prevalence herds, the present results did not support the use of milk ELISA for detection of infected herds in within-herd low-prevalence populations.

ACKNOWLEDGMENTS

Project funding was provided by Maritime Quality Milk, the dairy research program of the Atlantic Veterinary College (Charlottetown, PEI, Canada). The ELISA test kits used in this study were donated by Idexx Laboratories Inc. (Westbrook, ME; ELISA B) and IDvet Innovative Diagnostics (Montpellier, France; ELISA C). The authors acknowledge the Atlantic Veterinary College Farm Services Department and Maritime Quality Milk staff for their assistance with sample collection and laboratory analysis. Particular acknowledgement is given to Natasha Robinson, Theresa Andrews, and Heather Jack. The authors also thank the DHI technicians and dairy producers involved in this project for their participation and support.

REFERENCES

- CDIC (Canadian Dairy Information Centre). 2013. Dairy in Canada at the farm 2012. Accessed Oct. 30, 2013. http://dairyinfo.gc.ca/index_e.php?s1=dff-fcil&s2=farm-ferme&s3=nb.
- Collins, M. T., V. Eggleston, and E. J. Manning. 2010. Successful control of Johne's disease in nine dairy herds: Results of a six-year field trial. *J. Dairy Sci.* 93:1638–1643.
- Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. *J. Am. Vet. Med. Assoc.* 229:1912–1919.
- Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. *Clin. Diagn. Lab. Immunol.* 12:685–692.
- DairyCo. 2012. Dairy statistics: An insider's guide 2012. Accessed Sep. 25, 2013. http://www.dairyco.org.uk/non_umbraco/download.aspx?media=12943.
- Dohoo, I. R., S. W. Martin, and H. Stryhn. 2009. Veterinary Epidemiologic Research. 2nd ed. VER Inc., Charlottetown, PEI, Canada.
- Gardner, I. A., S. S. Nielsen, R. J. Whittington, M. T. Collins, D. Bakker, B. Harris, S. Sreevatsan, J. E. Lombard, R. Sweeney, D. R. Smith, J. Gavalchin, and S. Eda. 2011. Consensus-based reporting standards for diagnostic test accuracy studies for paratuberculosis in ruminants. *Prev. Vet. Med.* 101:18–34.
- Garry, F. 2011. Control of paratuberculosis in dairy herds. *Vet. Clin. North Am. Food Anim. Pract.* 27:599–607.
- Hendrick, S., T. Duffield, K. Leslie, K. Lissemore, M. Archambault, and D. Kelton. 2005. The prevalence of milk and serum antibodies to *Mycobacterium avium* subspecies *paratuberculosis* in dairy herds in Ontario. *Can. Vet. J.* 46:1126–1129.
- Kalis, C. H., J. M. T. Collins, H. W. Barkema, and J. W. Hesselink. 2004. Certification of herds as free of *Mycobacterium paratuberculosis* infection: Actual pooled faecal results versus certification model predictions. *Prev. Vet. Med.* 65:189–204.
- Kalis, C. H., J. W. Hesselink, H. W. Barkema, and M. T. Collins. 2000. Culture of strategically pooled bovine fecal samples as a method to screen herds for paratuberculosis. *J. Vet. Diagn. Invest.* 12:547–551.
- Klausen, J., A. Huda, L. Ekeröth, and P. Ahrens. 2003. Evaluation of serum and milk ELISAs for paratuberculosis in Danish dairy cattle. *Prev. Vet. Med.* 58:171–178.
- Lavers, C. J., S. L. B. McKenna, I. R. Dohoo, H. W. Barkema, and G. P. Keefe. 2013. Evaluation of environmental fecal culture for *Mycobacterium avium* subspecies *paratuberculosis* detection in dairy herds and association with apparent within-herd prevalence. *Can. Vet. J.* 54:1053–1060.
- Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. *Vet. Clin. North Am. Food Anim. Pract.* 27:525–535.
- Lombard, J. E., T. M. Byrem, B. A. Wagner, and B. J. McCluskey. 2006. Comparison of milk and serum enzyme-linked immunosorbent assays for diagnosis of *Mycobacterium avium* subspecies *paratuberculosis* infection in dairy cattle. *J. Vet. Diagn. Invest.* 18:448–458.
- McKenna, S. L. B., G. P. Keefe, H. W. Barkema, and D. C. Sockett. 2005. Evaluation of three ELISAs for *Mycobacterium avium* ssp. *paratuberculosis* using tissue and fecal culture as comparison standards. *Vet. Microbiol.* 110:105–111.
- Nielsen, S. S., and N. Toft. 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon- γ assay and faecal culture techniques. *Vet. Microbiol.* 129:217–235.
- Socket, D. C., D. J. Carr, and M. T. Collins. 1992. Evaluation of conventional and radiometric fecal culture and a commercial DNA probe for diagnosis of *Mycobacterium paratuberculosis* infections in cattle. *Can. J. Vet. Res.* 56:148–153.
- Tiwari, A., J. A. VanLeeuwen, S. L. B. McKenna, G. P. Keefe, and H. W. Barkema. 2006. Johne's disease in Canada: Part I: Clinical symptoms, pathophysiology, diagnosis, and prevalence in dairy herds. *Can. Vet. J.* 47:874–882.
- USDA (US Department of Agriculture). 2010 Uniform program standards for the Voluntary Bovine Johne's Disease Control Program. Accessed Sep. 25, 2013. http://www.aphis.usda.gov/animal_health/animal_diseases/johnes/downloads/johnes-ups.pdf.
- USDA (US Department of Agriculture). 2012. Farms, land in farms, and livestock operations: 2011 summary. Accessed Sep. 25, 2013. <http://usda01.library.cornell.edu/usda/nass/FarmLandIn/2010s/2012/FarmLandIn-02-17-2012.pdf>.

- Wells, S. J., S. M. Godden, C. J. Lindeman, and J. E. Collins. 2003. Evaluation of bacteriologic culture of individual and pooled fecal samples for detection of *Mycobacterium paratuberculosis* in dairy cattle herds. *J. Am. Vet. Med. Assoc.* 223:1022–1025.
- Wells, S. J., R. H. Whitlock, B. A. Wagner, J. Collins, F. Garry, H. Hirst, J. Lawrence, W. J. A. Saville, and A. L. Larew Naugle. 2002. Sensitivity of test strategies used in the voluntary Johne's disease herd status program for detection of *Mycobacterium paratuberculosis* infection in dairy cattle herds. *J. Am. Vet. Med. Assoc.* 220:1053–1057.
- Whitlock, R. H., S. J. Wells, R. W. Sweeney, and J. Van Tiem. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): Sensitivity and specificity of each method. *Vet. Microbiol.* 77:387–398.
- Yokomizo, Y., H. Yugi, and R. S. Merkal. 1985. A method for avoiding false-positive reactions in an enzyme-linked immunosorbent assay (ELISA) for the diagnosis of bovine paratuberculosis. *Nippon Juigaku Zasshi* 47:111–119.