Age-Specific Characteristics of ELISA and Fecal Culture for Purpose-Specific Testing for Paratuberculosis

S. S. Nielsen¹ and N. Toft

Department of Large Animal Sciences, The Royal Veterinary and Agricultural University, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark

ABSTRACT

Paratuberculosis is a chronic infection, and animals are not equally affected by it. Therefore, diagnostic tests that are able to detect different stages of the infection are needed for objective decision making. A longitudinal study was carried out to describe the ability of 2 tests to predict 2 conditions in dairy cattle: “infection” and “infectious,” exemplifying 2 different purposes of testing. “Infection” is the term of choice for certification and eradication purposes, and “infectious” is more relevant for control purposes. In the study period of 3 yr, repeated sampling of milk ($n = 23,219$) and feces ($n = 8,832$) was performed. A total of 1,985 Danish dairy cows provided material for the study. Milk samples were analyzed for antibodies using an ELISA, and fecal samples were analyzed for mycobacteria by culture. A reference test to correctly classify cattle antemortem does not exist; thus, “infection” and “infectious” were defined by repeated testing using one test as the condition to be detected by the other test. Fecal culture responses were evaluated against antibody status, and ELISA responses were evaluated against detected bacterial shedding. The results of this study indicate that the ability of both tests to detect “infection” increases almost linearly from 2 to 5 yr of age, whereas the ability of both tests to detect “infectious” is not affected by age. Purpose-specific tests are required to appropriately interpret and use test results for management of paratuberculosis, and relevant covariates, such as age, should be included when possible.

Key words: ELISA, paratuberculosis, immune response, bacterial shedding

INTRODUCTION

Paratuberculosis in cattle is a chronic infection caused by *Mycobacterium avium* ssp. *paratuberculosis* (Map; Chiodini et al., 1984). Infection is usually as-

sumed to take place in calves, and susceptibility to infection decreases as the animal ages (Taylor, 1953; Larsen et al., 1975). Efficient control of infection involves breaking the transmission routes through changes in management practices. Procedures that protect susceptible calves from infectious cows are described as keys in controlling paratuberculosis in infected herds, whereas test-and-cull strategies alone apparently do not decrease the prevalence (Groenendaal et al., 2002). Even so, diagnostic tests are used and may be a significant aid in risk management in infected herds to detect animals that contribute most to the bacterial load in the environment and in milk fed to calves. Diagnostic tests may also be used for certification and other purposes.

The diagnostic accuracy of a test is usually described by its diagnostic sensitivity (the probability that a test is positive given that infection is present) and the diagnostic specificity (the probability that a test is negative given that infection is not present). Underlying the sensitivity and specificity measure is an implicit assumption of an infection that is either present or absent in each animal at a level of interest for particular purposes and resulting decisions. The “infection” condition must reflect the purpose of testing as defined by the decision maker. This requires an understanding of the progressive stages of infection and accuracies for various testing schemes so that appropriate actions can be initiated subsequent to a positive test result.

In general, purpose-specific objectives could be 1) certification, 2) confirmation of clinical disease, 3) detection of infected animals, 4) detection of animals that are about to become an economic burden, and 5) detection of infectious animals that are shedding large amounts of bacteria to the environment, thus representing a risk to susceptible animals. Decisions subsequent to a test-positive result could be culling, treating the test-positive animals as high-risk animals, or confirmatory testing to provide further information.

For this study, it is assumed that paratuberculosis develops in 3 stages: 1) infection, which generally affects calves; 2) infectious (i.e., shedding infective doses of Map in feces or milk), which follows infection after some variable period of time; and 3) end-stage disease

Received July 10, 2005.

Accepted September 30, 2005.

¹Corresponding author: ssn@kvl.dk

Table 1. Description of 8 herds included in the study: herd size, housing system, production level, and apparent prevalence of *Mycobacterium avium* ssp. *paratuberculosis*

Herd	Housing system	Cow yr ¹	FCM (kg/cow yr ¹)	Prevalence (%) of fecal culture positive cows at one sampling			Age distribution at first sampling in herd			Age distribution at first calving		
				Minimum	Median	Maximum	Minimum	Median	Maximum	Minimum	Median	Maximum
1	Tie stall	66.0	8,246	1.6	9.9	17.2	2.0	3.9	9.2	1.8	2.2	3.0
2	Bed stall	105.6	10,060	0	2.2	5.3	1.8	3.6	10.8	1.8	2.2	3.4
3	Tie stall	119.9	7,943	0	2.1	5.9	2.2	3.6	7.3	1.7	2.4	3.6
4	Tie stall	71.2	7,383	0	5.0	9.1	2.1	3.8	11.0	1.9	2.3	3.0
5	Bed stall	82.7	7,313	0	10.4	18.1	2.1	3.4	7.9	1.7	2.1	3.5
6	Bed stall	260.2	8,314	4.4	7.5	13.7	2.2	4.0	12.9	1.7	2.4	3.5
7	Tie stall	68.7	5,922	0	12.5	34.8	1.9	3.7	11.0	1.8	2.1	3.1
8	Bed stall	81.5	8,138	0	4.8	10.8	2.1	3.9	8.1	1.9	2.4	3.7

¹In the period Oct. 1, 1999 to Sep. 30, 2000, one cow year is equivalent to 365 cow d.

with production losses. Obviously, the strategy of the decision maker influences the optimal choice of which of these 3 stages constitutes a truly positive or infected animal. To establish freedom from disease, information at stage 1 is necessary, but to control the infection to minimize losses, stage 2 might be adequate. Hence, it is necessary to define the purpose of the testing to establish the underlying condition that the test is intended to identify; only then may the diagnostic properties of the test be evaluated.

Two commonly used diagnostic techniques are 1) culture of Map from fecal samples or fecal culture (FC), and 2) detection of antibodies using ELISA. These tests are imperfect for detection of infection, primarily because of a long, and probably variable, incubation period. However, it is uncommon to assess the diagnostic information both relative to the purpose of testing and the chronicity of infection. Furthermore, it is known that other covariates, such as age, influence the test response, especially for the ELISA. Thus, as age of the animal is often readily available, such information should be used when evaluating a diagnostic test.

Although all infected animals must be assumed to be infectious to some degree, some animals may be more infectious than others. During the cell-mediated immune responses, some control of infection is still maintained (Stabel, 2000), and infectiousness is expectedly kept relatively low. During the subsequent humoral immune responses, the infectiousness is expected to be higher than during the cell-mediated immune responses. Thus, it can be assumed that the appearance of antibodies is a predictor of infectiousness. Therefore, an antibody (AB) test, such as an indirect ELISA, would make a good choice given that it can predict highly infectious animals.

The objective of this study was to describe 2 tests, a FC test and an indirect milk ELISA, for their ability to predict 2 conditions: "infection" and "infectious," when the tests are used as screening tests adjusted for

age as a covariate. The primary focus was describing the probability of detecting the conditions (sensitivity) and secondarily to describe the probability of absence of the conditions (specificity).

MATERIALS AND METHODS

Herds, Animals, and Observations

The sample population consisted of all cows that have had at least one calf and were present in 8 Danish dairy herds at any time during the study period from January 2000 to March 2003. In all 8 herds, Map had been isolated using FC of samples. During the study period, milk samples were obtained 11 times/yr from all lactating cows in each herd through the Danish milk recording system. Cows that were not lactating did not contribute milk samples on a given sampling date. Four times per year, fecal samples were collected from all lactating and nonlactating cows in each herd.

A summary of information regarding milk production and herd structure is given in Table 1. Information on milk production, breed, and age was obtained from the Danish Cattle Database. Date of birth was missing from 10 cows that were excluded from the study. The herds consisted of 6 different breeds, including crossbred cows. Breeds represented by few animals including 11 Red Danish, 1 Finnish Ayrshire, and 1 Old Danish were excluded from the study. The distribution of the breeds of the remaining cows was 1,430 Danish Holsteins, 435 Danish Jerseys, and 120 crossbreds. These 1,985 cows contributed a total of 23,219 milk samples and 8,832 fecal samples.

The number of samples per cow varied because of the observational nature of the study; i.e., all cows present at the date of first sampling constituted the starting sample. New cows entered at first calving, and older cows left when they were sold, culled, or died. Culling could have occurred because of FC results, as those results were communicated to the farmer. However,

Table 2. Distribution of 23,219 milk samples collected from 1,985 Danish dairy cows. Observations are cross-tabulated by age of the cow from which the milk samples were obtained when the cow tested positive (+) for *Mycobacterium avium* ssp. *paratuberculosis* in fecal culture (FC). Values in parentheses are mean OD_C values (corrected optical density) in the group

Age when cows were first detected as FC ⁺	Age at milk sampling					Total
	<2 yr	2 to 3 yr	3 to 4 yr	4 to 5 yr	>5 yr	
Never	64 (-0.01)	4,976 ¹ (-0.08)	5,484 ¹ (0.02)	3,396 ¹ (0.09)	4,712 (0.08)	18,632
2 to 3 yr of age	16 (-0.01)	555 ¹ (0.32)	300 ¹ (0.67)	54 (0.90)	0	925
3 to 4 yr of age	1 (-0.11)	469 ¹ (0.00)	724 ¹ (0.31)	334 ¹ (0.49)	85 (0.44)	1,613
4 to 5 yr of age	1 (-0.08)	126 (-0.14)	338 ¹ (0.00)	390 ¹ (0.44)	173 (0.62)	1,028
5 to 6 yr of age	0	11 (-0.15)	79 (-0.07)	143 ¹ (0.13)	328 (0.37)	561
>6 yr of age	0	0	7 (-0.10)	50 (-0.06)	403 (0.29)	460
Total	82	6,137	6,932	4,367	5,701	23,219

¹Observations used in the analyses, which corresponds to the samples obtained in the 2 yr following the year that a cow was detected FC⁺. Samples obtained from cows <2 yr of age were excluded.

culling based on a positive FC was not a common practice among the farms. About 15% of the fecal samples had contaminated culture tubes, and an accurate culture result could not be obtained. Cows from which only contaminated tubes were obtained were excluded from the study and were not included in the 1,985 cows.

The distribution of milk samples per cow was as follows: minimum = 1 sample per cow, median = 16 samples per cow, and maximum = 31 samples per cow. The distribution of fecal samples per cow was as follows: minimum = 1 sample per cow, median = 4 samples per cow, and maximum = 13 samples per cow. The distribution of milk samples per year of age is summarized in Table 2, and the distribution of fecal samples per year of age is summarized in Table 3.

Diagnostic Procedures

Milk samples were tested for presence of antibodies to Map using a milk AB ELISA. All samples were tested

in duplicate, and samples differing in corrected optical density (OD_C) by >0.1 were retested. The ELISA test and its performance have previously been described (Nielsen, 2002; Nielsen and Toft, 2002; Nielsen et al., 2002b,c), with an *M. avium* antigen and absorption with *Mycobacterium phlei*, both obtained from Allied Monitor (Ames, Fayette, MO). Nonspecific reactions are likely to occur, and the diagnostic sensitivity is affected by the age of the animal. The sensitivity and specificity of the test will be affected by the chosen detection limit. As ELISA response, the OD_C was obtained by subtracting the optical density value of a negative control from the mean optical density value of a given milk sample. These OD_C values were used as the outcome in the statistical analyses.

Fecal samples were processed as follows: 1 to 2 g of feces was decontaminated with 10 mL of NaOH solution followed by centrifugation at 1,300 × g. The supernatant was discarded, and the remaining material was dis-

Table 3. Distribution of 8,832 fecal samples collected from 1,985 Danish dairy cows. Observations are cross-tabulated by the age at which fecal samples were obtained with the age of the cow when the cow tested positive (+) for antibodies (AB) to *Mycobacterium avium* ssp. *paratuberculosis* in ELISA

Age when cow first became AB ⁺	Age at fecal sampling					Total
	<2 yr	2 to 3 yr	3 to 4 yr	4 to 5 yr	>5 yr	
Never	27	1,273 ¹	1,292 ¹	744 ¹	1,060	4,396
<2 yr of age	4	0	0	0	0	4
2 to 3 yr of age	3	292 ¹	138 ¹	25	2	460
3 to 4 yr of age	1	483 ¹	613 ¹	202 ¹	26	1,325
4 to 5 yr of age	1	212	437 ¹	363 ¹	128	1,141
5 to 6 yr of age	0	52	152	217 ¹	352	773
>6 yr of age	0	1	25	87	620	733
Total	36	2,313	2,657	1,638	2,188	8,832

¹Observations used in the analyses, which corresponds to samples obtained in the 2 yr following the year that a cow was detected AB⁺. Samples obtained from cows <2 yr of age were excluded.

solved in 5% oxalic acid with 0.1% malachite green. After a further incubation step, the material was centrifuged, and the supernatant was discarded. Neomycin sulfate and amphotericin B were added to the solution and incubated. After incubation and mixing, 3 to 4 drops of solution were applied on each of 4 tubes of medium. The fecal samples were cultured on Löwenstein-Jensen medium before July 2002 and on Herrold's egg yolk medium from August 2002 to the end of the testing period. Samples collected in July and August 2002 were tested in both media, and Herrold's egg yolk medium was more sensitive. Hence, it was decided to change and use the more sensitive medium. All positive cultures were confirmed for presence of the IS900 sequence in PCR. Further descriptions of the methods and the comparison of the 2 media are given in Nielsen et al. (2004).

Data Preparation

Data control was done by comparing the cow identity of a sampled cow with the information in the Danish Cattle Database to determine whether the cow had actually been in the herd at testing.

Figure 1 is a schematic representation of the statistical analyses. "Objective 1" refers to the investigations of ELISA as the diagnostic test, and "Objective 2" refers to FC as the diagnostic tests. The FC and ELISA were evaluated against the conditions: "Map infection" and "Map infectious." Because there is no antemortem reference test that can correctly determine the true status of an animal, the following definitions were used:

For objective 1 (evaluation of the ELISA) the condition "infection" was considered present in any cow that tested FC⁺ at least once in her life and "infectious" was the age at which a cow tested FC⁺. She could be classified as "infectious" as follows: never FC⁺; FC⁺ in third year of life (2 to 3 yr of age); FC⁺ in fourth year of life; FC⁺ in fifth year of life; FC⁺ in sixth year of life; or FC⁺ in seventh year of life or later.

For objective 2 (evaluation of the FC) the condition "infection" was considered present in any cow whose moving average of 2 consecutive OD_C values was >0.3. These animals were also referred to as AB positive. The moving average at a given age was calculated by taking the average of the present OD_C value and the OD_C value on the previous test of the cow. "Infectious" was defined as the age at which a cow obtained the "infection" status (AB⁺). She could be classified as "infectious" as follows: never AB⁺; AB⁺ in third year of life; AB⁺ in fourth year of life; AB⁺ in fifth year of life; or AB⁺ in sixth year of life; or AB⁺ in seventh year of life or later.

Statistical Analyses

Cross-tabulations showing the distributions of test-positive samples (FC⁺ or AB⁺) obtained by age group and by infectious group (age group at which an animal became test-positive) contributed to the basic descriptive statistics. For descriptive purposes, mean OD_C values were calculated for the various groups.

In the following analyses, cows contributed samples in the age group in which they were found FC⁺/AB⁺ and 2 age groups later, as shown in Table 2 (for analyses of the ELISA response) and Table 3 (for the analyses of the FC).

Predictions of the AB response as a function of age were estimated by nonparametric regression of OD_C values as a function of age, using cubic-smoothing splines (Hastie and Tibshirani, 1991) using the GAM procedure in SAS (Version 8.2, SAS Inst., Inc., Cary, NC). The predictions were performed for all observations in 2 groups (FC⁺ and FC⁻) in which the underlying condition was "infection" and were divided into groups depending on which age they turned FC⁺ or when the underlying condition was "infectious." The model used for each of these groups was as follows:

$$\text{OD}_C = \beta_0 + S(\text{age}), \text{ per FC group} \quad [1]$$

where OD_C was the OD_C value from the ELISA, β_0 was the baseline value of the OD_C, and S(age) was the smoothing function of age in year for each of the FC groups.

To assess the specificity of the ELISA, an additional analysis was performed. In this analysis, the noninfected cows were defined as cows with ≥ 8 negative FC obtained over a 2-yr period. Samples from the 2-yr period were excluded. Only samples from about the first year were included.

Subsequent to analysis of the OD_C response on a continuous scale, the OD_C were dichotomized at a cut-off of 0.3, which is the recommended laboratory cut-off (where OD_C > 0.3 are positive). Data were analyzed with the following model:

$$g[P(\text{OD}_C > 0.3)] = \quad [2]$$

$$\log \frac{P(\text{OD}_C > 0.3)}{1 - P(\text{OD}_C > 0.3)} = \beta_0 + S(\text{age}), \text{ per FC group}$$

where $P(\text{OD}_C > 0.3)$ was the probability of testing positive in the ELISA at cut-off 0.3, β_0 was the baseline probability of testing positive in ELISA, and S(age) was the smoothing function of age in years in each FC group.

As with the OD_C on a continuous scale, a separate analysis was done with cows that had ≥ 8 negative FC over a 2-yr period. Samples from the 2-yr period were

Figure 1. Schematic representation of statistical analyses. Map = *Mycobacterium avium* ssp. *paratuberculosis*, FC = fecal culture, AB⁺ = antibody-positive based on milk ELISA, and Y = year Y of life.

excluded. Only samples from about the first year were included.

A similar model was used for predictions of the probability of testing positive at any time in FC. The model was

$$g[P(\text{FC}^+)] = \log \frac{P(\text{FC}^+)}{1 - P(\text{FC}^+)} = \quad [3]$$

$$\beta_0 S(\text{age}), \text{ per AB}^+ \text{ group}$$

where $P(\text{FC}^+)$ is the probability of testing positive in FC, β_0 was the baseline probability of testing positive in FC, and $S(\text{age})$ was the smoothing function of age in years in each AB group.

The predictions for FC^+ were done for all observations divided into 2 groups: AB⁺ and AB⁻ with the underlying condition being “infection” and also when divided in groups depending on the age at which they turned “infectious.”

The effect of change in FC method was assessed by an extension of Equation 3, where culture method was

included as a parametric term, which could be either the old or the new method.

RESULTS

Table 2 gives the distribution of observations used for the analyses of OD_C values, cross-tabulated by actual age group at sampling and the age group in which the animal first became FC^+ , as well as the mean OD_C value in each of the groups. Table 3 gives the distribution of observations used for analyses of FC response; the cross-tabulation now includes the age at which an animal first turned AB⁺. A total of 302 cows contributed 2,485 samples to the specificity assessment of the ELISA. These cows had a minimum of 8 negative FC over a 2-yr period, and only samples obtained before that period were included.

In Figure 2a, the predicted OD_C values are shown for 2 groups of cows, one group never becoming FC^+ and the other group becoming FC^+ at some time during 2 to 6 yr of age. These are the predictions of the ELISA response for “infection.” In Figure 2b, the latter group

Figure 2. Predicted corrected optical density (OD_C) value in *Mycobacterium avium* ssp. *paratuberculosis* antibody ELISA used for milk samples and estimated for groups of cows divided on the basis of the age at which they tested positive in fecal culture (FC), if ever. The number of observations (n) that contributed to each prediction is given in the diagram. a) Predictions are shown for 3 groups of cows: never FC⁺ [minimum (min.) of 1 FC], never FC⁺ (min. of 8 FC), and FC⁺ sometimes from 2 to 6 yr of age. b) Predictions are shown for 5 groups in which FC⁺ cows were divided into groups depending on when they became FC⁺ [never, third year of life (Y), Y4, Y5, or Y6]. The shaded areas around each graph in (a) are the 95% confidence bands; these are omitted in (b) to improve readability.

was further divided into 5 groups, one for each age group in which cows became FC⁺. These are predictions of the ELISA-response for being “infectious.” The estimated probabilities of testing ELISA-positive are shown for the same groups in Figures 3a and b, subsequent to the dichotomization of the ELISA response.

Cows tested before 2 yr of age did not show indications of having antibodies, as the graphs for infected and noninfected cows have the same starting point at 2 yr of age (Figure 2a). The average OD_C value increases steeply from 2 to 2.5 yr of age. After 2.5 yr of age, the increase is less steep. A high OD_C value will typically be in the range 1.0 to 1.4, although OD_C values >2.0 were observed for 120 samples in the data. Thus, the average covers a wide variation of OD_C values. This variation can be examined further by analyzing the data in the age groups, based on when cows became FC positive (Figure 2b). The average OD_C in each of these groups increases more steeply than the average in Figure 2a, thus emphasizing the relation between immune response and bacterial shedding. The almost parallel appearance of the graphs (Figure 2b) is supporting evidence that the age of OD_C positivity is better explained by the age of FC positivity than the age itself; i.e., infectiousness is predicted by the appearance of an immune response.

In Figures 3 and 4, the sensitivity and specificity estimates obtained in the present study can be read on the vertical axis, although they are referred to there as probabilities given a condition. The sensitivity of the test for a given age is read directly as the probability for the positive groups, whereas the specificity given age is read as $1 -$ the probability for the group in which the condition was never found. This condition may be interpreted as infection for Figures 3a and 4a and infectious for Figures 3b and 4b.

The sensitivity of ELISA for prediction of “infection” increases from 0.06 at 2 yr of age to 0.50 at 5 yr of age. The specificity based on samples obtained from cows before 8 negative FC decreases from 0.997 to 0.93 in the same age span, where the specificity is calculated as $1 -$ the probability defined in the graph (Figure 3a). The sensitivity of the ELISA to predict current MAP shedding, or “infectiousness,” in a specific FC⁺ age group (cows FC⁺ between 2 to 3 yr of age) is increasing from 0.06 to 0.58 at 3.6 yr of age. Similar increases in sensitivities are observed in the other groups of cows, stratified into the age group when they became FC⁺ (Figure 3b). Had the age groups been more narrow; e.g., split into half-year intervals, the sensitivity had increased to 0.7 for cows FC⁺ in age range 2 to 2.5 yr of age and 0.9 for cows FC⁺ in age range 2.5 to 3.0 yr

Figure 3. Estimated probability of testing positive in an ELISA for detection of antibodies to *Mycobacterium avium* ssp. *paratuberculosis* for cows grouped according to the year in which they first tested positive in fecal culture (FC⁺). The number of observations (n) contributing to each predicted curve is given in the diagram. a) Predictions are shown for 3 groups of cows: never FC⁺ [minimum (min.) 1 FC], never FC⁺ (min. 8 FC), and FC⁺ sometimes from 2 to 6 yr of age. b) Predictions are shown for 5 groups in which FC⁺ cows were divided into groups depending on when they became FC⁺ [never, third year of life (Y), Y4, Y5, or Y6]. The shaded areas around each graph in (a) are the 95% confidence bands; these are omitted in (b) to improve readability.

Figure 4. Estimated probability of testing positive in a fecal culture test (FC⁺) for detection of *Mycobacterium avium* ssp. *paratuberculosis*, for cows grouped according to the age at which they first tested antibody positive (AB⁺) in an antibody ELISA. The number of observations (n) contributing to each predicted curve is given in the diagram. a) Predictions are shown for 2 groups of cows: never AB⁺ and AB⁺ sometimes between 2 and 6 yr of age. b) Predictions are shown for 5 groups in which AB⁺ cows were divided into groups depending on when they became AB⁺ [third year of age (Y), Y4, Y5, or Y6]. The shaded areas around each graph in (a) are the 95% confidence bands; these are omitted in (b) to improve readability.

of age (data not shown). However, because the data underlying these estimates were sparse, the uncertainty associated with the estimates requires that caution is used when interpreting the results. In Figure 4a, the estimated probability of testing FC⁺ [$P(\text{FC}^+ | \text{AB status})$] is shown for 2 groups of cows: those that never tested AB⁺ and those that tested AB⁺ some time during years 3 to 6 of life, corresponding to the “noninfected” and the “infected” group, respectively. In Figure 4b, $P(\text{FC}^+ | \text{AB status})$ is shown for 5 groups based on the age group in which they became AB⁺, if ever. These are predictions of the FC for being “infectious.” The sensitivity of FC increases with age. At 2 yr of age, the sensitivity is 0.05, and at 5 yr of age, the sensitivity is 0.21. The “specificity” was in the range 0.964 to 0.986 with higher specificity for young cows.

The effect of culture method was insignificant averaged over all groups of AB⁺ ($P = 0.22$). Hence, for ease of interpretation, the results given were averages of both methods.

DISCUSSION

Generally, there is a linkage between ELISA and FC. In this study, this linkage appears to be independent of age. Whereas there is a linkage between the 2 tests independent of age, the sensitivity of both tests still increases with age, and the specificity of ELISA apparently decreases with age. This is the first study describing the test responses of ELISA and FC as a function of age in a large population of naturally Map-infected cattle. The underlying condition for evaluating test responses was varied to allow purpose-specific test interpretation. The following assumptions were used: 1) Map infection takes place in calves, 2) Map infection lasts for life, and 3) some infected cows are not as infectious as others; i.e., the incubation period is variable and cow-specific. Purpose-specific test evaluation and interpretation with inclusion of the important covariate age could be an important extension in the use of imperfect tests in the control of paratuberculosis. The present study can be seen as a first step toward a purpose-specific test evaluation. There are a number of issues, which should be considered in the interpretation of the results, and a number of approaches can be used to improve the interpretation.

Disease Definition, Sensitivity, and Specificity

Given a unique disease definition (the underlying condition), the ability of a test to detect this condition can easily be characterized. However, for paratuberculosis, a suitable, unique disease definition does not exist. For decision makers seeking to eradicate Map from

a herd or region, the appropriate definition of paratuberculosis varies in that decision makers may merely be wishing to reduce economic losses or they may be hopeful of using the test results for appropriate management of infectious cows.

A frequent approach for the evaluation of tests for detection of Map infection is the selection of a sample population, for which infection status is determined postmortem. Such a population is often older than the population that is to be evaluated by the test subsequently. It may also be selected based on criteria, which favor agent-detecting methods (such as FC), because bacteria detected are deemed more indicative of infection than the finding of an immune response. Selection biases are difficult to avoid, because the test scheme chosen as the reference or gold standard essentially defines what constitutes absence and presence of infection (Nielsen and Toft, 2002).

Latent-class methods, which do not presume the definitive status of an animal, can be used, as exemplified in Nielsen and Toft (2002). Currently, these methods do not take into account the purpose of testing, which still need to be addressed. However, latent class methods may provide a promising alternative for the analyses carried out in the present study.

In this study, the repeated testing of animals by FC was used as a reference for the ELISA test, and the repeated testing of milk samples using ELISA was used as a reference for the FC. These are not perfect reference methods, but will underestimate the true sensitivity of each method because the low sensitivity of any reference method will falsely classify some test-positive as false-positive even though they are actually true positives that are not detected by the reference method if the reference method has a low sensitivity. However, to simplify the discussion, we will assume that when used for classification into “infected” or “infectious” the relevant reference method is perfect. Thus, for evaluation of ELISA, repeated FC is a perfect reference method and vice versa. The conditions “infected” and “infectious” were assumed to be the same irrespective of reference method.

Sensitivity and Specificity Given Disease Definition: Infection

Figures 2a, 3a, and 4a illustrate the test responses as a function of age when the underlying condition is infection. Given the aim is to detect an infected animal, these figures show which test responses can be expected with increasing age. For ELISA, the sensitivity increases almost linearly from 0.06 at 2 yr of age to 0.50 at 5 yr of age (Figure 3a), whereas the specificity decreases from 0.997 to 0.93 as the age of the tested animal in-

creases. This apparent drop in specificity is likely to be an artifact caused by the low sensitivity of FC (Figure 4a), although some of this potential artifact can be removed by testing more frequently with FC. The average sensitivity of FC increases from 0.05 at 2 yr of age to 0.21 at 5 yr of age. The specificity is high (0.964 to 0.984). A true false-positive is a positive test response from an animal that is not infected. It is possible, although unlikely, that a bacterium detected is not *Map*, as the cultured bacteria were confirmed with IS900 PCR. However, false-positives may be caused by pass-through, in which cows consume *Map* without becoming infected and lead to a possible false-positive test reaction. Another possibility is that ELISA has not detected cows that were shedding bacteria subsequent to *Map* infection. The conclusion at this stage must be that the ELISA is more sensitive for detection of *Map* infection than FC, but at an early age, it is still low. There may be a risk of obtaining many false-positives, but it is unlikely that they are all attributable to false-positive reactions of the ELISA. Many could be due to the low sensitivity of FC.

Sensitivity and Specificity Given Disease Definition: Infectious

The results shown in Figures 2b, 3b, and 4b are the test responses when the underlying condition is infectious. First, the ELISA response, on average, for a cow that is detected FC⁺ in her third year of life will also test positive within that year (Figures 2b and 3b), although initially, the sensitivity is only 0.06 at 2 yr of age, but at 3 yr of age, it has increased to 0.58 (Figure 3b). The pattern is similar for cows found positive in FC in their fourth and fifth year of life: initially, the sensitivity is low, but steeply increases within the following year. The increases in sensitivity are almost parallel for the different age groups. This indicates that a positive OD_C is coherent with a positive FC and vice versa. Thus, the cow may transfer to the infectious stage any time, regardless of age of the cow (although still in the interval 2 to 5 yr of age). For FC, the sensitivity is also increasing from 0.06 at 2 yr of age to 0.35 at 3 yr of age for cows that are AB⁺ in their third year of life. These results indicate that there is a good relationship between being infectious and AB⁺ and that both the FC and the ELISA may detect some of these animals. However, the ELISA has a greater sensitivity than the FC, although to some degree it may be at the expense of specificity, particularly for older animals. The apparent loss in specificity is likely caused by the generally low sensitivity of the FC method. This can partly be deduced from Figure 2b. A cow testing FC⁺ in her fifth year of life does not have AB before 3 to 3.5 yr of age, whereas

the average of the FC⁻ cows is higher. These would be expected to be similar if the FC⁻ cows were all correctly classified. Given that some FC⁻ animals are actually infected, the sensitivity of ELISA is underestimated. The magnitude of this underestimation could easily be 0.05 to 0.10, corresponding to the apparent drop in specificity.

If the OD_C is used as an approximation of AB production, it is observed that AB are generally present in low levels at 2 to 3 yr of age, whereas higher levels are not reached before the animals are 4 to 5 yr of age. This is consistent with previous findings (Nielsen et al., 2002a). Detected fecal shedding is maximum around 4 to 5 yr of age, with a maximum $P(\text{FC}^+|\text{AB}^+) = 0.17$. The maximum at 4 to 5 yr of age is consistent with the findings by Kalis et al. (1999). If AB was used as the gold standard in a test evaluation, the $P(\text{FC}^+|\text{AB}^+)$ would have corresponded to a diagnostic test sensitivity of 0.17. This is consistent with what has been obtained using methods in the absence of a gold standard (Nielsen and Toft, 2002; Nielsen et al., 2002c), in which sensitivity estimates in the range 0.2 to 0.4 were obtained. The lower sensitivity found here reflects the misclassification of false-positives introduced because of the lack of specificity of the ELISA.

Cows that are never AB⁺ have a probability of 0.02 to 0.04 of testing FC⁺. Hence, most FC⁺ cows will, at some point in time, test positive in ELISA. Those that do not become AB⁺ are potentially passive carriers of *Map* without ever being infected, as described by Sweeney et al. (1992). Another explanation is that they are unable to produce AB, perhaps because they were infected in utero while the immune system was developing. Hence, *Map* could produce persistent infections similar to those occurring with bovine virus diarrhea infections. A last possible explanation is censoring, i.e., FC⁺ cows have not been kept long enough after first shedding has been detected, and therefore, production of AB simply has not begun. The finding suggests that, in most cases, the ELISA will at some time detect infection, but not necessarily before the cow becomes infectious to other cattle.

Concluding Remarks

Discerning between “infection” and “infectious” can be highly relevant for a decision maker, as the infected cow may prohibit a declaration of freedom from infection, but this cow may never become an economic burden. Hence, if a farmer has no intention of declaring freedom from infection, he may only want to focus on detecting infectious animals to reduce spread of infection. Infectious animals can be detected with ELISA with a fair sensitivity in all age groups even with a

single testing. Repeated tests can be required to further characterize the infection state, e.g., whether or not the cow will experience production losses. Infected animals cannot be detected at younger ages, but this is less relevant if the objective only is to detect infectious animals. Thus, a clear aim of the testing strategy should be defined before testing. Given this strategy, an advisor may provide more specific estimates of the probability of detecting the condition, whether this specific condition is infection or infectiousness.

The nature of this study was descriptive; thus, to simplify the presentation of the results, some elements have been given less attention. Consequently, some of the results should be interpreted carefully. For some groups, the number of observations is low, and generally, the right ends of the curves are more uncertain than the rest because of the censoring of old cows. To simplify the results and because of the computational complexity, uncertainty estimates have generally been excluded from the presentation. However, 95% confidence limits are given in Figure 2a, 3a, and 4a. It should be noted that correlation between samples from the same cows has not been addressed, and single observations are not independent. Finally, the predictions presented are average values, which do not take into account the variation in the response to infection of individual cows as previously described (Nielsen et al., 2002b). This variation may be reflected in variation among herds. Variation attributable to herd could not be assessed because of the lack of statistical power. However, there were indications that cows of Herds 2 and 7 did not, on average, yield similar responses as cows in the other herds.

To handle some of the issues raised here, other techniques, such as time-to-event analyses, would be more appropriate, with the events being AB⁺ and FC⁺. Also assessing the relations between the 2 events would be important. However, the purpose of the current studies was primarily descriptive and should be considered as such.

Diagnostic sensitivity and diagnostic specificity are test properties normally characterizing diagnostic tests in control and eradication programs. With a clear disease definition and consensus on actions following diagnosis, sensitivity and specificity are adequate and informative characterizations of a test. However, some decision makers require tests for conditions other than those optimal in eradication programs, for example. The chronic nature of paratuberculosis and the lack of perfect diagnostic tests is a continuous challenge in the interpretation of the test results obtained in any testing. Chronicity affects the sensitivity in that progression of infection increases the sensitivity. In a given population, progression may not only be determined

by a fixed incubation time. It may be influenced by management factors (e.g., feeding strategy) that lead to variation in incubation times. Yet, the fixed incubation time approach can provide us with some information, e.g., by inclusion of age as factor in interpretation of the test results. Another factor is the definition of actions that are made subsequent to a diagnosis, when potential production losses may require one action whereas transmission of Map may call for another, with different requirements at different prevalences.

CONCLUSIONS

The results of this study indicate that the ability of both ELISA and FC to detect "infection" increases almost linearly from 2 to 5 yr of age. The ability of the 2 tests to detect "infectious" is not affected by age. Purpose-specific test evaluations should be conducted to appropriately interpret and use test results for management of paratuberculosis, and relevant covariates should be included when possible. For detection of infection, age is a relevant covariate

ACKNOWLEDGMENTS

The Danish Cattle Federation financed this study, partly via Kongeåprojektet, partly via CEPROS-II/7.

REFERENCES

- Chiadini, R. J., H. J. van Kruiningen, and R. S. Merkal. 1984. Rumiant paratuberculosis (Johne's disease): The current status and future prospects. *Cornell Vet.* 74:218–262.
- Groenendaal, H., M. Nielsen, A. W. Jalvingh, S. H. Horst, D. T. Galligan, and J. W. Hesselink. 2002. A simulation of Johne's disease control. *Prev. Vet. Med.* 54:225–245.
- Hastie, T. J., and R. J. Tibshirani. 1991. *Generalized Additive Models*. Chapman & Hall, London, UK.
- Kalis, C. H. J., J. W. Hesselink, E. W. Russchen, H. W. Barkema, M. T. Collins, and I. J. R. Visser. 1999. Factors influencing the isolation of *Mycobacterium avium* subsp. *paratuberculosis* from bovine faecal samples. *J. Vet. Diagn. Invest.* 11:345–351.
- Larsen, A. B., R. S. Merkal, and R. C. Cutlip. 1975. Age of cattle related to resistance to infection with *Mycobacterium paratuberculosis*. *Am. J. Vet. Res.* 36:255–257.
- Nielsen, S. S. 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subsp. *paratuberculosis* in dairy cattle. *J. Vet. Med. B* 49:384–387.
- Nielsen, S. S., C. Enevoldsen, and Y. T. Gröhn. 2002a. The *Mycobacterium avium* subsp. *paratuberculosis* ELISA response by parity and stage of lactation. *Prev. Vet. Med.* 54:1–10.
- Nielsen, S. S., Y. T. Gröhn, and C. Enevoldsen. 2002b. Variation of the milk antibody response to paratuberculosis in naturally infected dairy cows. *J. Dairy Sci.* 85:2795–2802.
- Nielsen, S. S., C. Grønbaek, J. F. Agger, and H. Houe. 2002c. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. *Prev. Vet. Med.* 53:191–204.
- Nielsen, S. S., B. Kolmos, and A. B. Christoffersen. 2004. Comparison of contamination and growth of *Mycobacterium avium* subsp.

- paratuberculosis* on two different media. J. Appl. Microbiol. 96:149–153.
- Nielsen, S. S., and N. Toft. 2002. Optimisation of the validity of ELISA and faecal culture tests for paratuberculosis: Selection of population or correction by population characteristics? Pages 400–405 in Proc. 7th Int. Colloquium on Paratuberculosis, Bilbao, Spain. R. A. Juste, M. V. Geijo, and J. M. Garrido, ed. Int. Assoc. Paratuberculosis, Madison, WI.
- Stabel, J. R. 2000. Transitions in immune responses to *Mycobacterium paratuberculosis*. Vet. Microbiol. 77:465–477.
- Sweeney, R. W., R. H. Whitlock, A. N. Hamir, A. E. Rosenberger, and S. E. Herr. 1992. Isolation of *Mycobacterium paratuberculosis* after oral inoculation of uninfected cattle. Am. J. Vet. Res. 53:1312–1314.
- Taylor, A. W. 1953. Experimental Johne's disease in cattle. J. Comp. Pathol. 63:355–367.