Short communication

Herd-level prevalence of *Mycobacterium avium* subsp. *paratuberculosis* infection in United States dairy herds in 2007

J.E. Lombard a, *, I.A. Gardner b, S.R. Jafarzadeh c, C.P. Fossler a, B. Harris d, R.T. Capsel d, B.A. Wagner a, W.O. Johnson e

a USDA, Animal and Plant Health Inspection Service, Veterinary Services, Centers for Epidemiology and Animal Health, Fort Collins, CO 80526-8117, USA
b Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Prince Edward Island C1A 4P3, Canada
c Department of Medicine and Epidemiology, University of California, Davis, CA 95616, USA
d USDA, Animal and Plant Health Inspection Service, Veterinary Services, National Veterinary Services Laboratories, Ames, IA 50010, USA
e Department of Statistics, University of California, Irvine, CA 92697, USA

**A R T I C L E   I N F O**

Article history:
Received 18 April 2012
Received in revised form 9 August 2012
Accepted 12 August 2012

Keywords:
*Mycobacterium avium* subsp. *paratuberculosis*
True herd-level prevalence
NAHMS
Composite fecal samples

**A B S T R A C T**

Testing of composite fecal (environmental) samples from high traffic areas in dairy herds has shown to be a cost-effective and sensitive method for classification of herd status for *Mycobacterium avium* subsp. *paratuberculosis* (MAP). In the National Animal Health Monitoring System’s (NAHMS) Dairy 2007 study, the apparent herd-level prevalence of MAP was 70.4% (369/524 had ≥1 culture-positive composite fecal samples out of 6 tested). Based on these data, the true herd-level prevalence (HP) of MAP infection was estimated using Bayesian methods adjusting for the herd sensitivity (HSe) and herd specificity (HSp) of the test method. The Bayesian prior for HSe of composite fecal cultures was based on data from the NAHMS Dairy 2002 study and the prior for HSp was based on expert opinion. The posterior median HP (base model) was 91.1% (95% probability interval, 81.6 to 99.3%) and estimates were most sensitive to the prior for HSe. The HP was higher than estimated from the NAHMS Dairy 1996 and 2002 studies but estimates are not directly comparable with those of prior NAHMS studies because of the different testing methods and criteria used for herd classification.

Published by Elsevier B.V.

1. Introduction

Estimates of the true herd-level prevalence (HP) and within-herd prevalence of infectious agents provide baseline data for assessment of the progress of disease control programs. HP can be estimated using a decision rule based on results of samples (e.g. serum, feces, milk or tissues) from multiple individual animals in each herd or using single or multiple composite samples (e.g. bulk tank milk, milk filters or composite fecal samples) (NAHMS, 1997; Christensen and Gardner, 2000; Adaska and Anderson, 2003; Warnick et al., 2003; Van Kessel et al., 2011).

Ideally, test-based prevalence estimates should be adjusted for the sensitivity and specificity of the selected diagnostic method to enable valid comparisons of estimates among studies that use different tests. There are few estimates of the HP of *Mycobacterium avium* subsp. *paratuberculosis* (MAP) in U.S. dairies. The National Animal Health Monitoring System’s (NAHMS) Dairy 1996 study, which is the only national estimate of MAP prevalence, reported 21.6% of dairy operations infected (NAHMS, 1997). The study was designed to have a 90% confidence of detecting ≥1 positive cow in herds with ≥10% of cows infected. Recent studies suggest that many herds have <10% of cows infected and that the sensitivity of the serum ELISA is lower than what was believed in 1996. Hence, the reported prevalence estimate was conservative (Adaska and Anderson, 2003; Hirst et al., 2004; USDA,
The NAHMS Dairy 2002 study also evaluated MAP infection but the low number of herds sampled and the fact that these herds were not randomly selected did not allow calculation of a national estimate (USDA, 2005). Of the 98 herds that were sampled in the 2002 study, 69 (70.4%) had at least one composite fecal sample that cultured positive for MAP (Lombard et al., 2006b), providing evidence that MAP prevalence in dairy operations was higher than previously reported.

Multiple studies have evaluated the use of composite fecal samples – samples from high traffic areas where manure from a large number of cows is deposited – to estimate herd-level and/or within-herd MAP prevalence in dairy herds (Raizman et al., 2004; Berghaus et al., 2006; Lombard et al., 2006b; Pillars et al., 2009; Aly et al., 2009). Although the number of composite fecal samples collected was not the same in all studies (2–6 samples/operation), all studies reported high herd sensitivity (HSe) (>70%) or detection of more than 70% of tested herds. False-negative results most likely occur when the within-herd prevalence and environmental load of MAP are low or when the only cows shedding MAP are non-lactating and not contributing feces to high traffic areas. False-positive culture results are considered rare and usually occur due to laboratory contamination.

The objective of the present study was to estimate the true HP of MAP infection in U.S. dairy herds in 2007 based on Bayesian analysis of culture results of composite fecal samples.

2. Materials and methods

2.1. Study overview

Serologic and culture data from the NAHMS Dairy 2002 study were used to construct a Bayesian prior for HSe of composite fecal cultures. The Bayesian prior for herd specificity (HSp) was based on expert opinion and allowed for rare false-positive results. The priors were then used to estimate true herd-level prevalence and predictive values of MAP from composite fecal samples collected from a nationally representative random sample of dairy operations during the NAHMS Dairy 2007 study.

2.2. Data sources

2.2.1. NAHMS Dairy 2007 study

In the 2007 study, a random sample of 3554 dairy operations from 17 major dairy states (California, Idaho, Indiana, Iowa, Kentucky, Michigan, Minnesota, Missouri, New Mexico, New York, Ohio, Pennsylvania, Texas, Virginia, Vermont, Washington, and Wisconsin) was eligible for participation. Of those 3554, 3304 (93%) were contacted by the National Agricultural Statistics Service. There were 2519 operations that completed the initial questionnaire and 1077 were eligible (30 or more cows on January 1, 2007) and consented to contact by a veterinary medical officer and potentially continue with the next phase of the survey. Of the 1077 operations that consented, 582 operations continued in the study and were eligible for testing of composite fecal samples for MAP (USDA, 2008a). Of the 582 eligible operations, 524 (90%) participated in the collection of composite fecal sample from areas on the dairy where manure from a majority of cows accumulated. Federal and state animal health officials collected samples from 6 different locations on each operation. Instructions were given to sample from areas specifically listed on the collection form (common alleyway, common pen, exit way from parlor, floor of holding pen, flush water, gutter cleaner, lagoon, manure pit, and manure spreader) rather than create samples from areas designated as ‘other’. Samples were sent overnight on ice to the USDA-APHIS-VS, National Veterinary Services Laboratories (NVSL) in Ames, IA for culture. Samples were cultured on Herrold’s egg yolk (HEY) agar (Becton Dickinson Diagnostic Systems, Sparks, MD). Two flasks containing HEY agar with Mycobactin J, two tubes of HEY agar with Mycobactin J, and one tube of HEY agar without Mycobactin J were inoculated. IS900 PCR was used to confirm positive cultures as MAP.

2.2.2. NAHMS Dairy 2002 study

Information from this study has been previously published (USDA, 2003; Lombard et al., 2006a,b). Briefly, a subset of 98 herds from 21 major dairy states (California, Colorado, Idaho, Illinois, Indiana, Iowa, Florida, Kentucky, Michigan, Minnesota, Missouri, New Mexico, New York, Ohio, Pennsylvania, Tennessee, Texas, Virginia, Vermont, Washington, and Wisconsin) was purposely selected from 1013 randomly selected operations that participated in the study. To be eligible to participate, operations had to have completed the initial questionnaire and have ≥30 dairy cows on January 1, 2002. Five composite fecal samples (compared with 6 in 2007) were collected by federal and state animal health officials from areas on these operations where manure accumulated from multiple adult cows. Samples were shipped on ice to the NVSL for culture. In contrast to the 2007 study, each sample was cultured by 3 methods (Herrold’s egg yolk agar, Becton Dickinson Diagnostic Systems, Sparks, MD; BACTEC™ 460TB System, Becton Dickinson Diagnostic Systems; and ESP® Culture System II, Trek Diagnostic Systems, Cleveland, OH). If a sample was positive by any method, the sample was deemed positive. If one or more cow samples were culture positive for a particular operation, the herd was designated as culture-positive for MAP. The HSp of culture of composite samples was assumed to be perfect. Two of the 98 herds were excluded from estimation of the HSe of testing of composite fecal samples because of improper sampling. In one herd, samples were collected only from heifer areas and in the other, individual cow samples were collected rather than composite fecal samples. In order to estimate HSe based on less than 5 composite fecal samples, models created in SUDAAN® software (Release 10.0.1 2010, Research Triangle Institute, Research Triangle Park, NC) using the hypergeometric distribution were used to estimate HSe for 1–4 samples drawn without replacement from the population of composite fecal samples.

Serum ELISA testing was performed on the whole herd or a sample of the herd from these 96 operations. Blood samples were shipped overnight on ice to NVSL for testing. A commercially available ELISA kit (Paracheck, Biocor Animal Health, Omaha, NE) was used for testing as directed.
by the manufacturer with the exception that samples were only tested in a single well, instead of in duplicate. Results were reported as negative or positive at the manufacturer-recommended threshold based on a computed ELISA score of 1.0. None of the herds used MAP vaccines although it is possible that some purchased cows might have been vaccinated. The apparent within-herd seroprevalence was calculated as the number of cows testing positive divided by the total number of cows tested for each herd. Based on the reported ELISA sensitivity of 29% and specificity of 99.7% (Collins et al., 2006), true within-herd prevalence for each herd was calculated from apparent prevalence (Rogan and Gladen, 1978). True prevalence was used to categorize herds in 3 groups based on the within-herd MAP estimate. True MAP within-herd prevalence ranges for the low, moderate, and high categories were >0–12%, >12–25%, and >25%, respectively.

2.3. Bayesian analysis

2.3.1. Prior distributions for HSe and HSp of composite fecal samples, and true HP

The priors for HSe were based on the conservative assumption that collection of 6 samples in 2007 would have the same HSe as 5 samples did in the 2002 study (Table 1) assuming a non-informative (beta 1, 1) prior before the study was done. Therefore, the corresponding HSe prior was beta (62, 19). The prior for HSp was based on the expert opinion of one of the coauthors (BH) who worked in the testing laboratory involved in the 2002 and 2007 studies. The HSp prior (beta (9999, 1)) allowed for rare false-positive results (approximately 1 in 10,000) if the 6 samples were all truly negative (i.e., from a non-infected herd) and cross-contamination of 1 or more samples occurred in the testing laboratory. The HP prior was non-informative (beta (1, 1)).

2.3.2. Bayesian model for herd prevalence

The model was based on the animal-level prevalence model of Branscum et al. (2004) where HP, HSe and HSp were substituted for their individual counterparts. Each herd’s infection status was assumed to be independently Bernoulli distributed such that

\[ Y_i \sim \text{Bernoulli}(p_i), \]

where \( Y_i \) was the test result (\( Y_i = 1 \) if positive and \( Y_i = 0 \) if negative) for herd \( i \), and \( p_i \) was the probability of a positive result for herd \( i \). Predictive values of herd-level negative and positive test results were calculated because a population-based design was used.

The default priors for HP, HSe, and HSp were as described in Section 2.3.1. Posterior distributions were approximated using Markov-chain Monte Carlo (MCMC) methods in WinBUGS (Lunn et al., 2000). Posterior medians and 95% probability intervals (PI) were used for inferences and were based on 50,000 iterates after a burn-in of 5000 iterates. Convergence of the MCMC chain after the burn-in period was assessed by evaluation of trace plots and running multiple chains from different starting values.

Sensitivity analysis was done by evaluating the effects of changing priors for HSe and HSp. For HSe, an expert (R. Whitlock, personal communication) indicated that he might expect a slight increase (2–4%) in HSe in culturing 6 rather than 5 composite samples without considering the data in Table 1. This slight increase in HSe is reasonable given the trend in HSe evident in the 2002 data. The increase in HSe corresponded to the detection of 1 additional infected herd in each of the 3 prevalence categories and hence, the corresponding prior for HSe was beta (65, 16). A second less optimistic prior for HSe (beta (56, 25)) was used. The median HSe value for this prior was 69.3%, which approximated the lower limit of values reported in prior published studies (see Section 1). The effect of a pessimistic prior for HSp (beta (99, 1)), which allowed for approximately 100 times the number of false-positives results compared with the default prior, was also evaluated.

3. Results

3.1. Test prevalence in 2007

Of the 524 tested herds in 2007, 369 (70.4%) had at least 1 of 6 samples that was culture positive for MAP (USDA, 2008b). In the culture-positive herds, the frequency distribution of 1–6 positive samples was 37 (10.0%), 38 (10.3%), 23 (6.2%), 48 (13.0%), 65 (17.6%), and 158 (42.8%), respectively.

3.2. Calculation of HSe using 2002 data

HSe was estimated using 1–5 composite fecal samples for 3 levels of herd prevalence and the results are presented in Table 1. As expected, as the number of composite samples tested increased and the estimates of within-herd true

### Table 1

<table>
<thead>
<tr>
<th>No. of composite fecal samples</th>
<th>Herd-level sensitivity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low within-herd prevalence (≥0–12.0%)</td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>5</td>
<td>55.6</td>
</tr>
<tr>
<td>4</td>
<td>55.6</td>
</tr>
<tr>
<td>3</td>
<td>48.1</td>
</tr>
<tr>
<td>2</td>
<td>44.4</td>
</tr>
<tr>
<td>1</td>
<td>33.3</td>
</tr>
</tbody>
</table>
prevalence increased, the HSe increased. Culture of 5 composite fecal samples across all herds had a HSe of 77.2%.

3.3. Herd prevalence and predictive values

For the base model using a beta (62,19) prior for HSe, the posterior median herd prevalence was 91.1% (95% PI, 81.6% to 99.3%). Sensitivity analysis indicated that posterior estimates changed most with changes in the HSe prior (Table 2) and minimally with a decrease in the HSp prior (results not shown). For all models, the herd-level positive predictive values were close to 1 but median herd-level negative predictive values ranged from 13% to 43% depending on the HSe prior. Probability intervals for herd-level negative predictive value were wide.

4. Discussion

In the present study, we adapted code developed by Branscum et al. (2004) to allow estimation of true HP of M. avium subsp. paratuberculosis and the corresponding herd-level positive and negative predictive values in United States dairy herds, as a function of HSe and HSp of culture of 6 composite environmental samples.

Within-herd prevalence and the level of MAP contamination of the herd environment (environmental load) are important covariates affecting the HSe of composite fecal samples for herd level detection. Prevalence and environmental load may not be related depending on the association between shedding level in individuals and the proportion of heavy shredders. The environmental load and probability of a pen testing positive from composite fecal sampling was positively but not statistically correlated with the number of animals in the pen shedding in 3 low-prevalence herds (Smith et al., 2011). This relationship is further complicated by the phenomenon of MAP super-shedding (Whitlock et al., 2005; Aly et al., 2012).

Composite fecal sampling has received quite a bit of attention over the past few years. It has been used in a number of studies to evaluate MAP infection, primarily in dairy herds. This sampling method has also been used to evaluate Salmonella at the herd level for dairy operations and performed similarly to individual animal sampling (Lombard et al., in press). Composite fecal sampling is less costly and resource intensive than sampling individual cows to determine the herd MAP infection status while maintaining a relatively high sensitivity (~70%) (Lombard et al., 2006b).

The NAHMS Dairy 2002 study, to the authors' knowledge, collected composite fecal samples from the largest number of operations (n = 96) prior to the Dairy 2007 study and also had individual animal sampling to determine within-herd prevalence. The estimates of HSe for composite fecal sampling used in this study were modeled using the 2002 data. Pillars et al. (2009) used a HSe of 81% and HSp of 100% but only collected 2 composite fecal samples from each of the 94 sampled operations. This is likely the reason the reported true prevalence was only 49%; however, a value of 67% would have been obtained if a more appropriate HSe value of 59.5% (Table 1) based on the 2002 data was used in the analysis.

Smith et al. (2011) compared quarterly composite fecal sampling results from 3 low prevalence herds in the Northeastern U.S. to individual cow fecal culture results and reported a HSe of 40% when 6 composite samples were collected. The NAHMS 2002 results suggest that collection of 5 composite fecal samples would result in a HSe of 55.6% – almost 50% higher than that calculated by Smith et al. (2011).

The posterior median HP (base model) was 91.0% (95% probability interval, 81.5 to 99.3%) and estimates were most sensitive to the HSe prior. The HP was higher than estimated from the NAHMS Dairy 1996 and 2002 studies but estimates are not directly comparable with those of prior NAHMS studies because of the different testing methods and criteria used for herd classification. Based on the NAHMS Dairy 2007 study and outcome of this modeling, it suggests the majority of dairy operations in the US are infected with MAP.

Conflict of interest statement

None.

Acknowledgements

Funding for data analysis was provided by the Johne's Disease Integrated Program (USDA-NIFA Coordinated Agricultural Project No. 2008-55620-18710). We thank Drs. Robert Whitlock and Michael Collins for providing expert opinion about sensitivity of composite sampling and the within-herd prevalence distribution, respectively. We also thank Patrick Camp, Philip Dykema, and Gabe Wilson for providing laboratory support in the culturing and identification procedures for the submitted samples.

Table 2

<table>
<thead>
<tr>
<th>Herd-level sensitivity</th>
<th>Prevalence (95% PI)</th>
<th>Negative predictive value (95% PI)</th>
<th>Positive predictive value (95% PI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base model</td>
<td>0.911 (0.815–0.993)</td>
<td>0.302 (0.025–0.593)</td>
<td>1(0.999–1)</td>
</tr>
<tr>
<td>Lower sensitivity</td>
<td>0.957 (0.866–0.998)</td>
<td>0.143 (0.006–0.427)</td>
<td>1(1–1)</td>
</tr>
<tr>
<td>Higher sensitivity</td>
<td>0.873 (0.785–0.978)</td>
<td>0.428 (0.079–0.687)</td>
<td>1(0.999–1)</td>
</tr>
</tbody>
</table>
Appendix A.

Code for the base model when herd sensitivity is modeled as independent of within-herd true prevalence. Abbreviations: hap = herd-level apparent prevalence; hse = herd sensitivity; hsp = herd specificity; hp = true herd-level prevalence; hppv = herd-level positive predictive value; and hnpv = herd-level negative predictive value.

model {
  y ~ dbin(hap, n)
  hap <- hp * hse + (1 - hp) * (1 - hsp)
  hse ~ dbeta(62, 19)
  hsp ~ dbeta(9999, 1)
  hp ~ dbeta(1, 1)
  hppv <- hse * hp / ((hse * hp + hp + hsp * (1 - hp)))
  hnpv <- hp / ((1 - hse) * hp + hsp * (1 - hp))
}

References


