A Bayesian model to investigate the effect of test variability on the prevalence estimate for Paratuberculosis

G. van Schaik[1*], Y.H. Schukken[2], C. Crainiceanu[3], J. Muskens[4], J.A. VanLeeuwen[5]
[1]Instituto de Medicina Preventiva Veterinaria, Universidad Austral de Chile, Valdivia, Chile [2]Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA,[3]Department of Statistical Sciences, Cornell University, Ithaca, USA,[4]Animal Health Service, Deventer, The Netherlands, [5]Department of Health Management, Atlantic Veterinary College, UPEI, Charlottetown, Canada

Summary
The objective of the study was to determine the effect of different definitions for infected herds on the herd-level true prevalence (TP) with a Bayesian model. Data of two seroprevalence studies for *Map* were available. The Bayesian model estimated the herd-level TP by calculating the estimated proportion of infected herds where the herd sensitivity (HSe_k) was over 50%. The HSe values were varied to 51%, 55% and 60% to investigate the changes in herd-level TP. The herd-level TP changed considerably with the different cut-off values and when the definition of an infected herd was changed from 1 test positive to 2 test positive cows.

Introduction
The ELISA tests that are available to detect an infection with *Mycobacterium avium* subsp. *paratuberculosis* (MAP) have a limited sensitivity (Se) and specificity (Sp). In many studies, the Se and Sp of the tests are treated as constants and this will result in an underestimation of the variability of the true prevalence (TP). A Bayesian model was developed to include test variability in the estimation of TP for MAP within herds, regions and locations. The objective of the current paper is to determine the sensitivity of the herd-level TP for different definitions for infected herds.

Study data
Data of two prevalence studies for *Map* using an ELISA in two countries were available for the analyses. In location 1, in approximately 90 randomly sampled herds in each of the four regions of the country, all cattle of at least 3 years old were sampled. In total, 15,468 cows from 371 non-vaccinating herds were tested with an absorbed IDEXX ELISA (IDEXX Skandinavia AB, Sweden). The optimal cut-off value for the test was obtained for the specific study population prior to the study and resulted in a Se for different MAP stages (sub-clinical to clinical) between 28 and 41%. The Sp was estimated at 99.1% with a range of 98.7-99.5%. In location 2, in 30 randomly sampled herds in each of three regions, approximately 30 randomly selected cows were sampled. A total of 2604 cows were tested in duplicate for antibodies with an absorbed IDEXX ELISA (IDEXX Laboratories, Westbrook, Maine) according to the manufacturer's recommendations. The mean Se was 50%, and ranged from 15.4% for sub-clinical cows to 88.1% for clinical cows. The mean Sp was 96.8% with a range from 91.6 to 100%.

Bayesian model

In the Bayesian analysis, all parameters were considered random variables whose variability was modelled by their probability distributions. Information about the unknown test Se, Sp, apparent and TP was incorporated into the model by joint probability distributions. The prior distribution of the AP was based on the data of the prevalence studies, the prior distributions of the test Se and Sp were based on earlier test validations, and the prior distribution of the TP was based on the uninformative prior distributions of herd-size, country and random herd-within-region effect. Posterior estimates were obtained by an iterative process, Markov Chain Monte Carlo (MCMC) simulation that estimates the parameters from the joint posterior distributions (the model parameters given the data and the prior distributions) through Bayes' formula. The model was implemented in WinBUGS software version 1.3.

In the model, a herd was defined as the adult milking cows in a herd and herd k was considered nested in region j and the regions were nested in location i. The model for the cow-level TP was as follows:

$NT+_k \mid P(T+)_k, n_k \sim \text{Binomial}(P(T+)_k, n_k)$

$P(T+)_k = Se_i*TP_k+(1-Sp_i)*(1-TP_k)$

$\text{logit}(TP_k) = \beta_1*N_k+ \beta_2*\text{location}_i+ t_{j(k)}$

$t_{j(k)} \mid r_{i(j)}, \sigma_t \sim \text{Normal}(r_{i(j)}, \sigma_t)$

$r_{i(j)} \mid \mu, \sigma_r \sim \text{Normal}(\mu, \sigma_r)$

$Se_i \mid a_i,b_i \sim \text{Uniform}(a_i,b_i)$

$Sp_i \mid c_i,d_i \sim \text{Uniform}(c_i,d_i)$

$NT+_k$ was the observed number of cows that tested positive in herd k, and was assumed to have a binomial distribution, conditional on the probability that a randomly selected cow from herd k tested positive, $P(T+)_k$, and the sample size n_k. $P(T+)_k$ was the proportion of test positive reactors in herd k, which was the animal-level apparent prevalence (AP). The TP_k was restricted between 0 and 1 and the logit of the TP_k was modeled as a function of the covariates and the random effects. Herd size was included in this model as an orthogonal covariate N_k (mean = 0, standard deviation = 1) to improve mixing of the model. Location was included as a fixed covariate in the model. Conditional on their mean, the random herd ($t_{j(k)}$) and region effects ($r_{i(j)}$) were assumed to be independent and normally distributed. The random herd effect quantified the between-herd variation within geographic regions and the random region effect quantified the between-region variation within location. The random herd effects $t_{j(k)}$ were centered on the random region effect $r_{i(j)}$; $r_{i(j)}$ was centered on μ. The cow level sensitivity and specificity, Se_i and Sp_i were assumed to have a uniform distribution with ranges in test Se_i and Sp_i obtained from the test evaluation in location 1 and location 2. The confidence intervals of test Se_i and Sp_i were used to express the variation in test validity. The model for the estimated herd-level TP was as follows:

$HSe_k = 1 - (P(T-)_k)^n \mid p(T+) > 0.5$

Where n = number of cows tested and $P(T-)_k = 1- P(T+)_k$.

The herd sensitivity (HSe) was the probability to declare an infected herd infected. We estimated the HSe_k for all herds because all herds had a probability of being infected, $P(T+)_k$, although when $P(T+)_k$ approached zero, the herd was very unlikely to be infected. For our purposes, when the probability of detecting at least one reactor in a herd (HSe_k) was larger than 50%, the herd was considered infected, otherwise the

herd was considered not infected. The model estimated the herd-level TP by calculating the estimated proportion of infected herds where HSe_k>50%. The sensitivity of the herd-level TP estimates for the HSe cut-off value of 50% was investigated by varying the values to 51%, 55% and 60%. In addition, a herd was considered infected when the probability of detecting at least two reactors in a herd ($HSe2_k$) was larger than 50%.
The priors were kept non-informative to avoid them influencing the outcomes of the model. The convergence of the MCMC samplers and the robustness of the model were assessed by considering the convergence of several parallel chains and by changing the number of iterations, discarding a number of initial iterations, and changing the prior distributions.

Results

The model was run for 20,000 simulations of which the first 4,000 simulations were discarded from the summarizing statistics. The model converged rapidly and was insensitive to increased numbers of iterations, to discarding a larger part of the initial iterations, and to different prior distributions (e.g. for test Se and Sp). The simulation results of the model parameters showed that herd size and location had no effect on the odds of being infected. The corrected cow-level TP was low, 5.8% and 3.6% in location 1 and 2, respectively. Certain regions within a location differed significantly in herd-level TP. The herd-level TP was 54.3% in location 1 and 32.9% in location 2. The variation in the herd-level TP estimate for location 2 was more than three times as large as the variation in location 1, mainly because of the relatively small number of investigated herds in location 2 (see Table 1).

Table 1: Mean herd-level TP in location 1 and 2 as reported in the original studies and for different scenarios with herd sensitivity to determine infected herds.

		Mean (%)	SD (%)	2.5% CL (%)	97.5% CL (%)
AP	Location 1	54.7	-	49.7	59.8
	Location 2	43.3	-	32.9	53.8
HSe>50%	Location 1	54.4	4.3	46.4	63.1
	Location 2	32.4	11.2	15.6	60.0
HSe>51%	Location 1	52.7	4.3	44.7	61.5
	Location 2	30.4	12.4	12.2	64.4
HSe>55%	Location 1	47.2	3.8	39.9	55.0
	Location 2	24.3	8.0	10.0	40.0
HSe>60%	Location 1	40.5	3.4	34.2	47.4
	Location 2	18.5	6.4	6.7	31.1
HSe2>50%	Location 1	18.4	1.9	14.8	22.4
	Location 2	5.2	2.3	0.0	11.1

The herd-level TP in the 2 locations changed considerably with the different cut-off values and when the definition of an infected herd was changed from 1 test positive to 2 test positive cows. Both MAP ELISAs were not very specific (mean Sp 96.8% and 99.1%), leading to false positive test results at the cow and herd levels. In future prevalence studies for MAP, sample size calculations should be based on a very low cow-level prevalence. Approximately 50% and 90% of the herds in location 1 and 2 had an estimated cow-level TP below 4% and 10%, respectively.