Evaluation and validation of the Koru Diagnostics ELISA test to determine *Staphylococcus aureus* infection in cows with elevated somatic cell count

Dr D.A.Yang
Department of Statistics, The University of Auckland
School of Veterinary Science, Massey University

June 2020

Summary

The Koru ELISA shows consistent accuracy of AUC = 0.987 in detecting *S. aureus* infection in HT1 and HT4 samples.

The Diagnostic Sensitivity (Se) and Diagnostic Specificity (Sp) for the combined HT1 and HT4 data were found to be 0.924 and 0.954 when a cut-off Sample/Positive (S/P) ratio of 0.30 was applied.

The Diagnostic Sensitivity (Se) and Diagnostic Specificity (Sp) for HT4 data only were found to be 0.906 and 0.942 when a cut-off Sample/Positive (S/P) ratio of 0.30 was applied.

Aims

The aim of this statistical analysis was to evaluate the performance of the novel KORU ELISA test in determining the *S. aureus* infection status of herd test milk samples.
It is important to note that a test result itself is fixed, however, given the test result, whether or not the animal is truly infected with *S.aureus* bacteria, is speculative. Therefore, a statement can be made regarding the probability of infection, using defined ELISA test S/P ratios.

Biological Assumptions

We assumed that the 'positive' figure in the ELISA S/P ratio was associated with the *S. aureus* infection. Particularly, we assumed that the mean S/P ratio was higher in the infected group than the uninfected group, which holds for most of the diagnostic tests. The *S. aureus* infection was unobserved, hence treated as a latent class, which was linked to the observed PCR results, where the proportion of PCR-Positive was considered as the 'apparent prevalence'.

Results from two screening tests (ELISA and PCR) were used, and neither of them was a perfect reference test. The two tests were assumed to be conditionally independent, with PCR detecting pathogens and the ELISA detecting antibodies. The conditional independence between ELISA and PCR is a widely accepted assumption in the literature of diagnostic test evaluation, thus it is reasonable to adopt this assumption in our case.

Model Development

We assumed that a suitable transformation was applied so that the transformed S/P ratio was a mixture of two normal distributions. The transformed S/P ratio for the i^{th} animal was denoted as R_i, and it was only dependent on the *S. aureus* infection status S_i, which was a Bernoulli random variable with the prevalence p_i. Therefore, we would have:

$R_i|S_i=k \sim N(\mu_k, \sigma^2_k)$, where k=2 indicates *S. aureus*-infected and k=1 indicates *S. aureus* non-infected.

The addition of information of the PCR results was denoted as Z_i, which was also a Bernoulli random variable with the parameter q_i measuring the probability of being test-positive.

By using the partition theorem, we have:

$q_i = p_i\, Se + (1 - p_i)(1 - Sp)$, where Se and Sp were the sensitivity and specificity of PCR test.

If there were no covariates and no random effects, then p_i = p(the overall prevalence)
If there were multiple herds, then $p_i = p_{j(i)}$, where j was the herd for cow i. We then have a random effect distribution for p_j and this could be modelled using a logit-normal distribution such that:

Logit $(p_{j(i)}) = \beta_0 + U_j$, where $U_j \sim N(0, \sigma^2_{herd})$

The inference we want to make is the predicted probability of *S. aureus* infection given the different S/P ratio of the ELISA, such as P(S|R).
Based on the Baye's theorem:

$$P(S = 1|R = r) = \frac{P(S = 1)f(R = r|S = 1)}{P(S = 1)f(R = r|S = 1) + P(S = 0)f(R = r|S = 0)}$$

where, $P(S = 1) = \frac{\exp(\beta_0)}{\exp(\beta_0)+1}$ was the overall prevalence of *S. aureus* infection and $f(R|S = k)$ was the probability density function of the normal random variable.

Eventually, we constructed the ROC curve in the absence of a gold standard. The sensitivity and specificity of the ELISA test were computed given distinct cut-off values c (transformed) as:

$$Se(c) = 1 - \Phi\left(\frac{c - \mu_2}{\sqrt{\sigma_2^2}}\right)$$

$$Sp(c) = \Phi\left(\frac{c - \mu_1}{\sqrt{\sigma_1^2}}\right)$$

where Φ is the cumulative distribution function of a standard normal random variable. The set of sensitivity and 1-specificity values can then be plotted to present a ROC curve.

Data Analysis

This analysis was based on the HT1 dataset.

PCR 'Suspect-positive' was combined with PCR-positive. The distribution of the ELISA S/P ratio grouped by PCR result was plotted. It could be seen from the plot that the distribution of ELISA S/P ratio in the PCR-negative animals was right-skewed with a long tail, whereas the distribution of ELISA S/P ratio in PCR-positive animal was more symmetrical (Fig 1). This suggested that the PCR has relatively low sensitivity but high specificity. This proved that combining the PCR 'Suspect-positive' with PCR-positive was justified. Combining PCR Suspect-positive with PCR-negative became less desirable as it would decrease the sensitivity of the PCR.

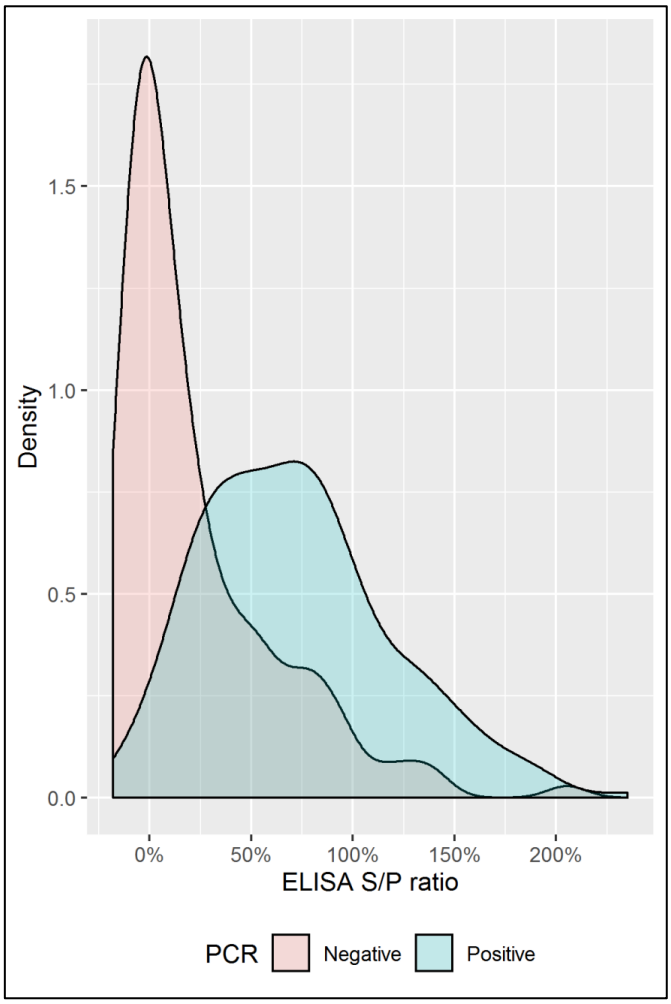


Fig 1. Density plot of untransformed ELISA S/P ratio grouped by PCR results.

The theoretical model assumed the ELISA results would be modelled as a mixture of normal distributions. Box-Cox transformation was used given that the raw ELISA S/P ratio is right-skewed. After transformation, the distribution was examined using a histogram. It was evident that Box-Cox transformation provided close approximation to a mixture of normal distributions (Fig 2).

Therefore, the transformed ELISA S/P ratio was used in the later analysis and denoted as R_i.

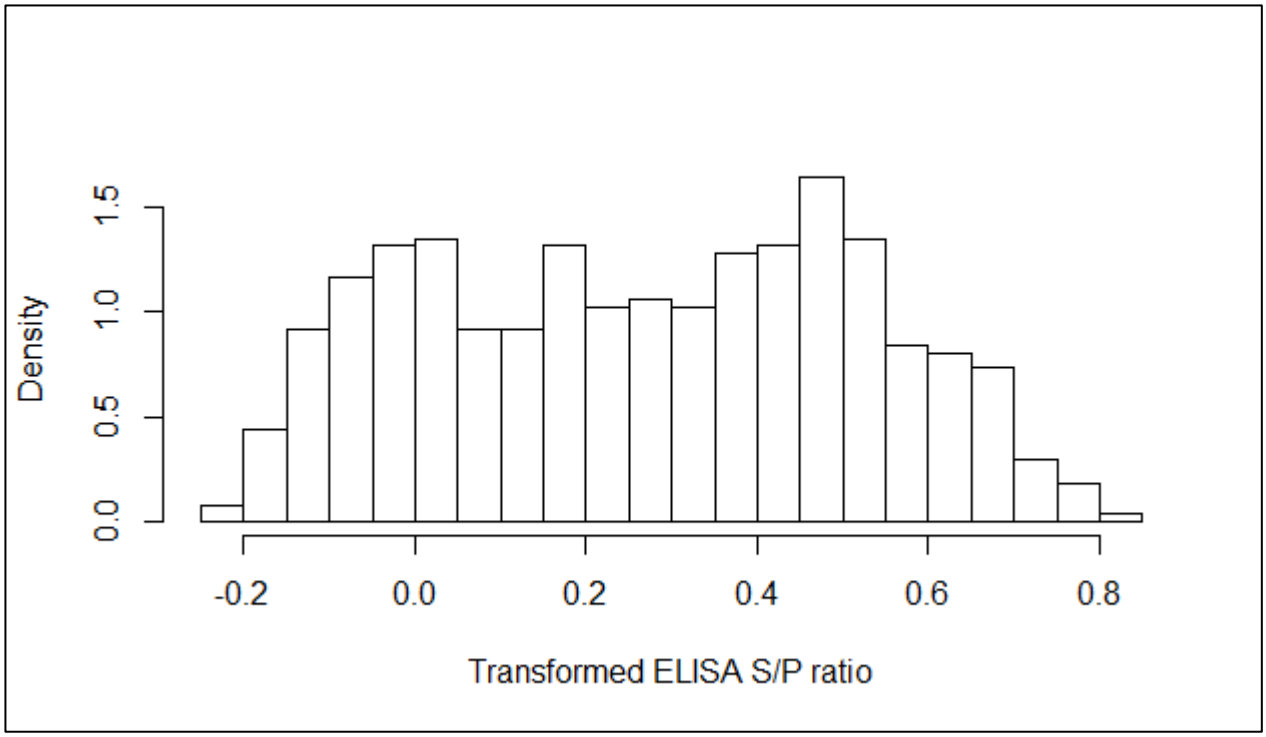
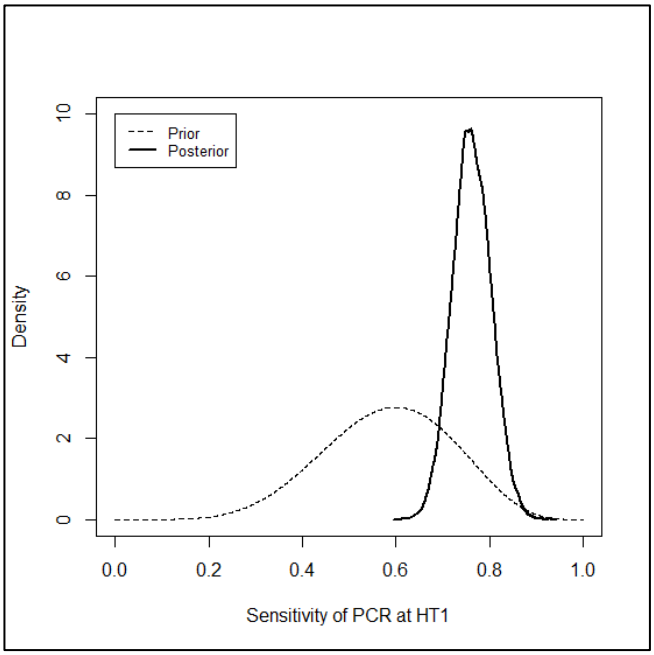
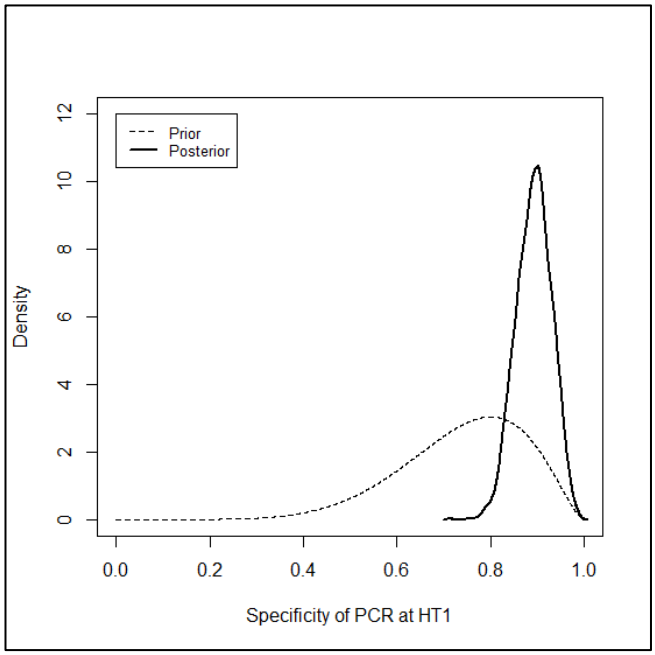


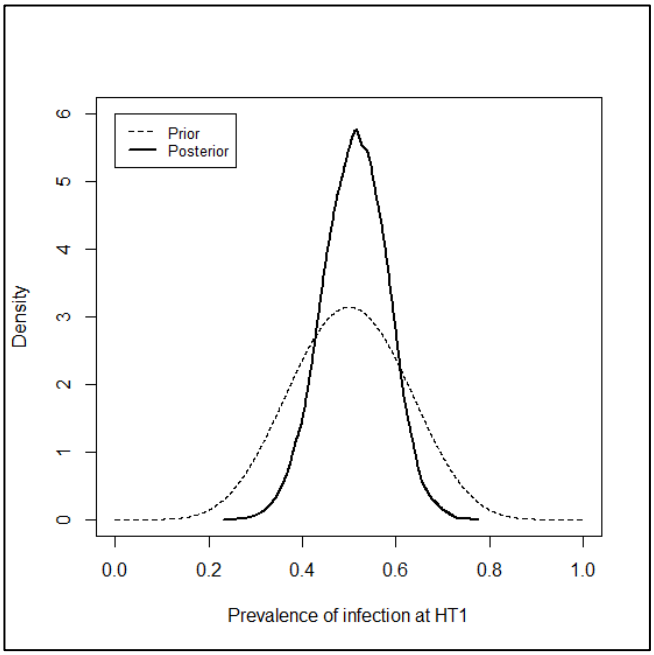
Fig 2. Histogram of transformed ELISA S/P ratio.

After the preliminary analysis, the aforementioned latent class model was fitted into the data. In the Bayesian analysis, priors were required. I used informative but only moderately strong priors for the Se and Sp of PCR and the overall prevalence of *S. aureus* infection across the herds. According to the previous information from the data, the best estimate of the Se was set at 0.6, with its 95th percentile = 0.8 and this corresponded to Beta (7.0375, 5.025) distribution. The best estimate of the Sp was 0.8 with 5th percentile = 0.5, corresponding to Beta (7.5485, 2.6371). Both priors allowed moderately high uncertainty, which enabled more contribution from the data to the posteriors. Beta (8, 8) was used for the overall prevalence; it was relatively less informative, allowing a wide range of values, however maintaining the best estimate of prevalence = 0.5.

The prior for the standard deviation of the random effect term σ_{herd} is specified as Unif (0,1). Several more diffuse uniform distributions were also placed to confirm its minimal impact on the posteriors. The posterior distributions simulated using Markov chain Monte Carlo were compared with the prior distributions. It was evident that the posterior distributions were obviously different from the prior distribution (updated by the data), suggesting insights were gained by performing the analysis on the data. It can be seen that the posteriors were narrower than the prior distributions and all the information contained in the posteriors was reasonable. As additional information, the differences between posteriors and priors of the PCR Se and Sp and the prevalence of infection are displayed below, however these plots are less important as they are not related to the performance of the ELISA.







The predicted probability of an *S. aureus* infection given the untransformed S/P ratio was displayed graphically (Fig 3). It was clear that when the S/P ratio was greater than 0.30, the probability of infection would be > 50%; if the S/P ratio was greater than 0.35, the probability would be >70%.

The predicted probability of infection given an ELISA S/P ratio, namely, P(S=1|R=r) is in general NOT the same as the sensitivity of the ELISA test, namely, P(R>c|S=1) based on a particular cut-off value c. The former is a posterior probability, computed using the Bayes' theorem, where the probability of a randomly selected individual being truly infected, namely, P(S=1) of a population is involved. Therefore, the predicted probability is affected by the P(S=1) of a given population, which means it is not consistent across populations. While the Se and Sp are the parameters describing the accuracy of a diagnostic test, they are independent of P(S=1) in a given population.

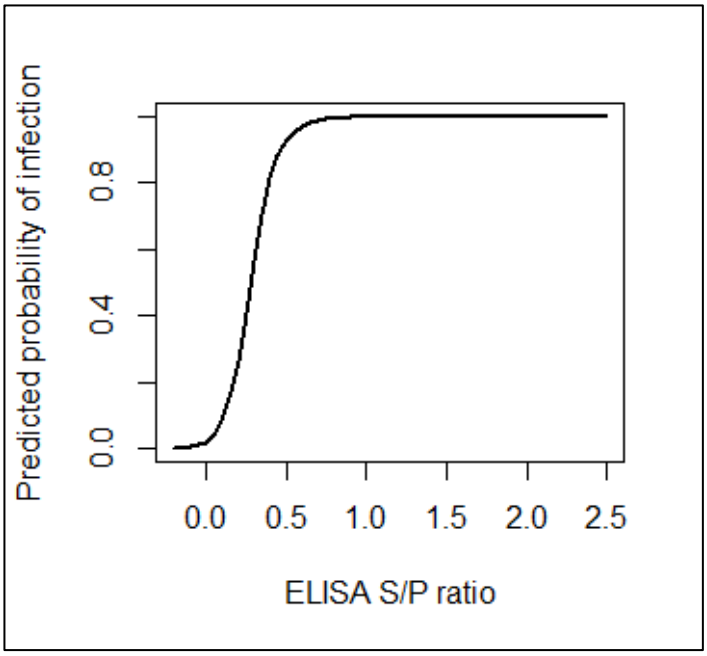


Fig 3. Predicted mean probability of infection as a function of ELISA S/P ratio

A ROC analysis suggested that the KORU ELISA has a high overall accuracy (AUC=0.9787). The ROC curve is presented in Fig 4. The optimal cut-offs were in the range of 0.30-0.35.
Table 1 provides a more detailed analysis of the distinct choices of candidate cut-off values in the aforementioned range.

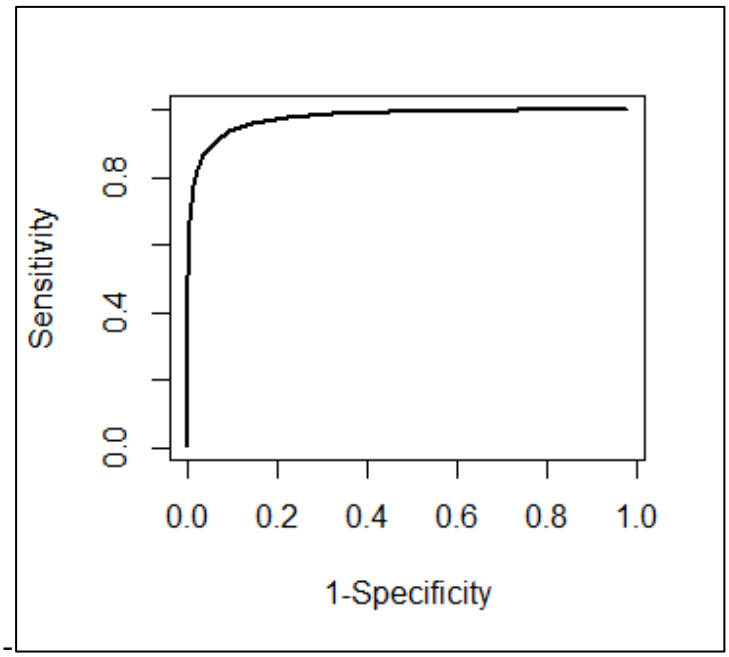


Fig 4. ROC curve for ELISA

Table 1. Sensitivity and specificity of ELISA for HT1 samples given a range of cut-off values. Values are posterior medians and the corresponding 95% probability interval.

Cut-offs	Sensitivity			Specificity		
	Median	95%PI		Median	95%PI	
0.30	0.906	0.852	0.947	0.945	0.893	0.976
0.31	0.899	0.844	0.942	0.950	0.901	0.979
0.32	0.892	0.835	0.936	0.955	0.908	0.981
0.33	0.884	0.826	0.931	0.959	0.915	0.984
0.34	0.877	0.817	0.925	0.963	0.921	0.986
0.35	0.869	0.807	0.918	0.967	0.927	0.988
0.36	0.860	0.798	0.912	0.970	0.932	0.989
0.37	0.852	0.788	0.905	0.973	0.937	0.991
0.38	0.843	0.778	0.898	0.976	0.942	0.992
0.39	0.834	0.768	0.890	0.978	0.946	0.993
0.40	0.825	0.758	0.883	0.980	0.950	0.994
0.41	0.816	0.748	0.875	0.982	0.954	0.995
0.42	0.806	0.738	0.867	0.984	0.958	0.995
0.43	0.796	0.727	0.858	0.986	0.961	0.996
0.44	0.786	0.716	0.850	0.987	0.964	0.996
0.45	0.776	0.706	0.841	0.988	0.967	0.997
0.46	0.766	0.695	0.832	0.990	0.969	0.997
0.47	0.756	0.684	0.822	0.991	0.972	0.998
0.48	0.746	0.674	0.813	0.992	0.974	0.998
0.49	0.735	0.663	0.803	0.992	0.976	0.998
0.50	0.724	0.652	0.793	0.993	0.978	0.999

Analysis of HT1 and HT4 Data

HT4 data were analysed independently and on a combined data set that consisted of HT1 and HT4 data. The same analysis steps were applied to both HT4 dataset and the combined dataset.

Comparing HT4 data to HT1 data revealed that some herds were tested on both occasions. However, for a same herd, proportions of PCR-positive animals at different herd tests were appreciably different. This was unlikely due to the inconsistent performance of PCR at different lactation stages however, it was more likely due to the change of the true prevalence of infection. The change of the true prevalence could be associated with many factors such as culling and treatment. However, this type of information was unknown and beyond the scope of the analysis.

Since the prevalence of infection for the same herd varied between herd tests, it made sense to account for this difference. One possible approach was to create a cluster ID by combining the herd ID and test ID. For example, the same herds tested at different herd tests were considered as distinct populations (distinct clusters). The other possible approach was to modify the current analytical model. Instead of modelling the true prevalence using a logit-normal distribution, we could model the proportion of PCR-positives as a function of herd test (consider it as a variable with two categories). The second approach was statistically correct (according to the data), as the proportion of PCR-positives did vary between herd tests. However, it was biologically unjustified, and it was even more difficult to come up with a proper prior for the regression coefficient representing the effect of herd test on the probability

of being PCR-positive. For example, it was difficult to know the odds of having PCR-positive was *X* times higher at HT1 compared to HT4. Therefore, the first option was preferred.

According to the density plot of the untransformed ELISA S/P ratio, the Se of PCR was expected to be higher when adding HT4 data than the result obtained based on the HT1 data, as more PCR-positive animals had a low S/P ratio.
The Sp of PCR was expected to be consistent with the result obtained based on the HT1 data as the distributions of ELISA S/P ratio in PCR-negative groups at both herd tests appeared to be similar (Fig 5).

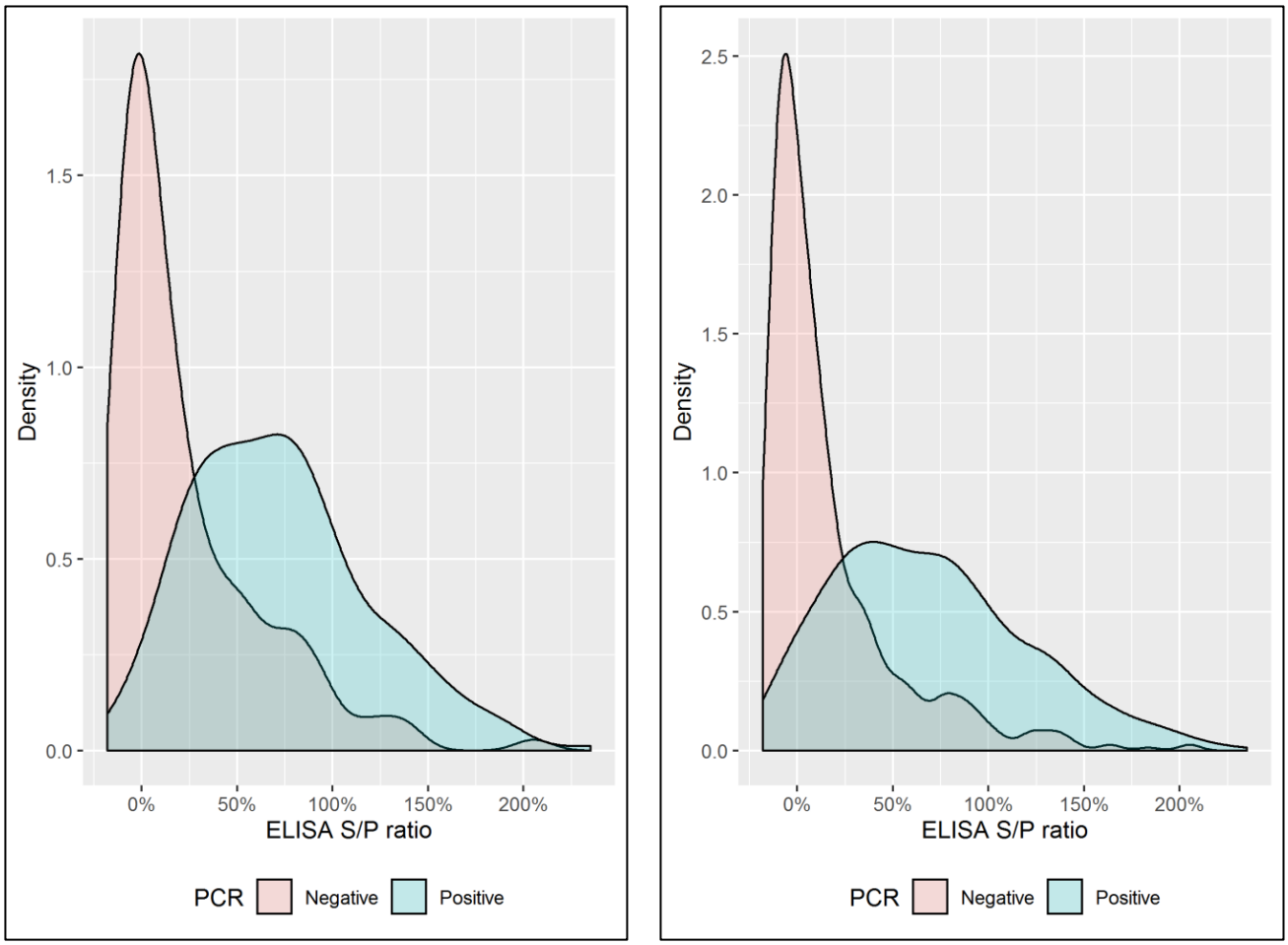


Fig 5. The distribution of ELISA S/P ratio in PCR-positive and PCR-negative animals. The left panel is based on HT1 data, the right panel is based on HT1 and HT4 data.

The ELISA performance was consistent with an AUC = 0.9867 for HT1 + HT4 and an AUC = 0.9789 for HT4.

The predicted probabilities given a range of S/P ratios remained extremely similar for all HT data. The optimal cut-offs remained in the same range – from 0.30 to 0.35 (Table 2 and Table 3).

Table 2. Sensitivity and specificity of ELISA for HT1+HT4 samples given a range of cut-off values. Values are posterior medians and the corresponding 95% probability interval.

Cut-offs	Sensitivity			Specificity		
	Median	95%PI		Median	95%PI	
0.30	0.925	0.886	0.955	0.955	0.926	0.975
0.31	0.918	0.877	0.950	0.959	0.931	0.978
0.32	0.910	0.868	0.945	0.962	0.936	0.980
0.33	0.903	0.858	0.939	0.966	0.941	0.982
0.34	0.895	0.849	0.933	0.969	0.945	0.984
0.35	0.886	0.839	0.926	0.971	0.949	0.986
0.36	0.878	0.829	0.919	0.974	0.953	0.987
0.37	0.869	0.818	0.912	0.976	0.956	0.989
0.38	0.859	0.808	0.904	0.978	0.959	0.990
0.39	0.850	0.797	0.896	0.980	0.962	0.991
0.40	0.840	0.786	0.888	0.982	0.965	0.992
0.41	0.830	0.775	0.879	0.983	0.968	0.993
0.42	0.820	0.764	0.871	0.985	0.970	0.994
0.43	0.809	0.753	0.861	0.986	0.972	0.994
0.44	0.799	0.741	0.852	0.987	0.974	0.995
0.45	0.788	0.730	0.842	0.988	0.976	0.995
0.46	0.777	0.718	0.832	0.989	0.978	0.996
0.47	0.766	0.707	0.822	0.990	0.979	0.996
0.48	0.754	0.695	0.812	0.991	0.981	0.997
0.49	0.743	0.683	0.801	0.992	0.982	0.997
0.50	0.732	0.672	0.790	0.993	0.983	0.997

Table 3. Sensitivity and specificity of ELISA for HT4 samples given a range of cut-off values. Values are posterior medians and the corresponding 95% probability interval.

Cut-offs	Sensitivity			Specificity		
	Median	95%PI		Median	95%PI	
0.30	0.908	0.854	0.948	0.943	0.917	0.964
0.31	0.899	0.843	0.941	0.947	0.922	0.966
0.32	0.888	0.830	0.934	0.950	0.926	0.969
0.33	0.878	0.818	0.926	0.953	0.930	0.971
0.34	0.867	0.805	0.917	0.956	0.934	0.973
0.35	0.856	0.792	0.908	0.959	0.937	0.975
0.36	0.845	0.779	0.899	0.961	0.940	0.977
0.37	0.833	0.766	0.889	0.963	0.943	0.978
0.38	0.821	0.753	0.879	0.966	0.946	0.980
0.39	0.808	0.740	0.869	0.968	0.949	0.981
0.40	0.796	0.727	0.858	0.969	0.951	0.982
0.41	0.784	0.713	0.847	0.971	0.954	0.984
0.42	0.771	0.700	0.836	0.973	0.956	0.985
0.43	0.758	0.687	0.825	0.974	0.958	0.986
0.44	0.745	0.673	0.813	0.976	0.960	0.987
0.45	0.733	0.660	0.801	0.977	0.962	0.987
0.46	0.720	0.647	0.789	0.978	0.963	0.988
0.47	0.707	0.634	0.777	0.979	0.965	0.989
0.48	0.694	0.622	0.765	0.980	0.966	0.989
0.49	0.682	0.609	0.753	0.981	0.968	0.990
0.50	0.669	0.597	0.741	0.982	0.969	0.991

PCR performance was obtained from the two different datasets and compared to the HT1 results. The posterior medians and their 95% probability interval were provided.

HT1

Sensitivity	0.7619 (0.6827-0.8436)
Specificity	0.8945 (0.8180-0.9627)

HT4

Sensitivity	0.8973 (0.8193 -0.9568)
Specificity	0.9814 (0.9516 -0.9960)

HT1 + HT4

Sensitivity	0.8400 (0.7701-0.9087)
Specificity	0.8978 (0.8300-0.9570)

The PCR results for HT4 and HT1 + HT4 were as expected. Se and Sp were higher in HT4 than at HT1. The Se was still higher in the combined dataset while the Sp values were nearly identical to HT1. The 95% probability intervals based on the combined dataset were narrower as more data points were used.

However, it is important to point out that the posteriors obtained from the combined data were extremely sensitive to the prior for the standard deviation of the random effect. Changing Unif (0,1) to Unif (0,2), we found Se of PCR = 0.15 and Sp of PCR = 0.20, which was unreasonable. Particularly, the model's identifiability was ensured under the condition that Se + Sp >1. Therefore, the information obtained from the combined PCR data needs to be interpreted with caution. From a statistical point of view, the results obtained based on HT1 data were much more reliable.