Evaluation of the Sensitivity and Specificity of an Enzyme-Linked Immunosorbent Assay for Diagnosing Brucellosis in African Buffalo (Syncerus caffer)

Citation	Gorsich, E. E., Bengis, R. G., Ezenwa, V. O., & Jolles, A. E. (2015). Evaluation of the Sensitivity and Specificity of an Enzyme-Linked Immunosorbent Assay for Diagnosing Brucellosis in African Buffalo (Syncerus caffer). Journal of Wildlife Diseases, 51(1), 9-18. doi:10.7589/2013-12-334
DOI	10.7589/2013-12-334
Publisher	Wildlife Disease Association
Version	Accepted Manuscript
Terms of Use	http://cdss.library.oregonstate.edu/sa-termsofuse

Evaluation of the sensitivity and specificity of an ELISA for diagnosing brucellosis in African buffalo (*Syncerus caffer)*.

Erin E. Gorsich[1,5], Roy G. Bengis[2], Vanessa O. Ezenwa[3], Anna E. Jolles[4]

1 Department of Integrative Biology, 3029 Cordley Hall, Oregon State University, Corvallis OR 97331, USA

2 Directorate of Animal Health, P.O. Box 12, Skukuza, 1350 South Africa

3 Odum School of Ecology and Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia, Athens, GA 30602, USA

4 Department of Biomedical Sciences and Department of Integrative Biology, Oregon State University, Corvallis OR 97331 USA.

5 Corresponding author (email: gorsiche@oregonstate.edu, Tel: +1-541-737-3705, fax: +1-541-737- 0501)

Running head: Gorsich et al.- Diagnosing brucellosis in African buffalo

Word Count: 3,597

Abstract

Brucellosis is a disease of veterinary and public health importance worldwide. In sub-Saharan Africa, where this disease has been detected in several free-ranging wildlife species, successful disease control may be dependent on accurate detection in wildlife reservoirs, including African buffalo (*Syncerus caffer*). This study estimates the sensitivity and specificity of a commercial enzyme-linked immunosorbent assay (IDEXX, Brucellosis Serum Ab Test) for brucellosis based on a dataset of 571 serum samples from 258 buffalo located within the Kruger National Park, South Africa. We defined a pseudo-gold standard test result as those buffalo that were consistently positive or negative on two additional serological tests, namely the rose bengal test (RBT) and the complement fixation test (CFT). The ELISA's cut-off value was selected using receiver operating characteristics (ROC) analysis, the pseudo-gold standard, and a threshold criterion that maximizes the total sensitivity and specificity. Then, we estimated the sensitivity and specificity of all three tests using Bayesian inference and latent class analysis. We estimated the ELISA to have a sensitivity of 0.928 (95% BCI from 0.869-0.974) and specificity of 0.870 (95% BCI from 0.836-0.900). Compared to the ELISA, the RBT had a higher estimated sensitivity of 0.986 (95% BCI from 0.928- 0.999), and both the RBT and CFT had higher specificities, estimated to be 0.992 (95% BCI from 0.971 to 0.996) and 0.998 (95% BCI from 0.992 to 0.999), respectively. Therefore, this study shows that no single serological test perfectly diagnosed infection. However, after adjustment of cut-off values for South African conditions, the IDEXX Brucellosis Serum Ab Test may be a valuable additional screening test for brucellosis in Kruger National Park's African buffalo.

Keywords: African buffalo, Bayesian, Brucellosis, Enzyme linked immunosorbent assay, Latent data, Sensitivity, Specificity

Introduction

Brucellosis is an important veterinary public health issue and one of the most common zoonotic diseases worldwide (McDermott and Arimi, 2002). *Brucella abortus,* the pathogenic bacteria responsible for bovine brucellosis, causes sub-acute to chronic disease in many ungulate species including African buffalo, elk, bison, eland, waterbuck, impala and cattle (Godfroid, 2002). Brucellosis transmission occurs primarily when bacteria are shed from infected animals around birthing periods. Bacteria are shed in birth products, aborted fetuses, and intermittently through unpasteurized milk (Rhyan et al., 2009). Infection is characterized by abortions, high morbidity rates, and context-dependent reductions in survival (Joly and Messier, 2005) and, as a leading cause of cattle morbidity worldwide, accurate disease detection is essential for public health (Godfroid et al., 2011). These concerns have motivated successful 'test-and-slaughter' programs in industrialized countries that have virtually eliminated the disease except in areas adjacent to wildlife reservoirs. Research efforts aimed at understanding infection in wildlife and minimizing transmission between wildlife and livestock are essential for disease management (Kilpatrick et al., 2009; Gomo et al., 2012). As such, the development and evaluation of reliable diagnostic tests for brucellosis in wildlife is a priority.

Brucellosis has been maintained endemically in African buffalo (*Syncerus caffer)* in Kruger National Park (KNP), South Africa (Chapparo et al., 1990), since its speculated introduction from European cattle (Gradwell et al., 1977). In African buffalo, diagnosis is based on three indirect diagnostic tests that measure the host's antibody response rather than the presence of *B. abortus* organisms: the rose bengal test (RBT), the complement fixation test (CFT), and the serum agglutination test (SAT; Herr and Marshall, 1981; Chapparo et al., 1990). We restrict our analysis to those tests routinely used in African buffalo (Chapparo et al., 1990)

70 although additional diagnostic tests have been used for brucellosis testing in cattle and American
71 bison (e.g. Gall et al., 2000). Information on antibody responses to *B. abortus* infection have
72 been determined from experimental infections in cattle (Nielsen et al., 1984). The Serum
73 Agglutination Test (SAT) was one of the first serological tests for brucellosis and is based
74 primarily on IgM antibodies because they are the most active agglutinins (Nielsen, 2002). This
75 test causes many false positives and has been discontinued by the World Organization for
76 Animal Health (OIE, 2008a). The RBT and CFT are often used in combination for accurate
77 diagnosis, with the RBT used as a screening test and the CFT used as a confirmatory test.
78 However, application of the CFT requires precise measurements and specialized reagents,
79 making it difficult to implement under field conditions. As a result, it is being replaced by
80 ELISA diagnostic tests (Godfroid et al., 2010). All three tests (RBT, CFT, ELISA) are
81 recommended by the OIE as valuable livestock diagnostic tests (OIE, 2008a), but the direct
82 application of these tests from cattle populations to African buffalo populations is problematic.
83 This is because test sensitivity and specificity will vary among species, and none of these tests
84 has been validated in African buffalo.

85 Traditional estimates of diagnostic test sensitivity and specificity are based on direct
86 comparisons against an established gold standard test (detection of *Brucella* organisms by culture
87 methods). Because true gold standard test results are often costly or impractical to obtain,
88 especially in wildlife systems, a new test's accuracy is commonly estimated by comparing it to a
89 reference test with a known error rate (Buck and Gart, 1966) or by comparison to multiple
90 imperfect diagnostic tests (Enoe et al., 2000). Techniques that estimate test accuracy or disease
91 prevalence when there is uncertainty in the test's sensitivity or specificity are called latent class
92 analyses because they use the observed frequency of diagnostic test results to estimate a latent

93 variable, the true disease status, from which the new diagnostic test can be evaluated (Branscum
94 et al., 2005). Accurate estimates of a test's sensitivity and specificity with latent class analysis
95 requires correctly representing whether the outcomes of two tests for a given animal are
96 independent or correlated (conditional upon the true state of the animal; Georgiadis et al., 2003).
97 Therefore, we consider potential correlations among tests in this analysis. This analysis also
98 follows an increasing trend in the use of a Bayesian inference with latent class analysis;
99 examples include the estimation of test accuracy for Foot and Mouth disease (Engel et al., 2008),
100 tuberculosis (Alvarez et al., 2012), and brucellosis in cattle (Matope et al., 2011). Bayesian
101 inference could also be useful for diagnostic test evaluation in wildlife because it incorporates
102 uncertainty about model parameters based on independently collected, or prior information.
103 These techniques are recommended by the OIE to estimate sensitivity and specificity, but
104 represent only one step in the validation process (OIE, 2008b). The assumptions and
105 modifications used in latent class analyses have been reviewed in general for latent class
106 techniques (Enoe et al., 2000) and more specifically for latent class techniques with Bayesian
107 inference (Branscum et al., 2005).

108 This paper aims to evaluate the utility of an ELISA (IDEXX Brucellosis Serum Ab Test)
109 for diagnosis of brucellosis in an important wildlife host, African buffalo. First, we selected an
110 ELISA cut-off value based on a pseudo-gold standard created from a subset of sampled buffalo
111 that consistently tested seropositive or seronegative on both the RBT and CFT. Second, we used
112 latent class modeling to estimate the sensitivity and specificity of the ELISA, RBT, and CFT
113 based on the entire dataset of diagnostic test results.

114 **Materials and Methods**

115 *Animal captures and test methods*

116 Serum samples were collected from a cohort of 202 female buffalo from herds in two
117 areas of southern Kruger National Park, South Africa, the Lower Sabie and the Crocodile Bridge
118 area. Buffalo were captured approximately every six months between 2008 and 2010 as part of
119 an ongoing disease study. Fifty-two animals died throughout the study period and were replaced
120 with additional buffalo, resulting in 571 samples collected from 254 buffalo. No buffalo were
121 sampled less than six months apart. We collected samples for diagnostic test evaluation between
122 June 2008 and August 2009 and again between March and October 2010. All buffalo captured
123 in those periods were tested with each diagnostic test. Animals were chemically immobilized by
124 research veterinarians and South African National Parks (SANParks) staff with M99 (etorphine
125 hydrochloride) and ketamine. Jugular blood was collected from each animal into blood tubes
126 and immediately stored on ice in a cooler for transportation back to the laboratory. The blood
127 was centrifuged at 6,000 g for 10 minutes and sera samples were separated and stored at -20°C
128 for subsequent disease testing. Animal capture and data collection protocols were approved by
129 Oregon State University, University of Georgia, and SANParks' Institutional Animal Care and
130 Use Committees.

131 We used three serological measures of brucellosis infection. The rose bengal test (RBT)
132 and complement fixation test (CFT) were conducted by the Onderstepoort Veterinary Institute's
133 diagnostic laboratories in South Africa according to OIE specifications (OIE, 2008a). Briefly,
134 the RBT is conducted by monitoring the agglutination response after mixing serum with rose
135 bengal stained *B. abortus* cells. The CFT is conducted by monitoring the degree of haemolysis
136 after incubating inactivated test serum, antigen, and exogenous complement with sensitized
137 sheep red blood cells (OIE, 2008a). The Brucellosis Serum Ab ELISA tests (IDEXX P04130)

were conducted in the field laboratory at KNP according to kit instructions. This assay detects antibodies to the lipopolysaccharide (LPS) antigen of smooth *Brucella* strains. Test results are determined by a sample's optical density (OD) read at 450nm and compared to the positive and negative controls according to this equation:

Cut-off%= [100× μ(OD450 of the paired sample wells)- μ(OD450 of negative control wells)]/ [μ(OD450 of positive control wells) - μ(OD450 of negative control wells)].

The cut-off value for determining seropositivity in cattle recommended by IDEXX is 120%. However, we explored test sensitivity and specificity at additional cut off values because as we were testing sera from a different species.

Selection of ELISA cut off values with ROC curve analysis

To select ELISA cut-off values, we defined a pseudo-gold standard that estimates true disease seroprevalence. We combined the results from the CFT and RBT into a composite reference standard (Alonzo and Pepe, 1999). Buffalo were identified as seropositive only if they remained both RBT and CFT positive over a six month period, and seronegative only if they remained negative on both tests over a six month period. Of the 254 individuals tested with all three diagnostic tests, 153 buffalo were sampled twice during a consecutive 6-month period and returned concordant test results using the RBT and CFT tests. The ELISA's test results at the end of the time period were compared to this pseudo-gold standard.

We used receiver operating characteristic (ROC) curves to select the ELISA's cut off value and two-graph receiver operating characteristic curves to display the relationship between sensitivity and specificity for various cut-off values. (TG-ROC; Gardner and Greiner, 2006). Selection of test cut-off values remains dependent on the intended use of the test, which may

160 vary for different decision-making situations (e.g. test-and-cull programs vs. surveillance). For
161 example, lower cut-off values may be advisable when there are consequences for false negative
162 test results while higher cut-off values may be preferred when there are high costs for false
163 positive test results (Greiner et al., 2000). We report the cut-off value that maximizes the total
164 sensitivity and specificity (Se+Sp). ROC analysis and the TG-ROC plot were conducted with
165 the package, DiagnosisMed (Brasil, 2010) for R statistical software (R Core Team, 2012).
166 Clopper-Pearson binomial confidence intervals were drawn for test accuracy in the ROC curve
167 analysis (Brasil, 2010). Because estimates from the pseudo-gold standard analysis only include a
168 subset of the animals with concordant test results on the RBT and CFT tests, the analysis may
169 overestimate ELISA test accuracy. This could occur if the reduced dataset excludes animals
170 with lower antibody responses or animals that became infected during the study. Thus, we used
171 latent class models to estimate ELISA sensitivity and specificity from the test results of all
172 collected samples.

173 *Latent class analysis and prior estimation*

174 Latent class analysis allows evaluation of diagnostic tests in the absence of a gold
175 standard. The simplest model presented here assumes that the outcomes of the tests for a given
176 animal are independent, conditional upon the true state of the animal. This model is referred to
177 as the conditional independence model and is described in detail in the appendix. A complete
178 model specification and review of Bayesian approaches to estimation can be found in Branscum
179 et al (2005); the models initial descriptions in two and three populations can be found in Hui and
180 Walter (1980) and Walter and Irwig (1988), respectively. The validity of assuming two tests are
181 conditionally independent requires further justification (Vacek, 1985). The results of diagnostic
182 tests that measure similar biological processes are likely to be correlated (conditional on the

183 animals true disease status; Gardner et al., 2000) and assuming independence may result in
184 incorrect estimates of test accuracy (Georgiadis et al., 2003). The RBT, CFT, and ELISA all
185 measure the hosts' antibody response to *Brucella* smooth LPS, but they use different methods of
186 antibody detection (Godfroid et al., 2010; Nielsen, 2002). Therefore, because we had little prior
187 knowledge about the potential correlation between test outcomes, we consider models assuming
188 both conditional independence and conditional dependence.

189 We used model selection based on Deviance Information Criteria (DIC) to compare the
190 fit of models assuming conditional independence and conditional dependence (e.g. Rahman et
191 al., 2013). DIC is a model assessment tool based on model fit and the effective number of
192 parameters (Link and Barker, 2010). Models with lower DIC values provide a better fit to the
193 data, and we chose the model with the lowest DIC value (Spiegelhalter et al., 2002). Prior
194 distributions for diagnostic test sensitivity and specificity were represented as beta distributions
195 and were defined using published results from test validations in cattle (Grenier et al., 2009;
196 Table 1). The prior distributions for each parameter are displayed in Table 1 and details of their
197 specification are given in the appendix. This prior information was combined with the full
198 dataset of 571 samples (Supplement 2). Median and 95% Bayesian credible intervals are
199 presented for all parameters in the best fitting model. We conducted sensitivity analyses on this
200 model by (1) increasing the mode and lower bound of the each tests' sensitivity and specificity
201 prior distributions by 5 percentage points, (2) decreasing the mode and lower bound of the each
202 tests' sensitivity and specificity prior distributions by 5 percentage points, and (3) by specifying
203 uninformative priors between the interval of zero to one, modeled as Beta (1,1), for each tests'
204 sensitivity and specificity parameter. We also compared estimates generated from models fit

with only the first sample point for each of the 254 buffalo sampled to explore if the pseudoreplication in our dataset influenced the estimates of test accuracy.

Results

Selection of ELISA cut-off values with ROC curve analysis

The pseudo-gold standard defined 28 positive and 123 negative animals (Table 2). Within this subset of buffalo, the sensitivity and specificity estimates when using the kit's defined cut-off of 120% were 1 (95% confidence interval from 0.82 to1.00) and 0.87 (95% confidence interval from 0.80 to 0.92), respectively. The ROC curve analysis shows that ELISA specificity was improved at higher cut-off values with minimal reductions in sensitivity. The cut-off value with the highest sensitivity and specificity (Se+Sp) was 159% (Figure 1). This cut-off is associated with a sensitivity of 1 (95% confidence interval from 0.82 to 1) and a specificity of 0.93 (95% confidence interval from 0.87 to 0.97).

Latent class analysis:

The diagnostic test results used for latent class models were calculated based on the ELISA cut-off value of 159% and all 571 samples (Table 3). The model assuming conditional dependence between the ELISA and CFT had the lowest DIC value (DIC= 59.24). Neither the model assuming conditional independence (DIC= 63.24) nor the models with additional dependence parameters had lower DIC values (Supplement Table S1). We, therefore, report the results of this model based on parsimony and model fit.

Test accuracy varied among the diagnostic tests (Figure 2). The ELISA's sensitivity and specificity were estimated to be, Se= 0.928 (95% BCI from 0.869-0.974) and Sp= 0.870 (95% BCI from 0.836-0.900). The RBT had the highest estimated sensitivity, (Se= 0.986, 95% BCI

from 0.928- 0.999), and both the ELISA and RBT had significantly higher sensitivities than the CFT (Se=0.374, 95% BCI from 0.294-0.460). However, both the RBT and CFT had significantly higher specificities than the ELISAs, with estimated values of 0.992 (95% BCI from 0.971 to 0.996) and 0.998 (95% BCI from 0.992 to 0.999), respectively. Prevalence in the Lower Sabie region was estimated as 0.235 (95% BCI from 0.183 to 0.292) and in the Crocodile Bridge region as 0.228 (95% BCI from 0.183 to 0.277).

Sensitivity analyses showed that decreasing the mode of the ELISA prior distribution by 5 decreased the median of the posterior distributions from sensitivity= 0.928 (95% BCI from 0.869-0.974) to 0.925 (95% BCI from 0.867- 0.971) and from specificity 0.870 (95% BCI from 0.836- 0.900) to 0.869 (95% BCI from 0.835- 0.900), with similar results when prior information was also relaxed to 70% (Table 4). Increasing the mode of the ELISA prior distributions by 5 resulted in only a minor increase to sensitivity= 0.930 (95% BCI from 0.871- 0.976) and specificity=0.870 (95% BCI from 0.836-0.901). The estimates of ELISA accuracy also remained similar when the prior values for RBT and CFT accuracy were relaxed (Table 4). When the model was fit to data with one test result per buffalo, test specificity remained similar but test sensitivity increased slightly to 0.960 (0.887- 0.993). The 95% credible intervals overlap despite these perturbations, suggesting that the estimates of ELISA sensitivity and specificity were influenced by the frequency of test results and, to a lesser extent, the prior information.

Discussion

The IDEXX ELISA was estimated to have a sensitivity of Se=0.928 (95% BCI from 0.869-0.974) and specificity of Sp= 0.870 (95% BCI from 0.836-0.900 when using the cut-off value of 159%. At this cut-off value, the results show that the ELISA has a higher median

249 sensitivity than the CFT, similar but lower sensitivity to the RBT, but a lower specificity than
250 both the RBT and CFT. The estimates of test accuracy in this study are based on the selected
251 ELISA cut-off value. The cut-off value that maximized the total sensitivity and specificity
252 (Se+Sp) was 159%. Because test sensitivity and specificity are inversely related at a given cut-
253 off value, a different cut-off would result in altered estimates of test accuracy. The selected cut-
254 off value should be taken into account when comparing diagnostic tests (Greiner et al., 2000).
255 For example, at the suggested cut-off value for cattle, 120%, the ELISA had a lower specificity
256 and a higher sensitivity. This result emphasizes the importance of test optimization for each
257 population and species to which it is applied.

258 In addition to species-specific differences, there are three nonexclusive factors that
259 explain why the cut-off value for cattle resulted in a higher number of miss-classified results.
260 First, the test is being applied under field laboratory conditions. Serum samples for these
261 analyses were collected and frozen in the field at -20°C for one to three years, with temperature
262 fluctuations possible due to a somewhat variable power supply (though to our knowledge no
263 outright freezer failure occurred during the storage period of these samples). Ideally, sample
264 storage would use consistent and colder (-80C) temperatures; as such, sub-optimal storage
265 conditions might have degraded the samples to some degree. Second, brucellosis is endemic in
266 this buffalo population and our sampling may have resulted in animals with a wider range of
267 times since infection than those used for the tests' validation in cattle. Finally, all diagnostic
268 tests are susceptible to cross-reactive antibodies. *Yersinia enterocolitica* O:9 shares common
269 antigenic epitopes with *B. abortus,* and is known to cross-react during diagnosis, but little is
270 known about *Yersinia*'s presence in buffalo populations (Godfroid et al., 2002). The evaluation

271 presented here allows these sources of variability to be incorporated into the estimates of test
272 accuracy, allowing the estimates to be robust to problems inherent in most field conditions.
273 Latent class analysis allows the quantification of test variability and accuracy in the
274 absence of a gold standard. Like all model-based analyses, their implementation involves a
275 tradeoff between the model complexity (number of parameters) and parsimony. The model
276 selection performed in this study shows that models including co-variance between the ELISA
277 and CFT had a better fit to the data compared to the model assuming conditional independence.
278 The lack of support for models representing dependence between the ELISA/ RBT and the RBT/
279 CFT, based on DIC values, is supported by the low conditional correlations among those tests.
280 Because the tests measure antibodies through different mechanisms (Nielsen, 2002), it is
281 plausible that the tests are conditionally independent of each other, given the true state of the
282 animal. However, those models also may have had higher DIC values because there was
283 minimal data to estimate the conditional dependence terms; there were few samples with ELISA-
284 , RBT- and CFT+ test results or ELISA-, RBT+, CFT+ test results. Previous work on brucellosis
285 in sheep represented conditional dependence between the RBT and the ELISA and between the
286 RBT and SAT (Rahman et al., 2013). Other systems, however, have found the conditional
287 independence model to be most appropriate (Muma et al., 2007, Rahman et al., 2013). The
288 results of this analysis show similar estimates of sensitivity and specificity in all models
289 regardless of the test correlation assumptions (Table S1) and suggests that these estimates were
290 robust to model assumptions.

291 The uncertainty in how any of the diagnostic tests relate to active infection in wildlife
292 represents a major hurdle to accurate diagnostic methods (Treanor et al., 2011). Owing to these
293 limitations, the results of this evaluation serve as a comparison among the serological tests

294 historically used. Additional assays for brucellosis, including the FPA (Gall et al., 2000), PCR
295 (Bricker, 2002), and alternative ELISA techniques (Nielsen, 2002) have shown improved
296 accuracy in other systems and should be considered for future testing in African buffalo.
297 Further, the estimates of test sensitivity and specificity presented in this analysis includes prior
298 information (Figure 2). Rather than a limitation, incorporating this information could be a
299 valuable tool for wildlife studies given the sample size requirements and potential identifiability
300 problems with latent class models (Dendukuri et al., 2010). Our analysis also assumes that test
301 sensitivity and specificity are consistent throughout the course of brucellosis infection and
302 between populations. As more information develops about the course of brucellosis in buffalo,
303 future diagnostic tests evaluations should incorporate variation in detection rates between
304 different stages of infection (Engel et al., 2010; Caraguel et al., 2012) or different populations
305 (Munoz et al., 2012).

306 The benefits of the ELISA are that it is relatively inexpensive, easy to perform in field
307 conditions, and results in a quantitative test result. The choice of an appropriate diagnostic test,
308 however, is dependent on its intended use. For example, with a specificity of 87%, the ELISA
309 may not present an ideal diagnostic tool for screening of commercial buffalo herds because it
310 would result in many false-positive animals being removed at an undesirably high cost to the
311 farmer. However, its use in combination with the RBT could improve current diagnostic
312 methods by avoiding misclassifications. For large-scale disease surveys, the ELISA's 93%
313 sensitivity and ease of use may make it a valuable screening tool for African buffalo. Given the
314 importance of brucellosis for public health in sub-Saharan Africa, further work establishing and
315 validating improved diagnostic methods is needed for detection of *B. abortus* in one of its
316 wildlife reservoirs, the African buffalo.

Acknowledgements

Buffalo capture was conducted by Veterinary Wildlife Services department of KNP, Rob Spaan, Jo Spaan, Kristie Thompson and Brianna Beechler. We would like to thank Brianna Beechler and Jo Spaan for additional data collection and Jennifer Joubert and Nomkhosi Mathebula for advice on laboratory techniques. We are also thankful to Dr. Louis van Schalkwyk, Dr. Nels Johnson, Dr. Adam Branscum, and Nick Greenfield for advice and discussions about this manuscript. This research was approved by Kruger National Park's Scientific Services committee and by Oregon State University and University of Georgia IACUC (Protocol numbers OSU: 3822; UGA: A201010-190-A1). This research was supported by the National Science Foundation through a NSF Ecology of Infectious Diseases award to V. Ezenwa and A. Jolles (EF- 0723918, DEB-1102493/EF-0723928), a NSF Graduate Research Fellowship to E. Gorsich, and an NSF Doctoral Dissertation Improvement award to E. Gorsich (DEB-121094).

Literature Cited

Albert PS, Dodd LE. 2004. A Cautionary Note on the Robustness of Latent Class Models for Estimating Diagnostic Error without a Gold Standard. *Biometrics* 60:427–435.

Alonzo TA, Pepe MS. 1999. Using a combination of reference tests to assess the accuracy of a new diagnostic test. *Stat Med* 18:2987–3003.

Alvarez J, Perez A, Bezos J, Marques S, Grau A, Saez JL, Minguez O, de Juan, L. 2012. Evaluation of the sensitivity and specificity of bovine tuberculosis diagnostic tests in naturally infected cattle herds using a Bayesian approach. *Vet Microbiol* 155:38–43.

Berkvens D, Speybroeck N, Praet N, Adel A, Lesaffre E. 2006. Estimating disease prevalence in a Bayesian framework using probabilistic constraints. *Epidemiology 17*:145–153.

Branscum AJ, Gardner IA, Johnson WO. 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Prev Vet Med* 68:145–163.

Brasil P. 2010. DiagnosisMed: Diagnostic test accuracy evaluation for medical professionals. R Package Version 02.3, http://CRAN.R-project.org/src/contrib/Archive/DiagnosisMed/

Bricker BJ. 2002. PCR as a diagnostic tool for brucellosis. *Vet Microbiol* 90:435–446.

Buck AA, Gart JJ. 1966. Comparison of a screening test and a reference test in epidemiological studies. I. indices of agreement and their relation to prevalence. *Am J Epidemiol* 83:586–592.

Caraguel C, Stryhn H, Gagne N, Dohoo I, Hammell L. 2012. Use of a third class in latent class modelling for the diagnostic evaluation of five infectious salmon anemia virus detection tests. *Prev Vet Med* 104:165–173.

Chapparo F, Lawrence JV, Bengis R, Myburgh JG. 1990. A serological survey for brucellosis in buffalo (*Syncerus caffer*) in the Kruger National Park. *J S Afr Vet Assoc* 61:110–111.

Dendukuri N, Belisle P, Joseph L. 2010. Bayesian sample size for diagnostic test studies in the absence of a gold standard: Comparing identifiable with non-identifiable models. S*tat Med* 29:2688–2697.

Engel B, Buist W, Orsel K, Dekker A, de Clercq K, Grazioli S, van Roermund H. 2008. A Bayesian evaluation of six diagnostic tests for foot-and-mouth disease for vaccinated and non-vaccinated cattle. *Prev Vet Med* 86:124–138.

Engel B, Backer J, Buist W. 2010. Evaluation of the Accuracy of Diagnostic Tests From Repeated Measurements Without a Gold Standard. *J Agric Biol Environ Stat* 15:83–100.

Enoe C, Georgiadis MP, Johnson WO. 2000. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med* 45:61–81.

Gall D, Nielsen K, Forbes L, Davis D, Elzer P, Olsen S, Balsevicius S, Kelly L, Smith P, Tan S, Joly D. 2000. Validation of the fluorescence polarization assay and comparison to other serological assays for the detection of serum antibodies to *Brucella abortus* in bison. *J Wildl Dis* 36:469–476.

Gardner IA, Greiner M. 2006. Receiver-operating characteristic curves and likelihood ratios: improvements over traditional methods for the evaluation and application of veterinary clinical pathology tests. *Vet Clin Pathol 35:*8–17.

371 Gardner IA, Stryhn H, Lind P, Collins MT. 2000. Conditional dependence between tests affects
372 the diagnosis and surveillance of animal diseases. *Prev Vet Med* 45:107–122.

373 Gelman A, Rubin DB. 1992. Inference from Iterative Simulation Using Multiple Sequences. *Stat*
374 *Scicence* 7:457–511.

375 Georgiadis MP, Johnson WO, Ramanpreet S, Gardner IA. 2003. Correlation-adjusted estimation
376 of sensitivity and specificity of two diagnostic tests. *J Appl Stat* 52:63–76.

377 Godfroid J. 2002. Brucellosis in wildlife. *Rev Sci Tech Off Int Epizoot 21:*277–286.

378 Godfroid J, Saegerman C, Wellemans V, Walravens K, Letesson J, Tibor A, McMillan A,
379 Spencer S, Sanna M, Bakker D, Pouillot R, Garin-Bastuji B. 2002. How to substantiate
380 eradication of bovine brucellosis when aspecific serological reactions occur in the course of
381 brucellosis testing. *Vet Microbiol* 90:461–477.

382 Godfroid J, Nielsen K, Saegerman C. 2010. Diagnosis of Brucellosis in Livestock and Wildlife.
383 *Croat Med J* 51:296–305.

384 Godfroid J, Scholz HC, Barbier T, Nicolas C, Wattiau P, Fretin D, Whatmore AM, Cloeckaert A,
385 Blasco JM, Moriyon I, Saegerman C, Muma JB, Al Dahouk S, Neubauer H, Letesson JJ. 2011.
386 Brucellosis at the animal/ecosystem/human interface at the beginning of the 21st century. *Prev*
387 *Vet Med 102:*118–131.

388 Gomo C, de Garine-Wichatitsky M, Caron A, Pfukenyi DM. 2012. Survey of brucellosis at the
389 wildlife-livestock interface on the Zimbabwean side of the Great Limpopo Transfrontier
390 Conservation Area. *Trop Anim Health Prod* 44:77–85.

391 Gradwell DV, Schutte AP, Vanniekerk CA, Roux DJ. 1977. Isolation of *Brucella abortus*
392 biotype 1 from African buffalo in Kruger National Park. *J S Afr Vet Assoc* 48:41–43.

393 Greiner M, Pfeiffer D, Smith RD. 2000. Principles and practical applications of receiver-
394 operating characteristic analysis for diagnostic tests. *Prev Vet Med* 45:23–41.

395 Grenier M, Verloo D, de Massis F. 2009. Meta-analytical equivalence studies on diagnostic tests
396 for bovine brucellosis allowing assessment of a test against a group of comparative tests. *Prev*
397 *Vet Med* 92:373–381.

398 Herr S, Marshall C. 1981. Brucellosis in free-living African buffalo (*Syncerus caffer*)- A
399 serological survey. *Onderstepoort J Vet Res* 48:133–134.

400 Hui SL, Walter SD. 1980. Estimating the error rates of diagnostic tests. *Biometrics* 36:167–171.

401 Joly DO, Messier F. 2005. The effect of bovine tuberculosis and brucellosis on reproduction and
402 survival of wood bison in Wood Buffalo National Park. *J Anim Ecol* 74:543–551.

403 Kilpatrick AM, Gillin CM, Daszak P. 2009. Wildlife-livestock conflict: the risk of pathogen
404 transmission from bison to cattle outside Yellowstone National Park. *J Appl Ecol* 46:476–485.

405 Link WA, Barker RJ. 2010. *Bayesian inference with ecological applications*. Academic Press,
406 London.

407 Matope G, Muma JB, Toft N, Gori E, Lund A, Nielsen K, Skjerve E. 2011. Evaluation of
408 sensitivity and specificity of RBT, c-ELISA and fluorescence polarisation assay for diagnosis of
409 brucellosis in cattle using latent class analysis. *Vet Immunol Immunopathol* 141:58–63.

McDermott JJ, Arimi SM. 2002. Brucellosis in sub-Saharan Africa: epidemiology, control and impact. *Vet Microbiol* 90:111–134.

Menten J, Boelaert M, Lesaffre E. 2008. Bayesian latent class models with conditionally dependent diagnostic tests: A case study. *Stat Med* 27:4469–4488.

Muma JB, Toft N, Olotya J, Lund A, Nielsen K, Samul K, Skjerve E. 2007. Evaluation of three serological tests for brucellosis in naturally infected cattle using latent class analysis. *Vet Microbiol* 125:187–192.

Munoz PM, Blasco JM, Engel B, de Miguel, MJ, Marin CM, Dieste L, Mainar-Jaime RC. 2012. Assessment of performance of selected serological tests for diagnosing brucellosis in pigs. *Vet Immunol Immunopathol* 146:150–158.

Nielsen K. 2002. Diagnosis of brucellosis by serology. *Vet Microbiol* 90:447–459.

Nielsen K, Heck F, Wagner G, Stiller J, Rosenbaum B, Pugh R, Flores E. 1984. Comparative assesment of antibody isotypes to *Brucella abortus* by primary and secondary binding assays. *Prev Vet Med* 2:197–204.

OIE. 2008a. Manual of Diagnostic Tests and Vaccines for Terrestrial Animals, Paris.

OIE. 2008b. Principles and methods of validation of diagnostic assays for infectious diseases, Paris.

Plummer M, Best N, Cowles K, Vines K, Sarkar D, Almond R. 2006. CODA: Output analysis and diagnostics for MCMC. *R News* 6:7–11.

R Core Team. 2012. R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing).

Rahman, AKMA, Saegerman C, Berkvens D, Fretin D, Gani MO, Ershaduzzaman M, Ahmed MU, Emmanuel A. 2013. Bayesian estimation of true prevalence, sensitivity and specificity of indirect ELISA, Rose Bengal Test and Slow Agglutination Test for the diagnosis of brucellosis in sheep and goats in Bangladesh. *Prev Vet Med* 110:242–252.

Rhyan JC, Aune K, Roffe T, Ewalt D, Hennager S, Gidlewski T, Olsen S, Clarke R. 2009 Pathogenesis and epidemiology of brucellosis in Yellowstone bison: serologic and culture results from adult females and their progeny. *J Wildl Dis* 45:729–739.

Van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KGM, de Groot JAH. 2014. Latent class models in diagnostic studies when there is no reference standard--a systematic review. *Am. J Epidemiol* 179:423–431.

Spiegelhalter DJ, Thomas A, Best NG. 2000. WinBUGS Version 1.3 User Manual. Cambridge: Medical Research Council Biostatistics Unit. http://www.mrc-bsu.cam.ac.uk/bugs

Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A. 2002. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B* 64:583–616.

Sturtz S, Ligges U, Gelman A. 2005. R2WinBUGS: A Package for Running WinBUGS. *R J Stat Softw* 12:1–16.

Toft N, Akerstedt J, Tharaldsen J, Hopp P. 2007. Evaluation of three serological tests for diagnosis of Maedi-Visna virus infection using latent class analysis. *Vet Microbiol* 120: 77–86.

449 Treanor JJ, Geremia C, Crowley PH, Cox JJ, White PJ, Wallen RL, Blanton DW. 2011.
450 Estimating probabilities of active brucellosis infection in Yellowstone bison through quantitative
451 serology and tissue culture. *J Appl Ecol* 48:1324–1332.

452 Vacek PM. 1985. The effect of conditional dependence on the evaluation of diagnostic tests.
453 *Biometrics* 41:959–968.

454 Walter SD, Irwig LM. 1988. Estimation of test error rates, disease prevalence and relative risk
455 from misclassified data: a review. *J Clin Epidemiol* 41:923–937.

456

Tables

Table 1. Prior distributions for the ELISA, rose bengal test (RBT), and complement fixation test (CFT), and the literature from which they were estimated. Prior distributions were represented as beta distributions and estimated by defining the mode and lower confidence bounds based on estimates in the literature. Population prevalence was defined for buffalo populations in the Lower Sabie region (LS) and the Crocodile Bridge region (CB). Prior values are given for each tests' sensitivity (Sens) and specificity (Spec).

Parameters		Mode	Lower Limit	Beta (a, b)	Source
ELISA>159	Sens	0.976	> 0.60	6.29, 1.13	a
	Spec	0.975	> 0.60	6.31, 1.14	a
RBT	Sens	0.981	> 0.21	1.94, 1.02	a, b
	Spec	0.998	> 0.688	8.08, 1.01	a, b
CFT	Sens	0.960	> 0.23	2.08, 1.05	a, b
	Spec	0.998	> 0.306	2.56, 1.00	a, b
Prevalence	LS	0.30	>0.10	2.35, 4.14	c
	CB	0.35	>0.10	1.96, 2.78	c

a. Grenier et al., 2009, b. Nielsen, 2002, c. Chapparo et al., 1990

Table 2. Pseudo-gold standard test result frequencies and 95 % confidence estimates of test accuracy. Results were calculated with the ELISA cut-off value recommended for cattle (cut-off= 120) and the cut-off value selected based on receiver operating characteristic analysis (cut-off= 159). Test accuracy was improved with a higher cut-off value.

	ELISA Cut-off > 120		ELISA Cut-off > 159	
Pseudo-gold standard status	**Positive**	**Negative**	**Positive**	**Negative**
Positive	**28**	**0**	**28**	**0**
Negative	**16**	**107**	**9**	**114**
# Misclassified/ Accuracy	**16/ 89.4%**		**9/ 94.0%**	
Sensitivity	**1 (0.82-1.00)**		**1 (0.82-1.00)**	
Specificity	**0.87 (0.80-0.92)**		**0.93 (0.87- 0.97)**	

471 **Table 3.** ELISA, rose bengal test (RBT), and complement fixation test (CFT) results classified

472 for the Lower Sabie region (LS) and the Crocodile Bridge (CB) region.

ELISA/RBT/CFT	+ / + / +	+ / + / -	+ / - / +	+ / - / -	- / + / +	- / + / -	- / - / +	- / - / -
Lower Sabie	**24**	**28**	**0**	**19**	**0**	**3**	**0**	**161**
Crocodile Bridge	**21**	**47**	**0**	**39**	**3**	**6**	**0**	**220**
Total	**45**	**75**	**0**	**58**	**3**	**9**	**0**	**381**

473

474 **Table 4.** Sensitivity analyses of prior information and model assumptions. Results include the
475 consequence of adjusting prior information about each tests' accuracy and re-fitting the model to
476 a subset of the samples were each of the 258 buffalo are represented once. In analyses adjusting
477 test accuracy, the mode and lower bound were increased/decreased by 5 percentage points.

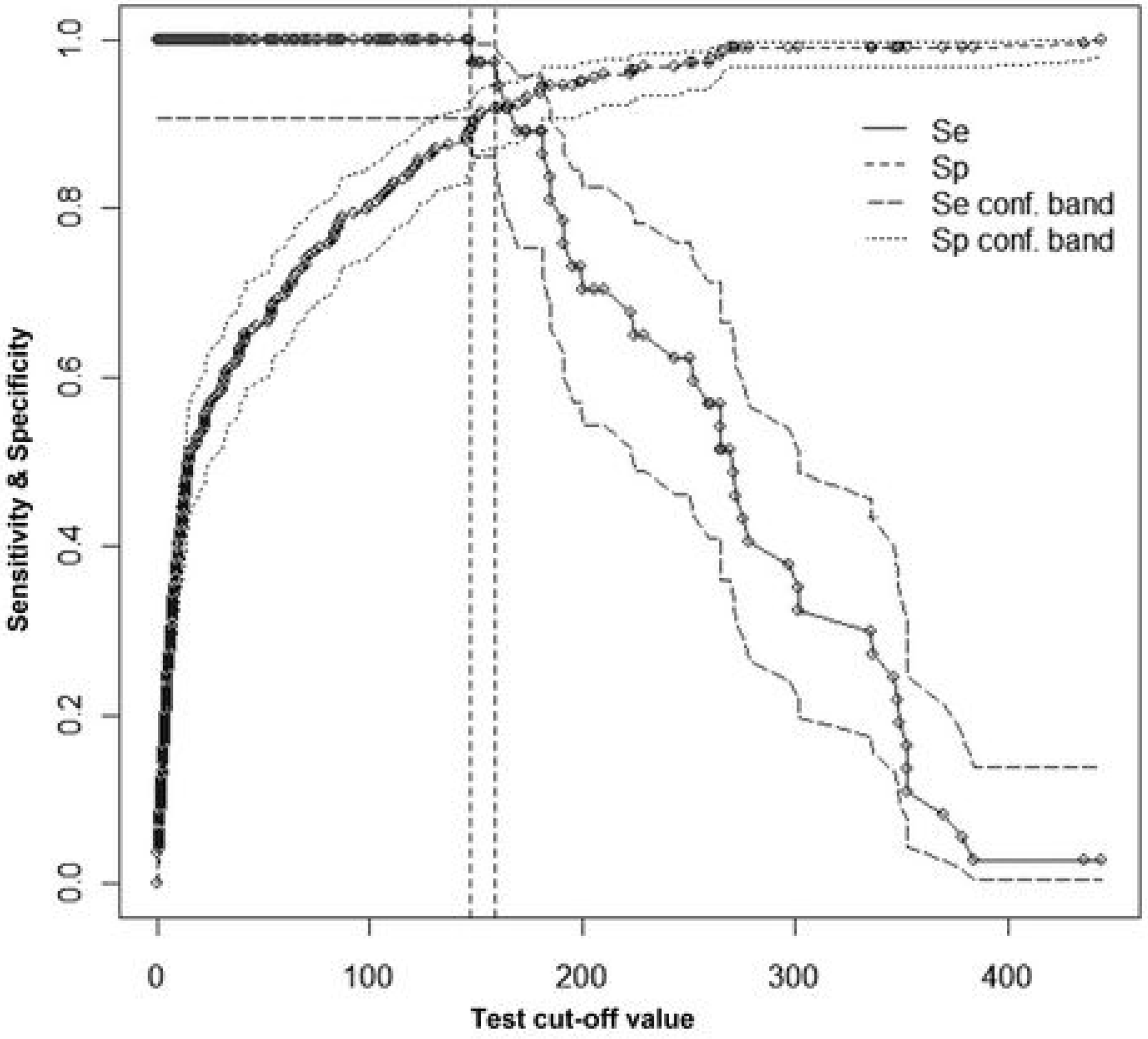
Model Specification	**ELISA Se (95%Crl)**	**ELISA Sp (95%Crl)**
CD between ELISA & CFT	0.928 (0.869- 0.974)	0.870 (0.836- 0.900)
Priors decreased by 5		
ELISA	0.925 (0.867- 0.971)	0.869 (0.835- 0.900)
RBT	0.933 (0.873- 0.977)	0.870 (0.836- 0.900)
CFT	0.927 (0.869- 0.974)	0.870 (0.836- 0.900)
Priors increased by 5		
ELISA	0.930 (0.871- 0.976)	0.870 (0.836- 0.901)
RBT	0.927 (0.868- 0.974)	0.870 (0.836- 0.974)
CFT	0.928 (0.869- 0.975)	0.870 (0.836-0.900)
Uniform Priors		
ELISA	0.925 (0.864- 0.974)	0.869 (0.834- 0.899)
RBT	0.928 (0.869- 0.975)	0.870 (0.836- 0.900)
CFT	0.927 (0.868- 0.974)	0.870 (0.836- 0.900)
No pseudo-replication	0.960 (0.887- 0.993)	0.855 (0.801- 0.900)

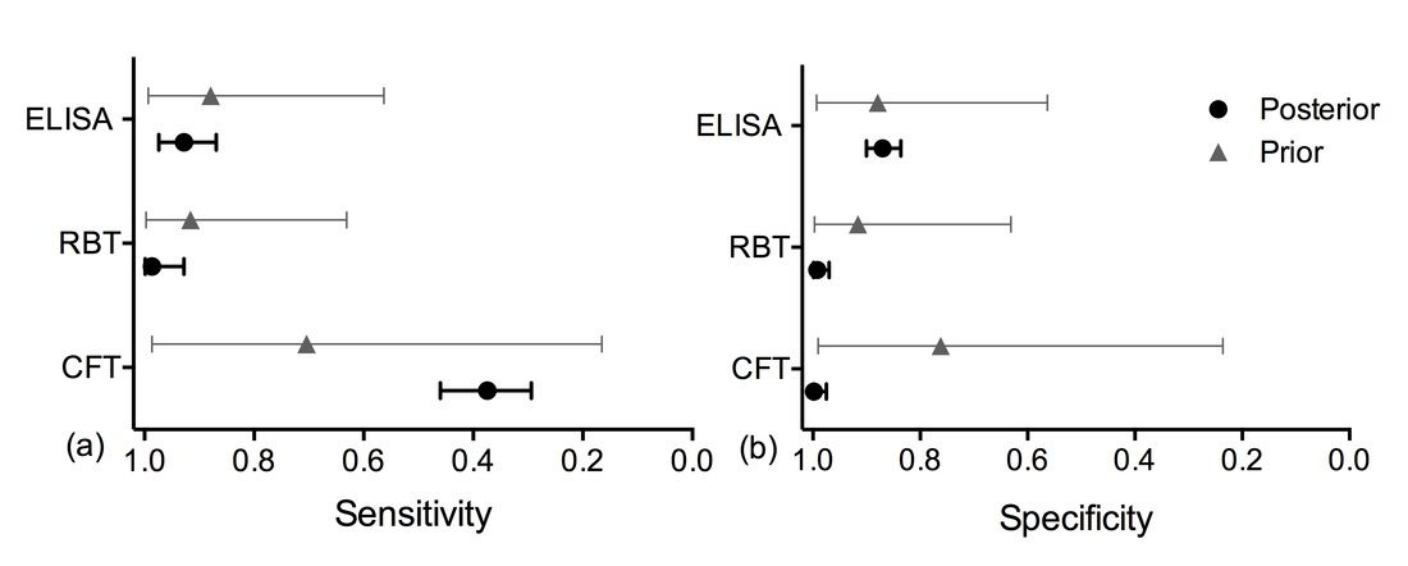
478

Figure Legends

Figure 1. Two-graph receiver operating characteristic curve that plots sensitivity (Se), specificity (Sp), and their non-parametric confidence bands as a function of test cut-off value. Vertical dashed lines show the cut-off value selected by ROC analysis for further investigation (cut-off= 159%).

Figure 2. Summary of prior and posterior distributions for latent class analysis of ELISA, rose bengal test (RBT), and complement fixation test (CFT) accuracy. Prior information for the sensitivity and specificity of each test is summarized by the median and 95th percentile of their distribution. Median parameter estimates and 95% Bayesian credible intervals for (a) sensitivity and (b) specificity are displayed for the model assuming conditional dependence between the ELISA and CFT.


Se
Sp
Se conf. band
Sp conf. band
Sensitivity & Specificity
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
Test cut-off value


ELISA
RBT
CFT
(a)
1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
ELISA
RBT
CFT
(b)
1.0
0.8
0.6
0.4
0.2
0.0
Specificity
Posterior
Prior