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

The Faculty of Oregon State University has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DOI</td>
<td>10.7589/2013-12-334</td>
</tr>
<tr>
<td>Publisher</td>
<td>Wildlife Disease Association</td>
</tr>
<tr>
<td>Version</td>
<td>Accepted Manuscript</td>
</tr>
<tr>
<td>Terms of Use</td>
<td><a href="http://cdss.library.oregonstate.edu/sa-termsofuse">http://cdss.library.oregonstate.edu/sa-termsofuse</a></td>
</tr>
</tbody>
</table>
Evaluation of the sensitivity and specificity of an ELISA for diagnosing brucellosis in 
African buffalo (*Syncerus caffer*).

Erin E. Gorsich¹,⁵, Roy G. Bengis², Vanessa O. Ezenwa³, Anna E. Jolles⁴

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

Traditional estimates of diagnostic test sensitivity and specificity are based on direct comparisons against an established gold standard test (detection of \textit{Brucella} organisms by culture methods). Because true gold standard test results are often costly or impractical to obtain, especially in wildlife systems, a new test’s accuracy is commonly estimated by comparing it to a reference test with a known error rate (Buck and Gart, 1966) or by comparison to multiple imperfect diagnostic tests (Enoe et al., 2000). Techniques that estimate test accuracy or disease prevalence when there is uncertainty in the test’s sensitivity or specificity are called latent class analyses because they use the observed frequency of diagnostic test results to estimate a latent
variable, the true disease status, from which the new diagnostic test can be evaluated (Branscum et al., 2005). Accurate estimates of a test’s sensitivity and specificity with latent class analysis requires correctly representing whether the outcomes of two tests for a given animal are independent or correlated (conditional upon the true state of the animal; Georgiadis et al., 2003). Therefore, we consider potential correlations among tests in this analysis. This analysis also follows an increasing trend in the use of a Bayesian inference with latent class analysis; examples include the estimation of test accuracy for Foot and Mouth disease (Engel et al., 2008), tuberculosis (Alvarez et al., 2012), and brucellosis in cattle (Matope et al., 2011). Bayesian inference could also be useful for diagnostic test evaluation in wildlife because it incorporates uncertainty about model parameters based on independently collected, or prior information. These techniques are recommended by the OIE to estimate sensitivity and specificity, but represent only one step in the validation process (OIE, 2008b). The assumptions and modifications used in latent class analyses have been reviewed in general for latent class techniques (Enoe et al., 2000) and more specifically for latent class techniques with Bayesian inference (Branscum et al., 2005).

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

**Materials and Methods**
Animal captures and test methods

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

We used three serological measures of brucellosis infection. The rose bengal test (RBT) and complement fixation test (CFT) were conducted by the Onderstepoort Veterinary Institute’s diagnostic laboratories in South Africa according to OIE specifications (OIE, 2008a). Briefly, the RBT is conducted by monitoring the agglutination response after mixing serum with rose bengal stained _B. abortus_ cells. The CFT is conducted by monitoring the degree of haemolysis after incubating inactivated test serum, antigen, and exogenous complement with sensitized 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:

\[
\text{Cut-off\%} = \left(100 \times \frac{\mu(\text{OD}_{450\text{ of the paired sample wells})} - \mu(\text{OD}_{450\text{ of negative control wells}})}{\mu(\text{OD}_{450\text{ of positive control wells})} - \mu(\text{OD}_{450\text{ of negative control wells}})}\right).
\]

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

Latent class analysis and prior estimation

Latent class analysis allows evaluation of diagnostic tests in the absence of a gold standard. The simplest model presented here assumes that the outcomes of the tests for a given animal are independent, conditional upon the true state of the animal. This model is referred to as the conditional independence model and is described in detail in the appendix. A complete model specification and review of Bayesian approaches to estimation can be found in Branscum et al (2005); the models initial descriptions in two and three populations can be found in Hui and Walter (1980) and Walter and Irwig (1988), respectively. The validity of assuming two tests are conditionally independent requires further justification (Vacek, 1985). The results of diagnostic tests that measure similar biological processes are likely to be correlated (conditional on the
animals true disease status; Gardner et al., 2000) and assuming independence may result in incorrect estimates of test accuracy (Georgiadis et al., 2003). The RBT, CFT, and ELISA all measure the hosts’ antibody response to *Brucella* smooth LPS, but they use different methods of antibody detection (Godfroid et al., 2010; Nielsen, 2002). Therefore, because we had little prior knowledge about the potential correlation between test outcomes, we consider models assuming both conditional independence and conditional dependence.

We used model selection based on Deviance Information Criteria (DIC) to compare the fit of models assuming conditional independence and conditional dependence (e.g. Rahman et al., 2013). DIC is a model assessment tool based on model fit and the effective number of parameters (Link and Barker, 2010). Models with lower DIC values provide a better fit to the data, and we chose the model with the lowest DIC value (Spiegelhalter et al., 2002). Prior distributions for diagnostic test sensitivity and specificity were represented as beta distributions and were defined using published results from test validations in cattle (Grenier et al., 2009; Table 1). The prior distributions for each parameter are displayed in Table 1 and details of their specification are given in the appendix. This prior information was combined with the full dataset of 571 samples (Supplement 2). Median and 95% Bayesian credible intervals are presented for all parameters in the best fitting model. We conducted sensitivity analyses on this model by (1) increasing the mode and lower bound of the each tests’ sensitivity and specificity prior distributions by 5 percentage points, (2) decreasing the mode and lower bound of the each tests’ sensitivity and specificity prior distributions by 5 percentage points, and (3) by specifying uninformative priors between the interval of zero to one, modeled as Beta (1,1), for each tests’ 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 to 1.00) 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
sensitivity than the CFT, similar but lower sensitivity to the RBT, but a lower specificity than both the RBT and CFT. The estimates of test accuracy in this study are based on the selected ELISA cut-off value. The cut-off value that maximized the total sensitivity and specificity (Se+Sp) was 159%. Because test sensitivity and specificity are inversely related at a given cut-off value, a different cut-off would result in altered estimates of test accuracy. The selected cut-off value should be taken into account when comparing diagnostic tests (Greiner et al., 2000). For example, at the suggested cut-off value for cattle, 120%, the ELISA had a lower specificity and a higher sensitivity. This result emphasizes the importance of test optimization for each population and species to which it is applied.

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

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

The uncertainty in how any of the diagnostic tests relate to active infection in wildlife represents a major hurdle to accurate diagnostic methods (Treanor et al., 2011). Owing to these limitations, the results of this evaluation serve as a comparison among the serological tests.
historically used. Additional assays for brucellosis, including the FPA (Gall et al., 2000), PCR (Bricker, 2002), and alternative ELISA techniques (Nielsen, 2002) have shown improved accuracy in other systems and should be considered for future testing in African buffalo. Further, the estimates of test sensitivity and specificity presented in this analysis includes prior information (Figure 2). Rather than a limitation, incorporating this information could be a valuable tool for wildlife studies given the sample size requirements and potential identifiability problems with latent class models (Dendukuri et al., 2010). Our analysis also assumes that test sensitivity and specificity are consistent throughout the course of brucellosis infection and between populations. As more information develops about the course of brucellosis in buffalo, future diagnostic tests evaluations should incorporate variation in detection rates between different stages of infection (Engel et al., 2010; Caraguel et al., 2012) or different populations (Munoz et al., 2012).

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


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).

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

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.

<table>
<thead>
<tr>
<th>Pseudo-gold standard status</th>
<th>ELISA Cut-off &gt; 120</th>
<th>ELISA Cut-off &gt; 159</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>16</td>
<td>107</td>
</tr>
<tr>
<td># Misclassified/ Accuracy</td>
<td>16/ 89.4%</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>1 (0.82-1.00)</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>0.87 (0.80-0.92)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. ELISA, rose bengal test (RBT), and complement fixation test (CFT) results classified for the Lower Sabie region (LS) and the Crocodile Bridge (CB) region.

<table>
<thead>
<tr>
<th>ELISA/RBT/CFT</th>
<th>+/+/+</th>
<th>+/+/-</th>
<th>+/-/+</th>
<th>+/-/-</th>
<th>-/+/+</th>
<th>-/+/-</th>
<th>-/-/+</th>
<th>-/-/-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Sabie</td>
<td>24</td>
<td>28</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>161</td>
</tr>
<tr>
<td>Crocodile Bridge</td>
<td>21</td>
<td>47</td>
<td>0</td>
<td>39</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>220</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>75</td>
<td>0</td>
<td>58</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>381</td>
</tr>
</tbody>
</table>
Table 4. Sensitivity analyses of prior information and model assumptions. Results include the consequence of adjusting prior information about each tests’ accuracy and re-fitting the model to a subset of the samples were each of the 258 buffalo are represented once. In analyses adjusting test accuracy, the mode and lower bound were increased/decreased by 5 percentage points.

<table>
<thead>
<tr>
<th>Model Specification</th>
<th>ELISA Se (95%CrI)</th>
<th>ELISA Sp (95%CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD between ELISA &amp; CFT</td>
<td>0.928 (0.869-0.974)</td>
<td>0.870 (0.836-0.900)</td>
</tr>
<tr>
<td>Priors decreased by 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELISA</td>
<td>0.925 (0.867-0.971)</td>
<td>0.869 (0.835-0.900)</td>
</tr>
<tr>
<td>RBT</td>
<td>0.933 (0.873-0.977)</td>
<td>0.870 (0.836-0.900)</td>
</tr>
<tr>
<td>CFT</td>
<td>0.927 (0.869-0.974)</td>
<td>0.870 (0.836-0.900)</td>
</tr>
<tr>
<td>Priors increased by 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELISA</td>
<td>0.930 (0.871-0.976)</td>
<td>0.870 (0.836-0.901)</td>
</tr>
<tr>
<td>RBT</td>
<td>0.927 (0.868-0.974)</td>
<td>0.870 (0.836-0.974)</td>
</tr>
<tr>
<td>CFT</td>
<td>0.928 (0.869-0.975)</td>
<td>0.870 (0.836-0.900)</td>
</tr>
<tr>
<td>Uniform Priors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELISA</td>
<td>0.925 (0.864-0.974)</td>
<td>0.869 (0.834-0.899)</td>
</tr>
<tr>
<td>RBT</td>
<td>0.928 (0.869-0.975)</td>
<td>0.870 (0.836-0.900)</td>
</tr>
<tr>
<td>CFT</td>
<td>0.927 (0.868-0.974)</td>
<td>0.870 (0.836-0.900)</td>
</tr>
<tr>
<td>No pseudo-replication</td>
<td>0.960 (0.887-0.993)</td>
<td>0.855 (0.801-0.900)</td>
</tr>
</tbody>
</table>
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.