

# No Gold Standard Estimation of the Sensitivity and Specificity of Two Molecular Diagnostic Protocols for *Trypanosoma brucei* spp. in Western Kenya

Barend Mark de Clare Bronsvort<sup>3\*</sup>, Beatrix von Wissmann<sup>1</sup>, Eric Maurice Fèvre<sup>2</sup>, Ian Graham Handel<sup>3</sup>, Kim Picozzi<sup>1</sup>, Sue Christina Welburn<sup>1</sup>

**1** Centre for Infectious Diseases, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, United Kingdom, **2** Centre for Infectious Diseases and Centre for Infection, Immunity and Evolution, School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom, **3** The Roslin Institute and The Royal (Dick) School of Veterinary Studies, University of Edinburgh, Roslin, Midlothian, United Kingdom

## Abstract

African animal trypanosomiasis is caused by a range of tsetse transmitted protozoan parasites including *Trypanosoma vivax*, *Trypanosoma congolense* and *Trypanosoma brucei*. In Western Kenya and other parts of East Africa two subspecies of *T. brucei*, *T. b. brucei* and the zoonotic *T. b. rhodesiense*, co-circulate in livestock. A range of polymerase chain reactions (PCR) have been developed as important molecular diagnostic tools for epidemiological investigations of *T. brucei* s.l. in the animal reservoir and of its zoonotic potential. Quantification of the relative performance of different diagnostic PCRs is essential to ensure comparability of studies. This paper describes an evaluation of two diagnostic test systems for *T. brucei* using a *T. brucei* s.l. specific PCR [1] and a single nested PCR targeting the Internal Transcribed Spacer (ITS) regions of trypanosome ribosomal DNA [2]. A Bayesian formulation of the Hui-Walter latent class model was employed to estimate their test performance in the absence of a gold standard test for detecting *T. brucei* s.l. infections in ear-vein blood samples from cattle, pig, sheep and goat populations in Western Kenya, stored on Whatman FTA cards. The results indicate that the system employing the *T. brucei* s.l. specific PCR ( $Se_1 = 0.760$ ) had a higher sensitivity than the ITS-PCR ( $Se_2 = 0.640$ ); both have high specificity ( $Sp_1 = 0.998$ ;  $Sp_2 = 0.997$ ). The true prevalences for livestock populations were estimated ( $p_{cattle} = 0.091$ ,  $p_{pigs} = 0.066$ ,  $p_{goats} = 0.005$ ,  $p_{sheep} = 0.006$ ), taking into account the uncertainties in the specificity and sensitivity of the two test systems. Implications of test performance include the required survey sample size; due to its higher sensitivity and specificity, the *T. brucei* s.l. specific PCR requires a consistently smaller sample size than the ITS-PCR for the detection of *T. brucei* s.l. However the ITS-PCR is able to simultaneously screen samples for other pathogenic trypanosomes and may thus be, overall, a better choice of test in multi-organism studies.

**Citation:** Bronsvort BmC, Wissmann Bv, Fèvre EM, Handel IG, Picozzi K, et al. (2010) No Gold Standard Estimation of the Sensitivity and Specificity of Two Molecular Diagnostic Protocols for *Trypanosoma brucei* spp. in Western Kenya. PLoS ONE 5(1): e8628. doi:10.1371/journal.pone.0008628

**Editor:** Photini Sinnis, New York University School of Medicine, United States of America

**Received:** October 26, 2009; **Accepted:** December 9, 2009; **Published:** January 7, 2010

**Copyright:** © 2010 Bronsvort et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was supported by the United Kingdom Department for International Development Animal Health Programme (DFID-AHP). EMF is funded by the Wellcome Trust (085308). IH is funded by the Biotechnology and Biological Sciences Research Council (BBSRC) (338BDD RA0762). The funding agencies had no role in conducting the study and in preparing the manuscript. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: Mark.Bronsvort@ed.ac.uk

## Introduction

Trypanosomiasis, or ‘nagana’, is an infectious disease of livestock caused by a range of protozoan parasites. *Trypanosoma vivax*, *Trypanosoma congolense* and *Trypanosoma brucei* are the three most important species of trypanosome, responsible for considerable production losses and livestock morbidity where they occur [3]. These parasites are transmitted by tsetse flies in the genus *Glossina*, in which they have obligate life cycle stages. Severity of infection with these trypanosomes depends on a range of factors; in local zebu cattle (*Bos indicus*) in western Kenya and elsewhere in East Africa, trypanosomiasis is an endemic disease, causing chronic anaemia [4], enlarged lymph nodes, staring coat, weakness and depression, and general loss of productivity and overall condition, including reduced milk yield and impaired fertility [5]. *T. brucei*, which is perhaps the least pathogenic of the three species in cattle [6], has three sub-species, namely *T. b. brucei*, *T. b. gambiense* and *T. b.*

*rhodesiense*; in Western Kenya and other parts of East Africa, *T. b. brucei* and *T. b. rhodesiense* co-circulate in cattle and other livestock species. As *T. b. rhodesiense* is the agent of the zoonotic form of human sleeping sickness, understanding the epidemiology of *T. brucei* s.l. in cattle is important both for understanding and controlling animal trypanosomiasis, but also with regards to estimating the size of the reservoir of human infective parasites.

Classical estimates of sensitivity and specificity are based on direct, empirical comparisons of test outcomes for different tests, where an index test is compared to an established “gold-standard” which has an assumed sensitivity and specificity of 100%. For field diagnosis in rural settings, microscopy-based techniques using direct observation of wet blood films, or concentration techniques such as the Buffy Coat Technique, BCT [7] and the Haematocrit Centrifugation Technique, HCT [8] are the most common method of parasite detection, and have historically been considered the gold standard. Recent studies [9] have illustrated,

**Table 1.** The apparent prevalence estimates for each species in the study based on the individual and cumulative test results from the *T. brucei* s.l. specific PCR, *T. brucei* s.l. ITS-PCR.

	<b>T<sub>1</sub>+ prevalence</b>	<b>(95%CI)</b>	<b>T<sub>2</sub>+ prevalence</b>	<b>(95%CI)</b>	<b>Cummulative T<sub>1</sub> &amp; T<sub>2</sub></b>	<b>(95%CI)</b>
Cattle (n = 1260)	0.071	(0.058–0.086)	0.060	(0.048–0.075)	0.087	(0.072–0.104)
Goats (n = 764)	0.005	(0.002–0.013)	0.004	(0.001–0.011)	0.008	(0.004–0.017)
Pigs (n = 311)	0.051	(0.032–0.082)	0.045	(0.027–0.074)	0.061	(0.039–0.093)
Sheep (n = 427)	0.002	(0.0001–0.013)	0.007	(0.002–0.020)	0.007	(0.002–0.020)

T<sub>1</sub> = *T. brucei* s.l. specific PCR; T<sub>2</sub> = *T. brucei* s.l. ITS-PCR.  
doi:10.1371/journal.pone.0008628.t001

however, that microscopy has a very poor sensitivity compared to diagnosis with molecular tools, highlighting that previous studies using these technologies are likely to have significantly underestimated both animal- and herd-level prevalence of these pathogens. This has clinical implications for the management of individual animals, but also important epidemiological implications regarding the zoonotic potential of *T. brucei* s.l. As a result, PCR-based diagnosis of *T. brucei* s.l. in livestock has now been used in a number of studies across Africa [2,10,11,12,13], using a number of different protocols and methods [1,2,14].

To enable comparisons between different studies, the relative performance of different testing systems needs to be quantified, preferably in such a way as to enable unbiased estimates of the true prevalence to be made, while accounting for uncertainty in the specificity and sensitivity of the system used (we refer to the *testing system* as the combination of the diagnostic protocol and method of sample collection and processing). We know of only one other study [15] that compared different PCR protocols (including those we examine here); while this was a valuable addition to the literature, their analysis assessed only agreement between tests and did not assess sensitivity or specificity, or indeed make estimates of the true prevalence based on the outcomes of the different tests. The development of a latent class model by Hui and Walter [16] to estimate sensitivity and specificity avoids the need for a “gold standard” which is rarely, if ever, genuinely perfect [17]. The extension of this into a Bayesian framework allows the uncertainty in the prior beliefs about the tests to be included [18] and full posterior distributions of the estimates to be given.

In the present paper, we compare two PCR-based testing systems for the detection of *T. brucei* s.l. in populations of cattle, pigs, sheep and goats in Western Kenya: a) a *T. brucei* s.l. specific primer pair [1] on material originating from ear-vein blood and stored on Whatman FTA filter cards [9]; b) ear-vein blood samples on Whatman FTA cards amplified using a single nested PCR targeting the Internal Transcribed Spacer (ITS) regions of ribosomal DNA [2]. We present estimates of sensitivity, specificity and predictive values for these two testing systems following the

**Table 2.** Test cross tabulation by species (T<sub>1</sub> = *T. brucei* s.l. specific PCR; T<sub>2</sub> = *T. brucei* s.l. ITS-PCR).

	<b>T<sub>1</sub>+/T<sub>2</sub>+</b>	<b>T<sub>1</sub>-/T<sub>2</sub>+</b>	<b>T<sub>1</sub>+/T<sub>2</sub>-</b>	<b>T<sub>1</sub>-/T<sub>2</sub>-</b>
Cattle (n = 1260)	55	21	34	1150
Goats (n = 764)	1	2	3	758
Pigs (n = 311)	11	3	5	292
Sheep (n = 427)	1	2	0	424

doi:10.1371/journal.pone.0008628.t002

STARD guidelines [19] and report on the estimated true prevalence of *T. brucei* s.l. in livestock in two areas within the Busia District of Western Kenya.

## Results

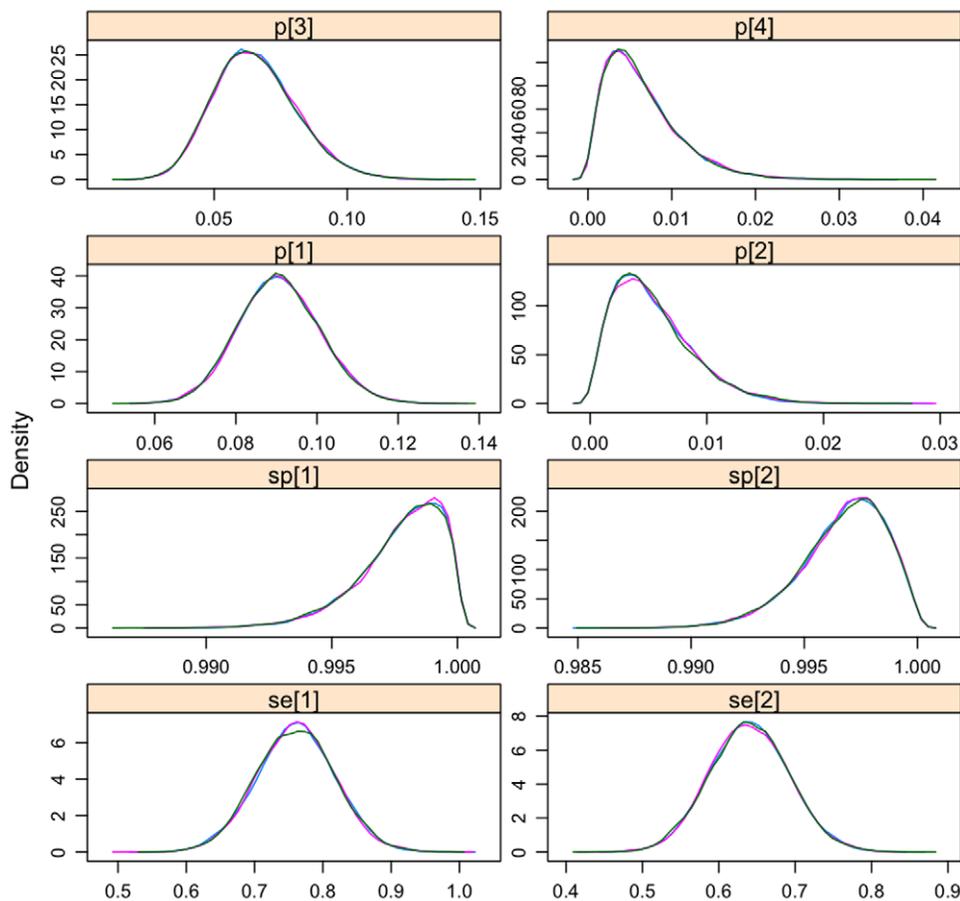
A total of 1,260 cattle, 764 goats, 311 pigs and 427 sheep were sampled across the two study sites and tested using both PCR techniques. The estimated apparent prevalence of *T. brucei* s.l. by each PCR method as well as the cumulative apparent prevalence are given in Table 1. The apparent prevalence is highest in cattle and pigs and lowest in sheep and goats. The estimates from each test differ only slightly, with the ITS-PCR appearing to be less sensitive than the *T. brucei* s.l. specific PCR. In this situation of low prevalence this difference in apparent prevalences is unlikely to be biologically important. However, the different tests are clearly classifying slightly different subsets of the population as infected/uninfected, highlighted by the cumulative prevalence being higher than the individual estimates (Table 1). The species concordant and discordant test classification results are given in Table 2: these form the input for the Hui-Walter model.

The unbiased estimates from the Hui-Walter model are given in Table 3. The estimates of sensitivity are low for both tests but the *T. brucei* s.l. specific PCR on average appears to be more sensitive. Both tests are highly specific. The estimated probability distributions of the test parameters are given in the density plots in Figure 1. These plot show the relative probability of the parameter taking a given value on the x axis and are effectively a smoothed

**Table 3.** Parameter estimates (and Bayesian 95% credibility intervals, BCI) for the *T. brucei* s.l. specific PCR (Se<sub>1</sub> and Sp<sub>1</sub>), *T. brucei* s.l. ITS-PCR (Se<sub>2</sub> and Sp<sub>2</sub>) and adjusted *T. brucei* s.l. prevalence estimates for cattle (p<sub>cattle</sub>), goats (p<sub>goats</sub>), pigs (p<sub>pigs</sub>) and sheep (p<sub>sheep</sub>) from the Hui-Walter model assuming conditional independence.

<b>Parameter</b>	<b>Mean</b>	<b>95% BCI</b>
Se <sub>1</sub>	0.760	0.648–0.873
Sp <sub>1</sub>	0.998	0.994–1.00
Se <sub>2</sub>	0.640	0.540–0.744
Sp <sub>2</sub>	0.997	0.992–1.00
p <sub>cattle</sub>	0.091	0.072–0.111
p <sub>goats</sub>	0.005	0.001–0.014
p <sub>pigs</sub>	0.066	0.038–0.099
p <sub>sheep</sub>	0.006	0.001–0.018

doi:10.1371/journal.pone.0008628.t003



**Figure 1. Probability density plots.** Probability density plots for each test parameter estimate for the *T. brucei* s.l. specific PCR (sensitivity = Se[1]; specificity = Sp[1]), *T. brucei* s.l. ITS-PCR (sensitivity = Se[2]; specificity = Sp[2]) and the adjusted prevalence estimates from the Hui-Walter model assuming conditional independence for cattle (p[1]), goats (p[2]), pigs (p[3]) and sheep (p[4]). The x axes give the parameter estimate and the y axis the relative likelihood of it taking that value. doi:10.1371/journal.pone.0008628.g001

histogram. The parameter value at the peak of the distribution represents the most likely value. The model behaved well with good mixing of the three chains as seen in the trace plots (Figure 2) which shows the samples for three chains for each parameter. They show that the chains are statistically stationary and are not autocorrelated. The Gelman-Rubin potential scale reduction factor (PSRF) statistic for all parameters was  $< 1.05$ . The PSRF is a measure of MCMC chain convergence and values substantially above 1 indicate lack of convergence [20]. Both tests have very high specificity with only moderate sensitivity, although the *T. brucei* s.l. specific PCR appears to be about 12% more sensitive than the ITS-PCR. The unbiased estimates of the true prevalence in each host species (Table 3) are higher for cattle and pigs than the apparent prevalence estimates for both those species, indicating that both PCR tests normally underestimate apparent prevalence. The prevalences are so low for sheep and goats that it is difficult to draw clear conclusions for these species. The estimates for sensitivity, specificity and prevalence were robust to removal of the cattle sub-population.

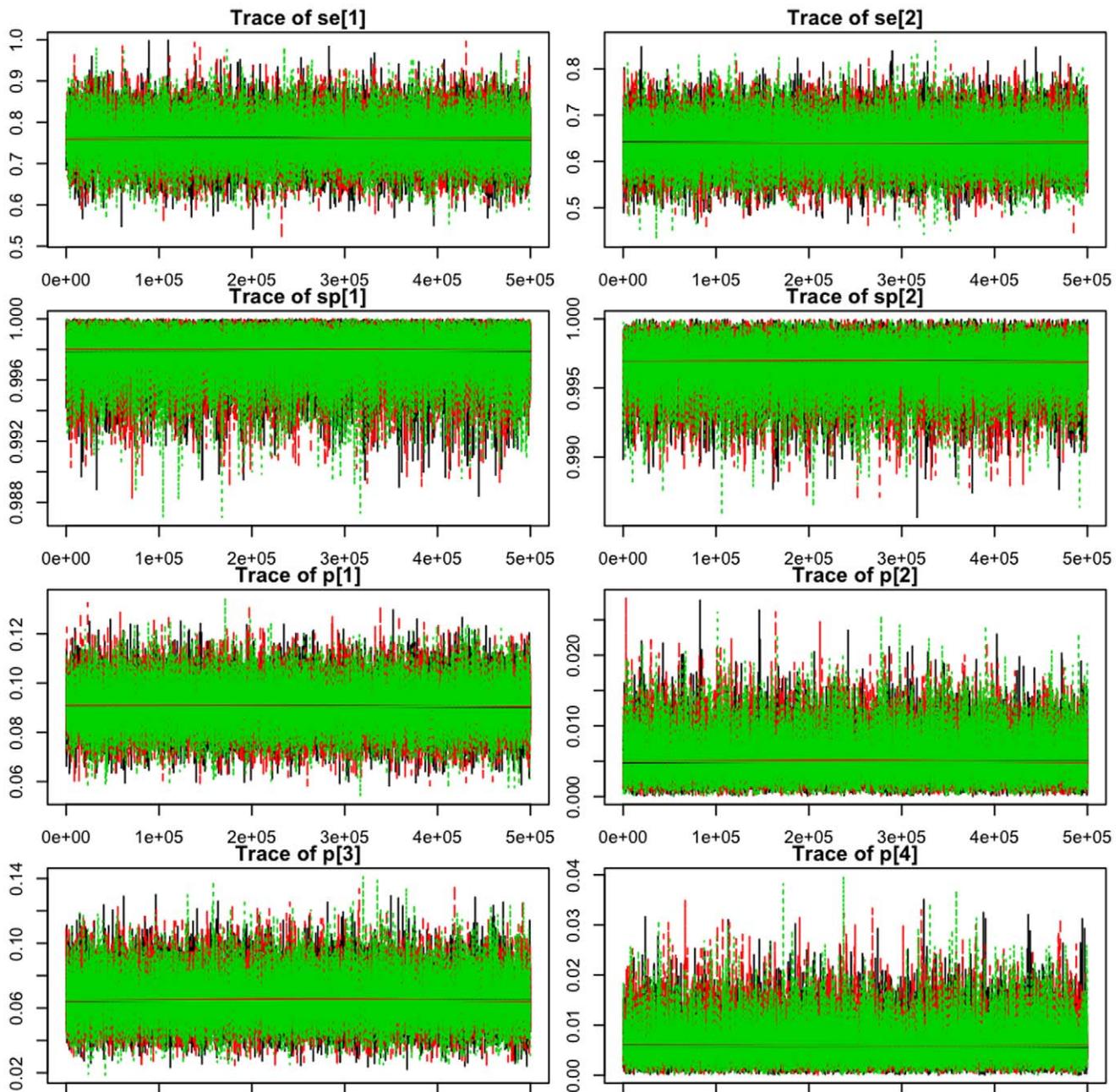
Having estimated the sensitivity and specificity for each test, we use these outputs to estimate the positive predictive values (PPV) and negative predictive values (NPV) for each test across a range of true prevalences (Figure 3). The PPV is the probability that an animal is truly positive given that it has had a positive test result. The NPV is exactly the inverse; that is, the probability that an

animal is truly negative given that it is test negative. The distribution of PPVs and NPVs across a range of prevalences clearly shows that the test performance in this regard is related to prevalence - both tests have high PPVs above 20% prevalence, but this decreases rapidly as the prevalence decreases. Inversely the NPV are extremely high at lower prevalences but decrease steadily as the prevalence increases. For both the PPV and the NPV, the *T. brucei* s.l. specific PCR has a higher predictive value at any given prevalence than the *T. brucei* s.l. ITS-PCR. The estimated PPV and NPV for each test in each sub population are shown in Table 4. From this it is clear that in these sub populations, the PPV of both tests is very high for cattle and pigs, in the region of 95%, because of their relatively high prevalence but decreases markedly for sheep and goats to around 50% due to the low prevalence in these species. For the prevalences estimated in these sub-populations, the NPVs were extremely high.

The impact of the test's performance on survey sample size are illustrated in Figure 4. The *T. brucei* s.l. specific PCR requires a consistently smaller sample size than the *T. brucei* s.l. ITS-PCR as expected with its higher sensitivity and specificity.

## Discussion

Molecular diagnostic tools, and in particular PCR, have vastly improved the detection of trypanosome infections over standard

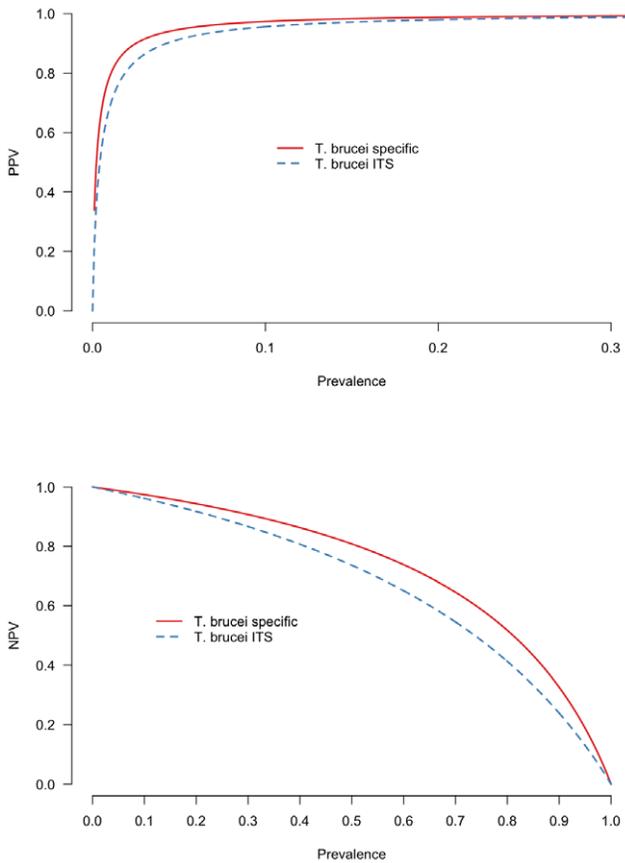


**Figure 2. Trace plots.** MCMC history plots for each parameter of the Hui-Walter model for parameter estimates and true prevalence estimates for the *T. brucei* s.l. specific PCR (sensitivity = Se[1]; specificity = Sp[1]), *T. brucei* s.l. ITS-PCR (sensitivity = Se[2]; specificity = Sp[2]), cattle (p[1]), goats (p[2]), pigs (p[3]) and sheep (p[4]). The plots record every 10th sample from 500,000 iterations and the x axis is the sequence of iterations and the y axis the parameter value from that iteration.

doi:10.1371/journal.pone.0008628.g002

parasitological techniques, by lowering the parasitaemia detection limit by several orders of magnitude. Even when applying concentration techniques, such as the haematocrit centrifugation technique (HCT) or the buffy coat technique (BCT), the analytical sensitivity of microscopy ranges between detectable parasitaemias of  $2.5 \times 10^2$  to  $5 \times 10^3$  parasites/ml of blood depending on trypanosome species [21], whereas PCR can detect the presence of parasite DNA equivalent to one trypanosome in 10ml of host blood [22]. The analytic detection limit of the *T. brucei* s.l. specific PCR has been shown to be as low as 1/10 of the genetic material of a single trypanosome per PCR reaction [1]. The ITS - PCR was been shown

to detect trypanosome DNA at a dilution equivalent to less than one parasite/ml of host blood [2]. Furthermore, primer design targeting precise DNA sequences ensures high specificity of PCR, making it independent of morphological differences required for speciation by microscopy. PCR has thus been incorporated as the diagnostic tool of choice into a wide number of studies investigating the epidemiology of trypanosomiasis, especially since advances in preservation methodology for biological samples have facilitated collection and stabilization of field samples of sufficiently high quality for molecular analysis. The field applications of PCR include estimating trypanosome prevalence for the monitoring of control programmes, though



**Figure 3. Predictive value plots.** The positive (PPV) and negative predictive value (NPV) of the *T. brucei* s.l. specific PCR and the ITS PCR over a range of prevalence of *T. brucei* s.l. PPV is the positive predictive value of the test at a given prevalence i.e. how likely the animal is to have the pathogen given that it has a positive test result. NPV is the negative predictive value of the test at a given prevalence i.e. how likely is an animal not to have the pathogen given that it had a negative test result. doi:10.1371/journal.pone.0008628.g003

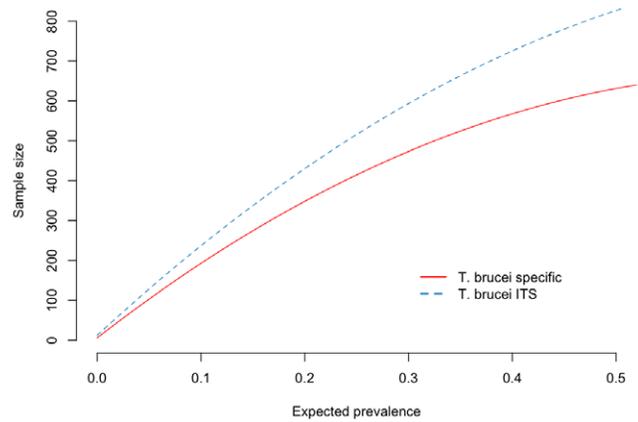
due to the cost and level of laboratory equipment involved, PCR is currently not suitable for diagnostic testing of individual animals for treatment decisions at the local level.

To enable comparison between different protocols, this study used a non-gold standard approach to quantify the parameters of the two PCR-based test systems: a *T. brucei* s.l. specific PCR [1]

**Table 4. Estimated predictive values of the two tests in the four sub-populations (Test 1 = *T. brucei* s.l. specific PCR; test 2 = *T. brucei* s.l. ITS-PCR).**

Sub-population	Prevalence	PPV		NPV	
		Test 1	Test 2	Test 1	Test 2
cattle	0.091	0.970	0.950	0.976	0.965
goats	0.005	0.625	0.505	0.999	0.998
pigs	0.066	0.957	0.929	0.983	0.975
sheep	0.006	0.659	0.534	0.998	0.998

PPV is the positive predictive value of the test at a given prevalence i.e. how likely the animal is to have the pathogen given that it has a positive test result. NPV is the negative predictive value of the test at a given prevalence i.e. how likely is an animal not to have the pathogen given that it had a negative test result. doi:10.1371/journal.pone.0008628.t004



**Figure 4. Sample size plots.** Approximate sample sizes for a simple random survey to estimate infection prevalence using the two tests. The samples sizes are calculated to estimate prevalence with a 95% confidence interval for an absolute precision of  $\pm 5\%$ . doi:10.1371/journal.pone.0008628.g004

and ITS-PCR [2] run in parallel on different punches of the same wholeblood samples from cattle, pigs, goats and sheep stored on FTA-cards. This approach simultaneously allowed for an unbiased estimate of the true prevalence of *T. brucei* s.l. infections in the different livestock species, whilst accounting for the uncertainties in the specificity and sensitivity of the testing systems used. As expected, based on the high target-specificity of the primers designed for the respective PCRs [1,2] the specificity of the test system was high (*T. brucei* s.l. specific PCR:  $Sp_1 = 0.998$ ; *T. brucei* s.l. ITS-PCR:  $Sp_2 = 0.997$ ) for both test systems. However, the sensitivities of the test systems were lower than expected ( $Se_1 = 0.76$ ,  $Se_2 = 0.64$ ) considering the detection limit of the PCRs themselves, which are capable of detecting very low parasitaemia. This low sensitivity may largely be due to the sample storage element of testing system - FTA cards preserve the sample by lysing cells and fixing the DNA *in situ* to the filter-paper matrix, and it has been shown that at low parasitaemias the parasite DNA is localized on the FTA card, with the result that using single punches for each PCR may result in negative PCR test results because the sub-sample (punch) of blood isolated for the specific reaction did not include any parasite DNA (Cox, PhD thesis 2007, University of Edinburgh). Other factors that may contribute to the low sensitivity of the testing system could include residual PCR inhibitors in the sampling material, such as haem, although the sample preparation protocol is designed to remove such known inhibitors [23]. Differences in sensitivity between the two testing systems may be attributable to a higher number of copies of the target sequence for the *T. brucei* s.l. specific PCR (10,000 copies/genome) as compared to the ITS-PCR (200 copies/genome) [1,2]. It may also be that the efficiency of PCR amplifications from the FTA filter paper matrix depend on the target sequence length (1250 base pairs for the ITS-PCR and 173 for the *T. brucei* s.l. specific PCR), though there is no empirical evidence that supports such a negative effect.

The Hui-Walter model assumes that the sensitivity and specificity of the tests are the same across different populations. In this case we have used the different species as the different populations. For an antibody based test this would be problematic as the types of antibody response may be quite different between species. Here, however, the test is detecting the parasite DNA on an FTA card. While differences in parasitaemia between host species have been reported, differentiating a real species-specific effect from the expected individual

animal level variation would be a significant undertaking [24], and we feel that it is safe for the purposes of this analysis to assume average parasite densities between host species are sufficiently similar to not affect this. Concern that co-infection of *T. theileri* in cattle may reduce specificity in this species was explored by re-estimation without the cattle population and estimates were found to be robust. Cross tabulation of the *T. brucei* tests with the *T. theileri* showed 27/1260 cattle and 1/427 sheep to be *T. theileri* positive. Only one cow was positive for both *T. brucei* and *T. theileri* with the ITS test. From these findings we find little evidence of *T. brucei* co-infection of *T. theileri* and therefore we conclude that specificity across the different sub-populations is unlikely to be influenced by *T. theileri* co-infection. The model also assumes conditional independence between the tests: i.e. given a truly positive animal, the results of the first test are independent of the second test. We believe this is justified with regards to sensitivity and specificity as the PCRs use different primer sets and target regions. The assumption that the specificity is  $>0.5$ , as reported in the methods to control the label switching issue, is justifiable as these techniques are widely accepted as highly specific because of the very nature of the technique.

The estimated true prevalences for both cattle ( $p_{\text{cattle}} = 0.091$ ) and pigs ( $p_{\text{pigs}} = 0.066$ ) were higher than those estimated from the results of the individual test systems, or indeed the cumulative prevalence derived for both tests (see Table 1), taking into account the uncertainties in the specificity and sensitivity. The under-estimation at the higher prevalences is a result of the low sensitivity and false negative results. The estimates in sheep ( $p_{\text{sheep}} = 0.006$ ) and goats ( $p_{\text{goats}} = 0.005$ ) are slightly lower than those from the cumulative test results reflecting the fact the specificity is not 1 which at very low prevalences results in low PPV for the tests and high risk of false positive results. The implications of these estimates are firstly that *T. brucei* s.l. is probably more widespread than currently implemented surveys based on such molecular tools suggest, both in Western Kenya [15] and elsewhere. Secondly, when designing, and assessing the impact of large scale interventions (eg [25]), the parameters of the collection and testing systems in use must be taken in to account to ensure that appropriate conclusions are drawn and recommendations made. Previous studies may have initially under-estimated the scale of the *T. brucei* s.l. reservoir in different livestock species, and may have under-estimated the impact that mass treatment activities have had in addressing it. Regarding the potential reservoir of zoonotic *T. b. rhodesiense*, cattle and pigs may be a more substantial risk than previously estimated [26] highlighting the need to specifically consider the test parameters for the *T. b. rhodesiense*-specific PCR protocols [14,27] in future studies. Finally, it is imperative to standardise protocols or establish, as we do here, the relative performance of different protocols across study populations and between testing centres, in order to make meaningful comparisons between different studies. This has largely been achieved for other diseases where standard protocols with known parameters exist [28] but is lacking for a large number of non-reportable infections such as *T. brucei* s.l.

The *T. brucei* s.l. prevalence estimated from PCR results can be adjusted for the test sensitivity and specificity. However, to obtain an equal precision of estimate for the prevalence, a larger number of samples would be required when using the ITS-PCR as compared to the more sensitive *T. brucei* s.l. PCR. For example, at a population prevalence of 10%, 23% more samples would be required when employing ITS-PCR than when employing the specific PCR, to achieve the same absolute precision of 5% (Figure 4). Particularly for large scale studies involving several thousand individual animals, this has implications on the costs and benefits of diagnostic test choice. While our study suggests that the ITS-PCR is less sensitive than the species-specific primers, it is able to simultaneously screen samples for other pathogenic

trypanosomes [2]; in multi-organism studies, it may be, overall, a better choice of test.

## Methods

### Ethical Statement

This study used biobank samples of blood from a number of livestock species collected from the ear vein. This non invasive approach requiring minimal restraint of the animals was approved by both the University of Edinburgh ethics review committee and the Kenyan Department of Veterinary Services.

### Study Sites

The samples were collected at two study sites within Busia District, Western Province, Kenya. Site 1, located in Funyula Division, comprised nine adjacent villages. Site 2, located in Butula Division, comprised ten adjacent villages. These two sampling areas were established field sites, originally chosen on the basis of a cattle trypanosomiasis prevalence of at least 6%, as established by a cross-sectional survey in 1997 [29] and were well characterised in terms of livestock-keeping dynamics and veterinary care seeking behaviour [30,31].

### Sampling

Census sampling targeting the entire livestock population of the two sampling sites was performed in July (Funyula site) and October (Butula site) 2004, by visiting all livestock keeping homesteads in all 19 sampling villages. Whole blood samples from ear-veins were collected from all cattle ( $n = 1260$ ), pigs ( $n = 311$ ), goats ( $n = 764$ ) and sheep ( $n = 427$ ) at each livestock keeping homestead. A total of 2762 livestock samples from 549 livestock-keeping homesteads were collected. The samples (100  $\mu\text{l}$ ) were directly applied to FTA<sup>®</sup> Cards (Whatman, Maidstone, Kent, UK) and allowed to air dry prior to storage at room temperature, an established method of preservation for sensitive detection of trypanosome infections by PCR [9].

### Laboratory Analysis

Laboratory analysis of all samples was carried out by B.v.W. in the course of her PhD, after one year of laboratory training by K.P..

### Sample Preparation and PCR

All blood samples were analysed by two Polymerase Chain Reactions (PCR) according to the published protocols. For each PCR reaction one 2 mm punch was cut from the samples on the FTA<sup>®</sup> Card and prepared according to the manufacturers instructions. Briefly, the discs were washed twice in FTA purification reagent to remove PCR inhibitors from the sample, followed by two washes with 1xTE buffer to remove residual FTA purification reagent. Once dried, the discs were transferred to PCR tubes to seed the reactions.

The first PCR targets the internal transcribed spacers (ITS) located within the ribosomal RNA genes (200 copies/genome) and detects and differentiates between the important pathogenic African trypanosome species affecting livestock, including *Trypanosoma brucei* s.l. [2]. The second PCR employed is specific for *T. brucei* s.l. with a satellite DNA target (10,000 copies/genome) [1]. One positive control (genomic DNA) and one negative control (blank FTA disc) were run with each set of reactions. PCR products were separated by electrophoresis in a 1.5% (w/v) agarose gel containing 0.5  $\mu\text{g/ml}$  ethidium bromide and visualised by ultraviolet light.

## Statistical Analysis

The Hui-Walter paradigm requires two (or more) tests evaluated in two (or more) populations. This model assumes that: (i) the prevalence of the disease is different within each population; (ii) the tests have the same properties across populations; and (iii) the tests are conditionally independent given the disease status. This Bayesian implementation of the Hui-Walter model [32] assumes that for the  $i$ th sub-population the counts ( $\mathbf{O}_i$ ) of the different combinations of test results,  $+/+$ ,  $+/-$ ,  $-/+$  and  $-/-$  for the two tests, follow a multinomial distribution:

$$\mathbf{O}_i | \text{Se}_j, \text{Sp}_j, p_i \sim \text{Multinomial}(\mathbf{Pr}_i, n_i) \text{ for } i=1,2,\dots,S \text{ and } j=1,2,\dots,T$$

where  $S$  is the number of subpopulations,  $T$  is the number of tests and  $\mathbf{Pr}_i$  is a vector of probabilities of observing the individual combinations of test results. Conditioning on the (latent) disease status, these probabilities can be specified using the  $\text{Se}_j$  and  $\text{Sp}_j$  of the tests and the prevalence ( $p_i$ ) of the sub-populations. The probabilities of observing each test combination in the  $i$ th subpopulation are given by:

$$\Pr(T_1+, T_2+) = \text{Se}_1 \text{Se}_2 p_i + (1 - \text{Sp}_1) (1 - \text{Sp}_2)(1 - p_i)$$

$$\Pr(T_1-, T_2+) = (1 - \text{Se}_1) \text{Se}_2 p_i + (1 - \text{Sp}_1) \text{Sp}_2 (1 - p_i)$$

$$\Pr(T_1+, T_2-) = \text{Se}_1 (1 - \text{Se}_2) p_i + \text{Sp}_1 (1 - \text{Sp}_2)(1 - p_i)$$

$$\Pr(T_1-, T_2-) = (1 - \text{Se}_1)(1 - \text{Se}_2) p_i + \text{Sp}_1 \text{Sp}_2 (1 - p_i)$$

In a Bayesian analysis all parameters are expressed as random variables. Prior distributions for the test properties (sensitivity and specificity) and the prevalence within the sub-populations must be specified. The sensitivity of the two tests and the prevalence in four species were given flat (Uniform(0,1) priors) as there were no published data to inform these estimates. Each test's specificity was given a uniform prior over the range 0.5–1.0. This assumption is still vague but by constraining the specificity above 0.5 we control the label switching issue of Markov chain Monte Carlo (MCMC). This issue is discussed in detail with reference to the estimation of Hui-Walters models by Toft *et al.* [32]. In order to explore the influence of the large cattle sub-population the model was re-estimated using only the small ruminant and pig data.

The model was estimated using the JAGS software [33] using the Runjags package [34] of the R statistical system [35]. Three MCMC chains were run for this analysis. The first 500,000 iterations were discarded as a burn-in and the following 500,000 iterations were kept and thinned to 50,000 for posterior inference. Convergence of the chains after the initial burn-in was assessed by visual inspection of the time-series plots for the parameter samples as well as Gelman-Rubin diagnostic plots using three sample chains with dispersed starting values [36]. The R package CODA [20] was used for analysis and graphing of the MCMC output.

The impact diagnostic test performance can be illustrated by estimation of predictive values, the probability of a tested individual having a given infection status condition on its test result. The positive and predictive values are given by the following formulae:

$$PPV = \frac{\text{Se} \times p}{(\text{Se} \times p) + (1 - \text{Sp})(1 - p)} \quad \text{equation 1}$$

$$NPV = \frac{\text{Sp}(1 - p)}{(1 - \text{Se})p + \text{Sp}(1 - p)} \quad \text{equation 2}$$

Where  $\text{Se}$  is the estimated test sensitivity,  $\text{Sp}$  is the estimated test specificity,  $P$  is the true seroprevalence in the population.

In a classical analysis the single, point estimates of the diagnostic test's sensitivity and specificity are used in these estimators. The Bayesian implementation of the Hui-Walters model produces full, joint posterior estimates of these parameters. We estimated the predictive values of the diagnostic test results over a range of prevalences by numerically integrating equation 1 and equation 2 over the paired estimates of sensitivity and specificity from the model. This approach incorporates the uncertainty and covariance structure of the test performance into the predictive values.

This methodology was extended to estimate the positive and negative predictive values of the two tests in the four study sub-populations by numerically integrating over the joint posterior distribution of prevalence and test sensitivity and specificity for each test/population combination. It is likely that a major application of these tests will be in large scale surveys for the estimation of infection prevalence. Estimates from such surveys will be uncertain due to sampling and imperfections of the diagnostic test. Conventionally these surveys are designed to estimate infection prevalence to a required precision [37]. The required sample size in a survey is a function of the expected prevalence, the diagnostic test performance and the required precision. To estimate the impact of the different diagnostic test performances we calculate approximate sample sizes that would be required for a simple survey over a range of prevalences using the two tests (for a given precision). The analysis uses the following formula for sample size [37] to calculate approximate sample sizes for a 95% confidence with a given absolute error.

$$n = \left( \frac{1.96}{d} \right)^2 \left( \frac{(\text{Se} \times p) + (1 - \text{Sp})(1 - p)}{\text{Se} + \text{Sp} - 1} \right) \left( \frac{(1 - \text{Se})p - (1 - \text{Sp})(1 - p)}{\text{Se} + \text{Sp} - 1} \right)$$

Where  $\text{Se}$  and  $\text{Sp}$  are the diagnostic test sensitivity and specificity,  $p$  is the expected prevalence and  $d$  is the absolute proportional error.

## Acknowledgments

We thank the District Veterinary Officer, Busia District, Kenya for his assistance and support during sample collection in Funyula and Butula Divisions. In particular, we would like to thank the animal health assistants, George Ouma, Robert Sanya, Alex Jonyo, Agrey Nying'uro and Thomas Situma and the ILRI driver Mr Kaigai Karanja for their hard work and assistance during the study, as well as Noreen Machila and Andrew P. Cox for their help with sample collection.

## Author Contributions

Conceived and designed the experiments: BMdCB BvW EMF KP SCW. Performed the experiments: BvW KP. Analyzed the data: BMdCB IGH. Wrote the paper: BMdCB BvW EMF IGH SCW.

## References

- Moser DR, Cook GA, Ochs DE, Baily CP, McKane MR, et al. (1989) Detection of *Trypanosoma congolense* and *Trypanosoma brucei* subspecies by DNA amplification using the polymerase chain reaction. *Parasitology* 99: 57–66.
- Cox A, Tilley A, McOdimba F, Fyfe J, Eisler MC, et al. (2005) A PCR based assay for detection and differentiation of African trypanosome species in blood. *Experimental Parasitology* 111: 24–29.
- Kristjansson PM, Swallow BM, Rowlands GJ, Kruska RL, de Leeuw PN (1999) Measuring the costs of African animal trypanosomiasis, the potential benefits of control and returns to research. *Agricultural Systems* 59: 79–98.
- Taylor K, Authié E (2004) Pathogenesis of animal trypanosomiasis. In: Maudlin I, Holmes PH, Miles MA, eds. *The Trypanosomiasis*. Wallingford, Oxfordshire: CABI Publishing. pp 331–353.
- Connor RJ (1994) The impact of nagana. *Onderstepoort Journal of Veterinary Research* 61: 379–383.
- Uilenberg G (1998) A field guide for the diagnosis, treatment and prevention of African animal trypanosomiasis. Rome: Food and Agricultural Organization of the United Nations.
- Murray M, Murray PK, McIntyre WIM (1977) An improved parasitological technique for the detection of African trypanosomiasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 71: 325–326.
- Woo PTK (1970) The haematocrit centrifugation technique for the diagnosis of African trypanosomiasis. *Acta Tropica* 27: 384–386.
- Picozzi K, Tilley A, Fèvre EM, Coleman PG, Magona JW, et al. (2002) The diagnosis of trypanosome infections: application of novel technology for reducing disease risk. *African Journal of Biotechnology* 1: 39–45.
- Clausen P-H, Wiemann A, Patzel R, Kakaire D, Poetzsch C, et al. (1998) Use of a PCR assay for the specific and sensitive detection of *trypanosoma spp.* in naturally infected dairy cattle in peri-urban Kampala, Uganda. *Annals of the New York Academy of Sciences* 849: 21–31.
- Ng'ayo MO, Njiru ZK, Kenya EU, Muluvi GM, Osir EO, et al. (2005) Detection of trypanosomes in small ruminants and pigs in western Kenya: important reservoirs in the epidemiology of sleeping sickness? *Kinetoplastid Biology and Disease* 4: 5.
- Pinchbeck GL, Morrison LJ, Tait A, Langford J, Meehan L, et al. (2008) Trypanosomiasis in The Gambia: prevalence in working horses and donkeys detected by whole genome amplification and PCR, and evidence for interaction between trypanosome species. *BMC Veterinary Research* 4.
- Solano P, Michel JF, Lefrançois T, de La Rocque S, Sidibe I, et al. (1999) Polymerase chain reaction as a diagnosis tool for detecting trypanosomes in naturally infected cattle in Burkina Faso. *Veterinary Parasitology* 86: 95–103.
- Picozzi K, Carrington M, Welburn SC (2008) A multiplex PCR that discriminates between *Trypanosoma brucei brucei* and *T.b. rhodesiense*. *Experimental Parasitology* 118: 41–46.
- Thumbi SM, McOdimba FA, Mosi RO, Jung'a JO (2008) Comparative evaluation of three PCR base diagnostic assays for the detection of pathogenic trypanosomes in cattle blood. *Parasites and Vectors* 1: 46.
- Hui SL, Walter SD (1980) Estimating the error rates of diagnostic tests. *Biometrics* 36: 167–171.
- Enœ C, Georgiadis MP, Johnson WO (2000) Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Preventive Veterinary Medicine* 45: 61–81.
- Bronsvort BMD, Koterwas B, Land F, Handel IG, Tucker J, et al. (2009) Comparison of a flow assay for brucellosis antibodies with the reference eELISA test in West African *Bos Indicus*. *PLoS One* 4.
- Bossuyet PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. (2003) The STARD Statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Annals of Internal Medicine* 138: W1–12.
- Plummer M, Best N, Cowles K, Vines K (2009) Output analysis and diagnostics for MCMC. R package version 0.13-4.
- Eisler MC, Dwingler RH, Majiwa PAO, Picozzi K (2004) Diagnosis and epidemiology of African animal trypanosomiasis. In: Maudlin I, Holmes PH, Miles MA, eds. *The trypanosomiasis*. Wallingford, Oxfordshire: CABI Publishing.
- Masake RA, Njuguna JT, Brown CC, Majiwa PAO (2002) The application of PCR-ELISA to the detection of *Trypanosoma brucei* and *T. vivax* infections in livestock. *Veterinary Parasitology* 105: 179–189.
- Whatman FTA Cards (Product Sheet) - Collect, archive, transport and purify nucleic acids at all room temperatures.
- Murray M (1989) Factors affecting duration and intensity of trypanosome infections of domestic animals. *Annales de la Societe Belge de Medecine Tropicale* 69: 189–196.
- Kabasa JD (2007) Public-private partnership works to stamp out sleeping sickness in Uganda. *Trends in Parasitology* 23: 191–192.
- Waiswa C, Olaho-Mukani W, Katunguka-Rwakishaya E (2003) Domestic animals as reservoirs for sleeping sickness in three endemic foci in south-eastern Uganda. *Annals of Tropical Medicine and Parasitology* 97: 149–155.
- Welburn SC, Picozzi K, Fèvre EM, Coleman PG, Odiit M, et al. (2001) Identification of human-infective trypanosomes in animal reservoirs of sleeping sickness in Uganda by means of serum-resistance-associated (SRA) gene. *The Lancet* 358: 2017–2019.
- OIE (2009) *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* 2009.
- INCO-DC (2000) Novel approaches to the epidemiology of resistance to drugs used in the control of bovine trypanosomiasis in East Africa. Final Report. Glasgow: International Cooperation with Developing Countries (1994–1998).
- Machila N, Wanyangu SW, McDermott J, Welburn SC, Maudlin I, et al. (2003) Cattle owners' perceptions of African bovine trypanosomiasis and its control in Busia and Kwale Districts of Kenya. *Acta Tropica* 86: 25–34.
- Thuranira CM (2005) Socio-economic factors influencing livestock keeping dynamics in a smallholder crop-livestock system in western Kenya [PhD Thesis]. Edinburgh: University of Edinburgh. 311 p.
- Toft N, Jørgensen E, Højsgaard S (2005) Diagnosing diagnostic tests: evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. *Preventive Veterinary Medicine* 68: 19–33.
- Plummer M (2003) JAGS: A program for analysis of Bayesian Graphical Models using Gibbs Sampling. Distributed Statistical Computing. Vienna.
- Denwood (2009) runjags: Run Bayesian MCMC Models in the BUGS syntax from within R. Available from <http://cran.r-project.org/web/packages/runjags/>.
- R Development Core Team (2009) R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Brooks SP, Gelman A (1998) Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7: 434–455.
- Thrusfield M (1995) *Veterinary Epidemiology*. Cambridge: Blackwell Science Ltd.