ELSEVIER

Preventive Veterinary Medicine 30 (1997) 61–73

PREVENTIVE VETERINARY MEDICINE

Impact of biological factors on the interpretation of bovine trypanosomosis serology

Matthias Greiner [a,*], T. Shivarama Bhat [a], Ralf J. Patzelt [a,b], David Kakaire [c], Gereon Schares [a], Ekkehard Dietz [d], Dankmar Böhning [d], Karl-Hans Zessin [a], Dieter Mehlitz [a]

[a] *Institute for Parasitology and Tropical Veterinary Medicine, Department of Tropical Veterinary Medicine and Epidemiology, Freie Universität Berlin, Berlin, Germany*

[b] *Collaborative Research Unit of the Freie Universität Berlin at the Animal Health Research Centre, Entebbe, Uganda*

[c] *Livestock Research Institute (LIRI), Tororo, Uganda*

[d] *Institute for Social Medicine, Department of Epidemiology, Freie Universität Berlin, Berlin, Germany*

Accepted 2 June 1996

Abstract

A total of 457 cattle from dairy farms in Mukono County, Uganda, were investigated for *Trypanosoma* antibodies by ELISA. The objective of the study was to identify explanatory covariate factors for seropositivity among nine farm-specific and four animal-specific variables. We used logistic regression models for parasitological and serological outcome variables and then compared the adjusted odds ratios for explanatory factors between the models. Age is positively correlated with seropositivity but not with the detection of the parasite. Therefore, age group-specific cut-off values were established using mixture-distribution analysis. This procedure, as well as a mixture-distribution-derived cut-off value for the total sample, resulted in a greater relative efficiency of the ELISA as compared to conventional interpretation (cut-off value defined using non-exposed negative controls). The relevance of age and other biological factors for the serological status is briefly discussed.

Keywords: Trypanosomosis; Cattle; Uganda; ELISA; Biological factors; Covariate analysis; Logistic regression; Fixed-effects model; Age effect; Cut-off value; Mixture-distribution analysis

* Corresponding author: Institute for Parasitology and Tropical Veterinary Medicine, Koenigsweg 67, 14163 Berlin, FRG. Tel.: (+49-30) 8108-2317; fax: (+49-30) 8108-2323; email: mgreiner@vetmed.fu-berlin.de

PII S0167-5877(96)01088-4

1. Introduction

By analogy with epidemiological terminology for causal structures we might regard infection and the elicited antibody response as independent and dependent events, respectively. However, in a diagnostic setting infection state is the variable of interest, and often its observation is biased due to false-negative results with the diagnostic methods involved. Therefore, inverse to the causal relationship, the prediction of infection on the basis of a given serodiagnostic result (diagnostic variable) is of interest. Empirical evidence exists that factors may influence diagnostic variables without being associated with the true state of infection. For example, age is a major factor for antibody test systems in infections with persistent antibodies. The probability of persisting antibodies can be assumed to increase with age in endemic situations. This hypothesis is supported by empirical data for *Gambiense* sleeping sickness seroepidemiology (Kegels et al., 1992). Unlike a confounding factor in risk analysis – which is associated with both the independent and dependent variable – age in our example represents a factor that may distort and invalidate diagnostic inferences from test results by virtue of its association with only the diagnostic variable.

The objective of our study was to investigate the impact of age and other biological variables on the probability of seropositivity using the bovine trypanosomosis model. Since there is no standard terminology for factors that cause variability in serodiagnostic test results we suggest to differentiate between biological (BF) and interfering factors (IF) and give some practical examples. Our data stem from a cross-sectional study of dairy farms in Mukono district, Uganda. The study variables comprised both animal and farm-specific factors. *Trypanosoma* antibodies were detected by enzyme-linked immunosorbent assay (ELISA) with the objective of estimating the prevalence of current infections. We present a set of comparative logistic multiple regression models for the identification of BFs.

2. Materials and methods

2.1. Study area, sampling

Mukono County is located in the south-eastern part of Uganda (0° 31'–0° 12' N; 32° 40'–32° 51' E) and covers an area of approximately 200 km^2. Mukono County is part of Mukono district and includes Mukono, Goma, Kawuga, Kyampsi, Nakisunga and Ntenjeru sub counties. The data derive from a cross-sectional pilot study launched and accomplished in June–July 1994 for a project on trypanocide resistance in the peri-urban dairy production near Kampala. The sampling frame consisted of 187 dairy farms existing in the region (data from census April 1994) from which 50 farms were selected at random using random number tables, stratified for three categories of herd size (small, 1–10 cattle; medium, 11–30 cattle; large, > 30). A total of 487 cattle were sampled systematically on the identified farms (all animals on farms with 1–10 animals; every second animal otherwise). Sera for *Trypanosoma* antibody ELISA were available from 457 animals.

2.2. Questionnaire data

Farm-specific information provided by farm managers and/or owners was collected by one principal interviewer via a field-tested questionnaire (available upon request). The questions concerned farm identification and location (five questions), farm owner and personnel (ten questions), management in general (five questions), production and marketing (eight questions), farm structure (nine questions) feeding/drinking practices (six questions), reproduction management (three questions) and disease prevention (16 questions). For the present study, only nine variables were established using questionnaire data: PROPEXOT (proportion of exotic animals in the herd in percent; 0 = 0, 1 = > 0 to 60; 2 = > 60); WATER (0 = no access to shores of lakes and rivers, 1 = access to shores); PASTURE (farm area in acre (1 acre = 0.405 ha) under pasture divided by herdsize; 0 = zero grazing, 1 = > 0 to 2, 2 = > 2 to 4, 3 = > 4); FEEDSUPPLY (kg dairy meal and other concentrate feed supplement per animal per day; 0 = 0, 1 = > 0 to 2, 2 = > 2 to 3, 3 = > 3); PRETREAT (days after the last prophylactic or curative trypanocide treatment; 0 = > 200, 1 = 91 to 200, 2 = 29 to 90, 3 = < 29); DIMINAZENE, HOMIDIUM, ISOMETAMIDIUM (application of the respective drugs by the farmers; 0 = no, 1 = yes); TICKCONT (frequency of tickcontrol by dip or spray (DS) or pour-on (PO); 0 = DS once per week or PO once a month, 1 = DS twice per week or PO once in every three weeks, 2 = DS three times per week or PO at least once in 2 weeks). Questions from the questionnaire were selected for analysis on the basis of assumptions concerning the biological background of bovine trypanosomosis.

2.3. Clinical and parasitological data

The following variables were animal-level: BREED (0 = local; 1 = local cross breed, 2 = exotic); SEX (0 = female, 1 = male); AGEGROUP (age in years; 0 = ≤ 0.6; 1 = > 0.6 to 1.2, 2 = > 1.2 to 2, 3 = > 2 to 5, 4 = > 5), packed red cell volume, PCV (0 = ≤ 26, 1 = > 26 to 29, 2 = > 29 to 32, 3 = > 32 vol.%). Jugular blood was examined for trypanosomes by haematocrit centrifugation technique (HCT) and by miniature anion-exchanger centrifugation technique (m-AECT). The detection of trypanosomes is indicated by the outcome variable TRYPS (0 = negative, 1 = positive), irrespective of the *Trypanosoma* species involved.

2.4. Trypanosoma antibody ELISA (Ab-ELISA)

Soluble crude antigen from *T. brucei* (MSUS/CI/78/TY0196) was prepared from procyclic forms maintained in vitro as described by Mehlitz and Tietjen (1988). The trypanosomes were disintegrated by three times freezing/thawing and ultrasonification in the presence of N-*p*-tosyl-L-lysine chloromethyl ketone (TLCK, 0.1mM). Soluble antigen was obtained as supernatant after centrifugation (30000 *g*, 30 min, 4°C). The Ab-ELISA was performed as described elsewhere (Greiner et al., 1994). Optimised test conditions were 7.5 μg/ml coating concentration on polystyrene microtiterplates (Greiner No. 655061, high binding capacity, Germany) and 1/250 serum dilution. Plate-to-plate variation was corrected as described by Franke et al. (1994); results are

indicated as optical density (OD) at 450nm. A conventional cut-off value was established as mean plus three-fold standard deviation of 86 bovine sera from Germany (non-endemic controls; assumed to be free of trypanosomosis). This cut-off was used for definition of the serological outcome variable SERORESP (0 = negative, 1 = positive). Computer-assisted mixture analysis (C.A.MAN, Böhning et al., 1992) was applied to the ELISA data to assess distribution heterogeneity. The programme optimises parameter estimates of a mixing distribution (number of subpopulations and their weights) by maximum likelihood estimation. A cut-off (referred to as intrinsic cut-off) is established that differentiates between identified – but originally unobserved – subpopulations. C.A.MAN can be applied to bi- or multimodally-distributed quantitative test data to describe the mixed distributions of low and high responders in the sample (Greiner et al., 1994). In our study, C.A.MAN was used to define the outcome variable SEROSUSP (0 = negative, 1 = positive; indicating an ELISA value as belonging to the high responder group of values). Mixture distribution analysis was then repeated for age strata subsamples to establish the variable SEROAGE (0 = negative, 1 = positive; indicating an ELISA value as belonging to the high responder group of values according to age group-specific intrinsic cut-off values). The mixture-distribution analysis is described in Appendix A.

2.5. Statistical model

Four ordinary multiple-logistic regression (LR) models were set up using the parasitological outcome variable and the three serological outcome variables and the explanatory variables described above. All variables that exhibited an effect on any of the outcome variables ($P < 0.05$; two-tailed Mantel-Haenszel summary odds ratios, stratification for AGEGROUP, PRETREAT, ISOMETAMIDIUM; Bhat, 1994) were forced into the models in order to obtain effect profiles that can be compared between the four models. Adjusted odds ratios (OR) were calculated. The significance of variables in the models was assessed based on the comparison of the Wald statistic with a chi-square distribution with one degree-of-freedom (df). The significance of the complete model was tested by a likelihood ratio statistic (LRS) that compares the fit of the model with only the intercept term versus the complete model with h variables. LRS is then compared to a chi-square distribution with h df. In order to account for extra-binomial variation at the herd level, mixed-effects logistic regression models were applied that included both fixed and a random term and, thus, adjust for herd effects. Specifically, a modified LRS was established for logistic-binomial models for distinguishable data (LBDD) by comparing the fit of the models with and without the random effects term (Atwill et al., 1995). Here, the square root of LRS was compared to a one-tailed normal distribution. All models were computed using the programme EGRET (Statistics and Epidemiology Research Corporation, SERC, 1988).

2.6. Further statistical methods

The distribution of ELISA values for subsamples was described by notched box-plots (STATGRAPHICS, Manugistics). The non-parametric Kruskal-Wallis test was used to

assess whether the ELISA values of different age groups come from the same population (STATGRAPHICS, Manugistics). The chi-square test was used to compare proportions. A kappa (κ) statistic was used to compare parasitological and serological results (BIAS; Ackermann, 1992). Relative measures of test accuracy were established using TRYPS as reference method (sensitivity = $TP/(TP + FN)$, specificity = $TN/(TN + FP)$, efficiency = $(TP + TN)/(TN + TP + FN + FP)$; with TN, TP, FN, FP indicating the frequency of true-negative, true-positive, false-negative and false-positive results, respectively).

3. Results

3.1. Parasitology and Ab-ELISA

The parasitological prevalence was 17.9% (95% binomial confidence interval 14.6–21.6%) in the total sample ($n = 487$). The trypanosome species involved in apparent infections were *Trypanosoma brucei* (84.6%) and *T.vivax* (15.4%). The prevalence was higher ($P < 0.001$) in local breeds (29.3%, 18.1–42.7%) than in exotics (10.6%, 6.4–16.3%) but not different between sexes ($P = 0.658$) and age groups ($P = 0.603$). The overall seroprevalence estimate based on a conventional cut-off (SERORESP) was 77.2% (73.1–81%), whereas the proportion of high responders according to mixture-distribution analysis (SEROSUSP) was only 45.3% (40.7–49.9%). According to age-specific cut-off values (SEROAGE), the overall seroprevalence was 43.9% (39.4–48.7%). A support size of two subpopulations in the mixture-distribution analysis (see Appendix A) was a consistent finding for all age strata as well as for the combined sample. Seroprevalence apparently increased with age; the distribution of ELISA values in the five age groups is shown in Fig. 1. ELISA values were different between age groups for unadjusted data and after stratification for TRYPS (Kruskal-Wallis, all $P < 0.001$).

The level of agreement beyond chance was poor for the comparison of parasitological

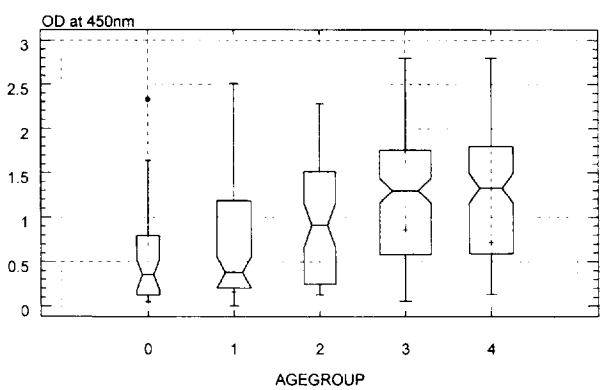

Fig. 1. Box-plots of *Trypanosoma* antibody ELISA results (OD values) for five age groups of cattle (total $n = 457$) from Mukono County, Uganda (1994). The height of the boxes indicates the interquartile range (IQR), whereas the base width is correlated to the square root of the respective sample size. The vertical notch width represents 95% CI of the median value; the mean is indicated (+).

Table 1
Relative measures of test accuracy of an *Trypanosoma* antibody ELISA applied to cattle in Mukono County, Uganda (1994, 457 cattle) [a]

Interpretation [b]	Sensitivity	Specificity	Efficiency
SERORESP	0.92 (0.84–0.97)	0.26 (0.22–0.31)	0.37 (0.33–0.42)
SEROSUSP	0.63 (0.52–0.74)	0.58 (0.53–0.63)	0.59 (0.55–0.64)
SEROAGE	0.63 (0.52–0.74)	0.60 (0.55–0.65)	0.61 (0.56–0.65)

[a] Point estimates with 95% binomial confidence limits indicated in parenthesis.
[b] Interpretation of the Ab-ELISA results was based on a conventional, intrinsic or age-specific intrinsic cut-off value (SERORESP, SEROSUSP and SEROAGE, respectively).

vs. serodiagnostic results (SERORESP: $\kappa = 0.077$; SEROSUSP: $\kappa = 0.133$; SEROAGE: $\kappa = 0.145$; $P < 0.001$ for all combinations). Considering TRYPS as reference method, a greater relative efficiency of the Ab-ELISA could be obtained when only high responders (according to mixture-distribution analysis) were considered seropositive (37%, 59%, 61% efficiency for SERORESP, SEROSUSP and SEROAGE, respectively; Table 1).

3.2. Biological factors

Fixed-effects models rather than random-effects models were applied since the random term associated with herds was not significant ($P > 0.05$) when logistic-binomial models for distinguishable data (LBDD) were used. Among 13 variables

Table 2
Adjusted odds ratios (OR) and 95% confidence intervals (CI) for potential biological factors for the parasitological and serological outcomes for cattle trypanosomosis, Mukono County, Uganda (1994, 487 cattle) [a]

Variable	TRYPS		SEROPRESP		SEROSUSP		SEROAGE	
	OR	CI	OR	CI	OR	CI	OR	CI
BREED	1.0	0.5– 2.0	1.3	0.7–2.6	1.3	0.8–2.1	1.3	0.8–2.2
SEX	0.3	0.1– 1.1	1.3	0.5–3.3	1.2	0.5–2.9	1.6	0.7–3.5
AGEGROUP	0.9	0.7– 1.2	2.3***	1.7–3.0	1.8***	1.4–2.3	1.6***	1.3–2.0
PCV	0.6***	0.4– 0.8	0.9	0.6–1.2	0.7**	0.5–0.9	0.7**	0.6–0.9
PROPEXOT	0.5*	0.2– 0.9	0.5*	0.2–0.9	0.5**	0.3–0.9	0.6*	0.3–1.0
WATER	0.3*	0.1– 0.9	1.0	0.5–2.4	1.4	0.7–2.8	1.3	0.6–2.4
PASTURE	0.3	0.1– 1.3	1.4	0.7–2.9	1.3	0.6–2.6	1.4	0.7–2.9
FEEDSUPPLY	0.6*	0.4– 1.0	0.9	0.6–1.4	1.0	0.7–1.4	0.9	0.6–1.3
PRETREAT	0.5***	0.3– 0.7	0.9	0.6–1.2	0.9	0.7–1.2	0.8	0.6–1.1
DIMINAZENE	1.3	0.5– 3.1	2.2*	1.1–4.7	1.3	0.7–2.4	1.4	0.8–2.5
HOMIDIUM	3.2	0.2–47.3	1.6	0.3–7.9	0.2	0.0–1.5	0.2	0.0–1.5
ISOMETAM	0.3	0.1– 1.3	0.3	0.1–1.4	1.0	0.3–3.7	1.2	0.3–4.2
TICKCONT	0.7	0.4– 1.1	0.6*	0.4–1.0	0.9	0.6–1.2	0.9	0.6–1.3

$P < 0.05$ (*); $P < 0.01$ (**); $P < 0.001$ (***).
[a] The outcome variables TRYPS, SERORESP, SEROSUSP and SEROAGE refer to the parasitological and serological diagnosis of infection. The variables BREED, SEX, AGEGROUP,..., TICKCONT refer to explanatory variables in the logistic regression models.

included in the model, increases in the variables PCV, PROPEXOT, WATER, FEEDSUPPLY and PRETREAT could be identified as protective factors against the parasitological detection of trypanosomes (Table 2).

Four variables contributed to the explanation of the observed variation of seropositivity according to a conventional cut-off value (SERORESP; Table 2). Increases in AGEGROUP and application of DIMINAZENE were identified as risk factors whereas an increase in PROPEXOT and TICKCONT had protective effects on SERORESP. The effect profile for seropositivity was different when intrinsic cut-off values (derived from mixture-distribution analysis) were used (SEROSUSP). Here, increased AGEGROUP was a risk effect whereas increases in PCV and PROPEXOT were protective. The same pattern was observed for the outcome variable based on age specific intrinsic cut-off values (SEROAGE).

4. Discussion

Our intention was to contribute to an understanding of the biological background that should be considered for the interpretation of sero-epidemiological data for trypanosomosis in the Ugandan peri-urban dairy production systems. For this purpose, we define all factors that in vivo modulate the true level of the analyte (in our case *Trypanosoma* antibodies) as biological factors (BF). Thus, we consider infection as the specific BF (diagnostic principle) whereas – without specific assumptions concerning the physiological pathways – age, sex, breed, pregnancy, nutritional status, previous chemotherapy, vaccination, passive immunisation, immune status and self-cured infection are considered a priori as potential non-specific BFs. Here, the terms specific and non-specific are used in the diagnostic (infection as gold standard) rather than analytical (true analyte level as gold standard) sense. The detection of persisting antibodies after self-cure for example must be considered true-positive in terms of an analytical evaluation but false-positive in terms of a diagnostic evaluation. What we call non-specific BFs is the source of what is described as intra- and inter-individual variation in clinical chemistry (Kringle and Johnson, 1986). In contrast to clinical chemistry our objective is to identify biological variables that can be used to explain variability of test results on a population rather than on an individual level. Other factors may interfere only with the analyte detection (measurement bias) without in vivo modulating the true analyte level (referred to as interfering factors, IF). The distinction between BF and IF is important since only IF can be controlled technically through modifications in the pre-analytical and analytical phase of the test. Some examples of BF and IF are given in Fig. 2. Although of tremendous importance for test design, optimisation and standardisation, IF are beyond the scope of our study. However, it is reasonable to assume that IF may account for unexplained variation in the correlation of BFs and test outcomes. The relevance of BF for optimisation of cut-off values for serodiagnostic tests was recently emphasised (Greiner and Böhning, 1994). We postulate that a test interpretation that accounts for non-specific BFs could improve the reliability of diagnostic decision making. Note that analytical random errors additionally add noise to the relation between test results and state of infection and are addressed with the parameters of test

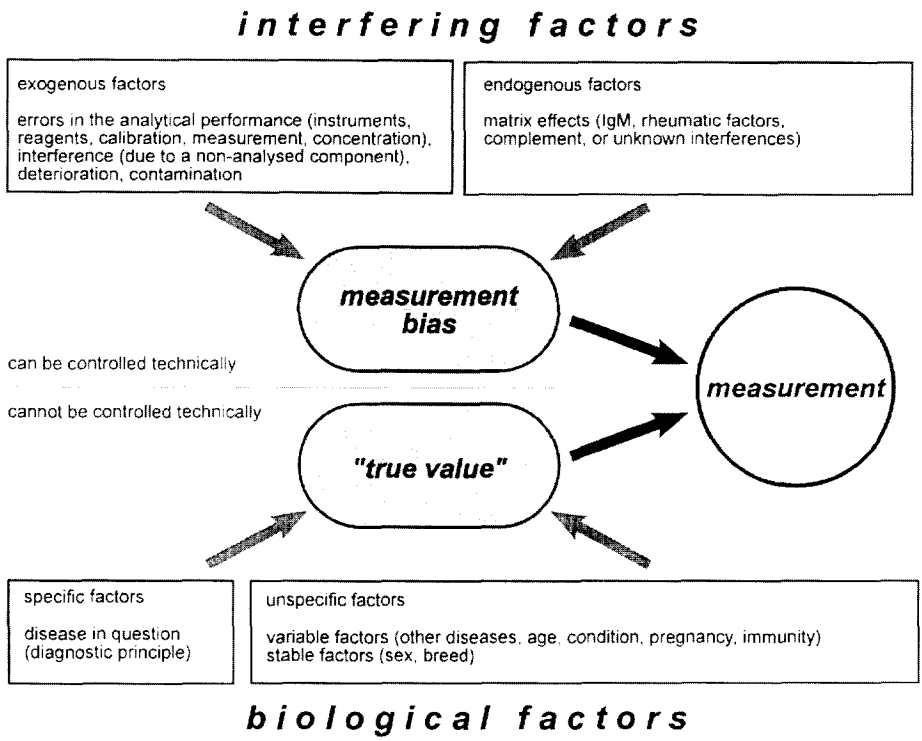

Fig. 2. Impact of biological (BF) and interfering factors (IF) on the accuracy of serodiagnostic methods. Measurement bias and the true value of the analyte (e.g., antibody level) are theoretical concepts involving a set of different components (IF and BF, respectively). Apart from the specific component (state of infection) BF and IF may produce false positive and/or false negative results and, thus, potentially invalidate inferences from test results.

precision. Analytical bias and analytical random errors were also summarised as analytical variation (Kringle and Johnson, 1986).

Putative biological factors for *Trypanosoma* seropositivity were investigated systematically using logistic regression models. It was recently emphasised that mixed-effects models rather than fixed-effects models should be chosen for analysis of clustered (e.g., herd) data (Kristula et al., 1992, Curtis et al., 1993, McDermott et al., 1994, Atwill et al., 1995). Since a considerable set of factors that operate on farm level were included in the models used here and due to the low contagiousness of trypanosomosis, a strong intra-herd correlation was not expected and not detectable when we formally tested for it.

The comparison of adjusted odds ratios from fixed effects logistic regression models for the parasitological and serological outcome variables indicated that seroprevalence but not parasitological prevalence increases with age. The levels of *Trypanosoma* antibodies apparently increased with age. It could be inferred therefore that non-specific (not associated with current infection) antibody titers occurred more often in older age groups than in younger animals. A potential bias due to an (unobserved) decrease in the

diagnostic sensitivity of the parasitological detection method with age cannot be excluded. However, an age effect on seropositivity can be inferred from a study by Dwinger et al. (1988). Also in human trypanosomosis (sleeping sickness), age is described as an important factor for antibody detection (Kegels et al., 1992). Increased seroprevalences in older age groups may be partly due to increased false-positive proportions in these strata. The probability of a recent abortive or self-cured infection increases with age in endemic situations. Maternal antibodies were described in the case of bovine trypanosomosis (Dwinger et al., 1992). Therefore, a certain proportion of elevated antibody levels in cattle below the age of 3 months may be false-positive. In their study of the epidemiology of bovine trypanosomosis in Ghibe valley, Ethiopia, Rowlands et al. (1993) reported an age effect also for the parasitological prevalence. The predominant trypanosome species of that region is *T. congolense*, however, which was not detected in our study.

A further aspect of the study was the comparison of methods for Ab-ELISA interpretation (conventional vs. intrinsic cut-off). The approach of mixture-distribution analysis as described previously (Greiner et al., 1994) was extended by age-specific intrinsic cut-off values (SEROAGE) since an age effect for Ab-ELISA results was already assumed from the results of univariable analysis (Fig. 1). The comparison between parasitological and serological results gave evidence for an increased relative diagnostic efficiency (Table 1) for the mixture-distribution approach. The test efficiency was maximal when age-specific intrinsic cut-off values were used (SEROAGE) – but not importantly different from the mixture-analysis approach for the entire sample (SEROSUSP). Since the parasitological diagnosis – due to its limited sensitivity – cannot be considered as the gold standard, the efficiencies reported in Table 1 can only be regarded as relative measures. Therefore, these results were additionally expressed in terms of the kappa index with the conceptional advantage of not considering any of the methods as the gold standard. From the empirical data, we concluded that the optimal cut-off value for the concerned test is not a fixed threshold value but a set of flexible, age-group-specific operating points (intrinsic cut-off values increased with age; data not shown). Further evidence for the appropriateness of intrinsic cut-off values can be inferred from the adjusted odds ratio for the packed cell volume (PCV). Interestingly, an increase in PCV was associated with a protective effect for the parasitological outcome variable as well as with the serological outcome variables as far as the latter were defined by mixture-distribution analysis (SEROSUSP, SEROAGE). If we consider the PCV value as marker for clinically-apparent trypanosomosis, we can hypothesise that high responder status (SEROSUSP or SEROAGE) – rather than seropositivity defined using an arbitrarily selected cut-off (SERORESP) – can be regarded as a predictor of apparent trypanosomosis.

Empirical data were not available to explain the protective effect of increased proportions of exotic animals in the herds (PROPEXOT) that was identified consistently for both parasitological and serological outcome variables. It is known that exotic breeds are more susceptible to patent trypanosomosis than local breeds. Obviously, this finding is confounded by unobserved management factors. Sample truncation due to high mortality rates in exotic animals appears to be unlikely – given the fact that also the crude (parasitological) prevalence in exotics was less than in local breeds. The access of

animals to shores of rivers and lakes (WATER) appeared to be a protective rather than a risk factor against patent trypanosomosis (TRYPS) which was surprising, given the fact that the study area is infested with riverine tsetse species. Hence, the variable WATER is clearly not a direct measure of infection challenge. In contrast, the protective effect of increasing amounts of FEEDSUPPLY for patent trypanosomosis can easily be explained by the nutritional benefit for animal condition. According to the data, shorter time intervals of trypanocide application (PRETREAT) are associated with a protective effect on patent trypanosomosis, whereas application of the trypanocide DIMINAZENE appears to indicate a paradoxical risk effect for seropositivity according to the conventional cut-off value (SERORESP; the most sensitive definition of seropositivity). It can be assumed that application of diminazene by the farmers reflects the clinical relevance of the disease on the farm. Positive association with the serodiagnostic variable but not with the parasitological variable may be indicative for an effective chemotherapy. The protective effect of tick control measures (TICKCONT) on seropositivity (SERORESP) could be attributed to an efficiency of acaricides for tsetse control (Fox et al., 1993, Okello-Onen et al., 1994), but may be additionally confounded by other management factors. It should also be considered that farm-specific variables are potentially distorted due to recall bias.

From our results we conclude that a conventional cut-off value (SERORESP, established using non-endemic controls) strongly overestimated the seroprevalence and was therefore not suitable for diagnostic purposes. The intrinsic cut-off values (SEROSUSP, SEROAGE, established by mixture-distribution analysis) provided more conservative estimates of seroprevalence. In statistical terms, intrinsic cut-off values can be assumed to provide unbiased estimates of the proportion of animals with elevated antibody levels as far as bimodally distributed data are concerned. It should be stressed that due to the lack of a suitable gold standard, representative reference populations from the study area were not available that would be the precondition to fully evaluate the serodiagnostic interpretations.

5. Conclusion

In a cross-sectional seroepidemiologic study on bovine trypanosomosis in Mukono County, Uganda, age was identified as major biological factor for *Trypanosoma* antibody detection by ELISA. The data suggest that antibody titers not associated with current infection increase with age. Mixture-distribution analysis was used as a strictly data-based approach for selecting intrinsic cut-off values in the absence of reference populations. This approach could be shown to accommodate both unstratified and stratified (according to age) analysis of distribution heterogeneity and the generation of common and age specific intrinsic cut-off values, respectively. Although a better test efficiency for age specific as compared to common intrinsic cut-off values failed to be statistically significant, the comparison with a conventional cut-off value (established using a non-endemic reference population) gave evidence for a superior relative test efficiency when applying mixture distribution cut-off values. The comparison of the profiles of empirical weights for explanatory variables between comparative logistic

regression models for serodiagnostic and parasitological outcome variables was found to be a tool for identifying biological factors for serodiagnosis and gave further evidence for the appropriateness of intrinsic cut-off values. The findings support earlier reports of an overestimation of seroprevalence when non-representative negative populations are used for establishing cut-off values (Franke et al., 1994, Greiner et al., 1994).

Acknowledgements

We acknowledge the constructive comments and suggestions of the reviewers and, especially, of the Editor-in-Chief that improved the paper's form and content. Furthermore, we thank Dr. M. Eisler, University of Glasgow, for the helpful comments. U. Tietjen's skilful handling of trypanosome cultures and the technical assistance of Mrs. I. Kyriakopoulos, Dipl.-Biochem. is appreciated. We thank the farm owners, managers and herdsmen for their co-operation. A part of the study was supported by the German Federal Ministry For Economic Co-operation (BMZ; PN 94.7860.3-01.100).

Appendix A. Mixture-distribution analysis (definition of SEROSUSP and SEROAGE)

The mild assumption made for modelling heterogeneity in the ELISA data was that the sample is coming from an unknown number k of subpopulations resulting in a density function f that can be represented as a mixture-distribution density $p_1 f(x,\vartheta_1) + \ldots + p_k f(x,\vartheta_k)$ of k unobserved normal densities with the subpopulations' means $\vartheta_1,\ldots,\vartheta_k$ and subpopulations' weights $p_1,\ldots,p_k$. The variance parameter σ of the normal distribution is not of specific interest here (nuisance parameter) but must be estimated from the data. The number of subpopulations of the mixture model was assumed to be unknown (flexible support size) although the a priori assumption of two subpopulations of ELISA-data (low and high responders) may be plausible. However, a strictly data-oriented approach was required to estimate the support size from the data. A well-motivated method for estimation of the model parameters ($\vartheta_1,\ldots,\vartheta_k$; $p_1,\ldots,p_k$; $\sigma_1,\ldots,\sigma_k$) is maximum likelihood (Dempster et al., 1977, Laird, 1978). The maximum likelihood estimators are found as those parameters that maximise the likelihood

$$\prod_{i=1}^{n}\left[\sum_{j=1}^{k} f(x_i,\vartheta_j)\,p_j\right]$$

The computer algorithms used to find the form estimators are described in detail elsewhere (Böhning et al., 1992 and the references given therein). If the flexible support size approach indicates a support size of 2 (two subpopulations), the parameter estimates can be used in the fixed support size approach. The maximum likelihood estimates can then be used to construct a threshold value (x_0, referred to as intrinsic cut-off) that differentiates between the two subpopulations such that Eq. (1) holds.

$$p_1\int_{x_0}^{1-\infty} f(x,\vartheta_1)\,\mathrm{d}x = p_2\int_{-\infty}^{x_0} f(x,\vartheta_2)\,\mathrm{d}x \qquad (1)$$

The intrinsic cut-off, therefore, is chosen to provide an unbiased estimate of seroprevalence if a single decision threshold is to be used to classify the objects of the mixture-distribution as belonging to one of the subpopulations (more details are given by Böhning et al., 1992 and Greiner et al., 1994). The programme C.A.MAN is available upon request (please send a preformatted diskette).

References

Ackermann, H., 1992. Analyse von Stichproben. Validierung und Beispiele. Epsilon-Verlag, Hochheim-Darmstadt, Germany.

Atwill, E.R., Mohammed, H.O., Scarlett, J.M. and Mcculloch, C.E., 1995. Extending the interpretation and utility of mixed effects logistic regression models. Prev. Vet. Med. 24: 187–201.

Bhat, T.S., 1994. Seroepidemiological investigations on *Trypanosoma* specific antibody levels in dairy cattle in Mukono County, Uganda. M.Sc. Thesis in Epidemiology and Preventive Veterinary Medicine, Free University of Berlin, Berlin, Germany.

Böhning, D., Schlattmann, P. and Lindsay, B., 1992. Computer-assisted analysis of mixtures (C.A.MAN): statistical algorithms. Biometrics, 48: 283–303.

Curtis, C.R., Mauritsen, R.H., Kass, P.H., Salman, M.D. and Erb, H.N., 1993. Ordinary versus random effects logistic regression for analyzing herd-level calf morbidity and mortality data. Prev. Vet. Med., 16: 207–222.

Dempster, A.P., Laird, N.M. and Rubin, D.B., 1977. Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). J. R. Statist. Soc. Series B, 39: 1–38.

Dwinger, R.H., Grieve, A.S., Jeannin, P., Agyemang, K. and Faye, J., 1988. Anti-trypanosomal antibodies in sequentially collected sera of N'dama cattle under natural trypanosomiasis risk in The Gambia. Proceedings of an ILCA/ILRAD Meeting on Livestock Production in Tsetse affected Areas of Africa, ILCA/ILRAD, 1987, Nairobi. pp. 100–109.

Dwinger, R.H., Grieve, A.S., Snow, W.F., Rawlings, P., Jabang, B. and Williams, D.J.L., 1992. Maternal antibodies in N'dama calves kept under natural trypanosomiasis risk in The Gambia. Parasite Immunol., 14: 351–354.

Fox, R.G.R., Mmbando, S.O., Fox, M.S. and Wilson, A., 1993. Effect on herd health and productivity of controlling tsetse and trypanosomosis by applying deltamethrin to cattle. Trop. Anim. Health Prod., 25: 203–214.

Franke, C.R., Greiner, M. and Mehlitz, D., 1994. Investigations on naturally occurring *Trypanosoma evansi* infections in horses, cattle, dogs and capybaras (*Hydrochaeris hydrochaeris*) in Pantanal de Pocone (Mato Grosso, Grazil). Acta Trop., 58: 159–169.

Greiner, M. and Böhning, D., 1994. Notes about determining the cut-off value in enzyme linked immunosorbent assay (ELISA) - reply. Prev. Vet. Med., 20: 307–310.

Greiner, M., Franke, C.R., Böhning, D. and Schlattmann, P., 1994. Construction of an intrinsic cut-off value for the sero-epidemiological study of *Trypanosoma evansi* infections in a canine population in Brazil: a new approach towards an unbiased estimation of prevalence. Acta Trop., 56: 97–109.

Kegels, G., Criel, van Lerberghe, W., Mentens, H., Magnus, E. and van Balen, H., 1992. Screening for *Trypanosoma brucei gambiense* antibodies with the indirect fluorescent antibody test (IFAT). Ann. Soc. Belg. Med. Trop., 72: 271–281.

Kringle, R.O. and Johnson, G.F., 1986. Statistical Procedures. In: N.W. Tietz (Editor), Textbook of clinical chemistry, W.B. Saunders, Philadelphia, PA, USA, pp. 287–355.

Kristula, M.A., Curtis, C.R., Galligan, D.T. and Bartholomew, R.C., 1992. Use of a repeated-measures logistic regression model to predict chronic mastitis in dairy cows. Prev. Vet. Med., 14:57–68.

Laird, N., 1978. Nonparametric maximum likelihood estimation of a mixing distribution. J. Am. Statist. Assoc., 73: 805–811.

McDermott, J.J., Schukken, Y.H. and Shoukri, M.M., 1994. Study design and analytic methods for data collected from clusters of animals. Prev. Vet. Med., 18: 175–191.

Mehlitz, D. and Tietjen, U., 1988. Trypanosome infection rates in tsetse midguts using a short-term in vitro culture technique. Acta Trop., 45: 183–184.

Okello-Onen, J., Heinonen, R., Ssekitto, C.M.B., Mwayi, W.T., Kakaire, D. and Kabarema, M., 1994. Control of tsetse flies in Uganda by dipping cattle in deltamethrin. Trop. Anim. Health Prod., 26: 21–27.

Rowlands, G.J., Mulatu, W., Authie, E., D'Ieteren, G.D.M., Leak, S.G.A., Nagda, S.M. and Peregrine, A.S., 1993. Epidemiology of bovine trypanosomiasis in the Ghibe valley, southwest Ethiopia. 2. Factors associated with variations in trypanosome prevalence, incidence of new infections and prevalence of recurrent infections. Acta Trop., 53: 135–150.

Statistics and Epidemiology Research Corporation and Cytel Software Corporation., 1988. Epidemiological graphics, estimation, and testing package (EGRET). Analysis module (PECAN; Version 25.6), Washington, USA.