 doi:10.1111/j.1365-3156.2006.01743.x

VOLUME 11 NO 12 PP 1777–1784 DECEMBER 2006

Comparative evaluation of freeze-dried and liquid antigens in the direct agglutination test for serodiagnosis of visceral leishmaniasis (ITMA-DAT/VL)

Diane Jacquet[1], Marleen Boelaert[2], Jill Seaman[3], Suman Rijal[4], Shyam Sundar[5], Joris Menten[2] and Eddy Magnus[1]

1 *Applied Technology and Production Unit, Institute of Tropical Medicine, Antwerp (ITMA), Antwerp, Belgium*
2 *Department of Public Health, Institute of Tropical Medicine, Antwerp (ITMA), Antwerp, Belgium*
3 *International Medical Relief Fund (IMRF), Watsonville, CA, USA*
4 *Department of Internal Medicine, BP Koirala Institute of Health Sciences (BPKIHS), Dharan, Nepal*
5 *Institute of Medical Sciences, Banaras Hindu University (IMSBHU), Varanasi, India*

Summary

OBJECTIVE The direct agglutination test (DAT) for visceral leishmaniasis (VL) with liquid (LQ) antigen is known to be only moderately reproducible because of inter-observer and batch-to-batch variability as well as its sensitivity to temperature and shaking during transport. We evaluated a DAT with freeze-dried (FD) antigen and compared it with the LQ antigen version.
METHODS Blood samples of clinical VL suspects and healthy endemic controls were collected in Sudan, Nepal and India. Both test versions were performed in duplicate in the respective countries and in the reference laboratory. Interbatch variability and stability tests were conducted and agreement was examined within and between centres on a dichotomic scale by Cohen's kappa as well as on a continuous scale through Bland–Altman plots.
RESULTS The FD antigen remains fully active even after storage at 45 °C for 24 months. Using a cut-off titre of 1 : 6400, the agreement between the FD and the LQ formats was excellent.
CONCLUSION The major advantages of FD antigen are its better stability at higher temperatures and its longer shelf life, which make it much more suitable than the LQ version for use in the field.

keywords visceral leishmaniasis, direct agglutination test, freeze-dried antigen, liquid antigen

Introduction

A direct agglutination test (DAT) for the serodiagnosis of visceral leishmaniasis (VL) was developed by A.E. Harith at the Royal Tropical Institute (KIT), Amsterdam, the Netherlands (Harith *et al.* 1986). Simultaneously and in collaboration with Harith, a slightly modified DAT was developed at the Institute of Tropical Medicine in Antwerp (ITMA), Belgium. In both test versions, the patient's serum was incubated with a suspension of trypsinized, formaldehyde-fixed and Coomassie-blue-stained promastigotes of *Leishmania donovani*. The presence of antibodies in the test serum was demonstrated by agglutination of the parasites. Performance of the DAT with this liquid (LQ) antigen improved by adding 0.8% of 0.1 M 2-mercaptoethanol (2ME) to the sample diluent, while stability of the antigen increased by supplementing the suspension medium with 0.056 M sodium citrate (Harith *et al.* 1988). The diagnostic capacity of the DAT with LQ antigen was consistently reported as very high in laboratory studies (Zijlstra *et al.* 1992). An evaluation of the DAT with LQ antigen showed, however, that while repeatability was excellent at 1 : 500 dilution, reproducibility between field and laboratory was less good (Boelaert *et al.* 1999). This was attributed partly to insufficient training of the laboratory technicians, partly to deterioration of the LQ antigen during transport and storage. Although often quoted as an issue, few data are available on the batch-to-batch variability of the DAT. More generally, reproducibility is an often neglected aspect in the evaluation of diagnostics.

In 1995, Meredith *et al.* (1995) of the KIT in Amsterdam developed a freeze-dried (FD) version of the DAT antigen, and Oskam *et al.* (1999) reported good diagnostic performance of this format. Because of its relative user-friendliness compared with other serological tests and the possibility to perform the DAT on dried blood samples collected on filter paper, the test became crucial in the control of VL outbreaks in southern Sudan (Seaman *et al.* 1996) and Somalia (Marlet *et al.* 2003).

Several attempts were made to transfer the technology of antigen production to centres in disease-endemic countries,

© 2006 Blackwell Publishing Ltd

but sustainability of this production has remained problematic. In Europe, there are, today, two non-commercial suppliers of DAT antigen: the KIT in Amsterdam and ITMA in Antwerp. Up till now, the KIT was the unique supplier of FD antigen. In order to back up the continuity of supply, the Applied Technology and Production Unit at the ITMA decided to develop a FD version of the DAT, as well. A prototype kit with FD antigen was ready by mid-2002 and showed excellent test results in comparison with the LQ form when conducted on well-characterized sera of kala-azar patients and negative controls. Preliminary stability tests yielded satisfactory results. However, this type of evaluation tends to give an overly optimistic view of the reproducibility of the FD DAT, as differences in titres are clinically most relevant in the spectrum of 1 : 800 to 1 : 12 800, the so-called 'borderline' titres. Therefore, the next step in the evaluation was to design a study that would include a sufficient number of borderline patients with intermediate titres. We conducted this evaluation with partners in Sudan, Nepal and India and report here on the performance of ITMA's FD DAT compared with ITMA's LQ DAT. The aim of our study was to conduct a thorough assessment of the reproducibility of the FD DAT, including its stability, its batch-to-batch variability, its agreement with the LQ DAT (between-format comparison) as well as the within-centre and between-centre reproducibility.

Materials and methods

Patients and samples

In Sudan, 200 clinically suspected kala-azar cases and 100 healthy endemic controls were consecutively recruited from Keew Central Upper Nile by the International Medical Relief Fund (IMRF) between October 2003 and February 2004. Of these 300 samples, 20 were excluded from the study for the following reasons: 15 for missing field results, 3 samples were contaminated, 1 because of a doubtful LQ DAT result and 1 filter paper missing at ITMA. The remaining 280 samples were tested in Sudan and at ITMA and included in the data analysis.

In Nepal, 118 confirmed kala-azar cases and 104 healthy persons were consecutively recruited by the BP Koirala Institute of Health Sciences (BPKIHS) between March 2003 and April 2004. The samples of all of the 222 persons enrolled were tested at ITMA and at BPKIHS and used for data analysis. Moreover, 49 samples collected during a community survey and showing borderline LQ DAT titres were included in the panel, bringing the total up to 271 samples. Unfortunately, the 49 borderline samples could only be tested at ITMA.

All persons were examined clinically and blood samples were collected in duplicate on filter paper (Whatman no. 3). The blood-filter papers were dried at ambient temperature and stored in sealed plastic bags containing silica gel as a desiccant. The duplicate filter papers were sent to ITMA.

In India, 150 confirmed kala-azar cases, 153 endemic controls, 100 healthy controls, 105 patients with other diseases and 10 post-kala-azar dermal leishmaniasis patients from the Kala-Azar Medical Research Center, Institute of Medical Sciences, Banaras Hindu University (IMSBHU), at its two sites at Varanasi and Muzaffarpur, were recruited in 2004. Blood was collected, serum samples prepared and stored at −30 °C until use. No duplicate samples were sent to ITMA.

Antigen

The antigen was prepared at the ITMA in the Applied Technology and Production Unit. Promastigotes of *L. donovani* 1-S strain were cultured in glucose–lactalbumin–serum–haemoglobin medium, harvested in the log phase at a concentration of 1×10^7–2×10^7 parasites/ml by centrifugation at 1590 ***g*** at 2 °C for 20 min and washed twice in phosphate–glucose–saccharose (PGS) buffer solution, pH 8.0–8.1. Digestion of surface proteins was carried out with 0.4% trypsin for 60 min at 37 °C with gentle orbital shaking (100 rpm). After trypsinization, the parasites were washed in PGS. A 1% suspension (vol/vol) was prepared in cold PGS and fixed with 4% formaldehyde for 20 h at 4 °C. Excess fixative was removed through repeated washing with phosphate-buffered saline (PBS), pH 7.2. The organisms were stained with 0.025% Coomassie brilliant blue in PBS containing 0.034 M sodium citrate for 20 h at 4 °C. A final 5×10^7 parasites/ml suspension was prepared in formaldehyde–citrate saline. The volume of each batch produced was between 1000 and 2000 ml.

This suspension was then divided into two equal volumes: one volume was used as LQ antigen. The other volume was centrifuged, resuspended at a concentration of 1.25×10^8 parasites/ml in lyophilization medium, freeze-dried in 1 ml volumes per vial and used as FD antigen. Both LQ and FD antigens were tested against a panel of well-characterised reference sera.

Buffer and diluent

The buffer for the LQ DAT contains saline, gelatine and 0.1% sodium azide. The buffer for the FD DAT is composed of PBS, pH 7.2, supplemented with protein and 0.1% sodium azide. In both test versions, the diluent is prepared by adding of 0.1 M 2ME to the respective buffer solution.

 © 2006 Blackwell Publishing Ltd

Control sera

The positive control was a diluted, FD kala-azar-positive serum with a DAT titre of about 1 : 102 400. The negative control was a kala-azar-negative serum with a DAT titre below 1 : 200. Both control sera had been screened for HIV and hepatitis B.

All reagents for both LQ and FD test versions (antigen, buffer, diluent, positive control, negative control and 2ME solution) were sent to the IMRF in Sudan, to the BPKIHS in Nepal and to the IMSBHU in India. A part of the same batches was kept at ITMA to test duplicate samples. For India, no duplicate tests were performed in Antwerp. The reagents were supplied as a test kit, designed to suit field and laboratory conditions together with all accessory materials and the protocols for the different test versions.

Test execution

In Sudan, Nepal and India, all samples were tested with the FD DAT and LQ DAT versions and read by the same person. In the same way, the duplicate blood samples collected in Nepal and Sudan were tested at ITMA without prior knowledge of the country's results.

Eluates of blood-filter papers (Sudan, Nepal and ITMA). blood protein eluates were prepared from the filter papers in V-shaped microplates (Greiner, 96-well). From each blood spot, a 5-mm disc was punched out and placed in a well of the microplate. Then, 125 μl of buffer was pipetted into each well. The plate was covered to avoid evaporation and incubated overnight at 4 °C. Assuming a mean haematocrit value of 50%, a 5-mm disc of blood-impregnated filter paper (Whatman no. 3) contains approximately 5 μl blood or 2.5 μl serum. The eluate in 125 μl of DAT buffer thus corresponds to a serum dilution of 1 : 50.

From the blood eluates, serial twofold dilutions of 1 : 100 up to 1 : 51 200 were then made in a freshly prepared diluent.

Dilution of serum samples (India). Serum samples were thawed and homogenized by gently mixing. Serial twofold dilutions of 1 : 100 up to 1 : 51 200 were made in a freshly prepared diluent in V-shaped microplates.

The ready-for-use LQ antigen just needed gentle shaking to homogenize the suspension. The FD antigen was reconstituted by adding 2.5 ml of buffer to a vial that was gently shaken before use. The tests with LQ and FD antigens were carried out simultaneously. To every fifth plate, a positive and a negative control serum were added. For each test sample, an 'antigen control', containing 50 μl of diluent per well, was included. Fifty microlitres of the respective antigens were pipetted into the wells containing the test sample dilutions and to the control wells. The microplates were covered and incubated overnight at ambient temperature.

Results were read the next day. Agglutination patterns in each individual well were scored as negative (small, sharp, dark blue dot identical to the antigen control) or positive (pale blue film covering the entire bottom of the well or dot, larger than the antigen control). The highest dilution with a positive score was considered as the end titre.

We explicitly draw the attention of the reader to the fact that the expression of the dilution in this article takes into account the volume (50 μl) of antigen added to the sample dilution (50 μl). Accordingly, a 1 : 200 dilution in this article corresponds to a 1 : 100 dilution expressed by El Harith *et al.* The 1 : 3200 dilution recommended as a cut-off value by El Harith *et al.* (serum dilution before addition of antigen), therefore, corresponds in this article to 1 : 6400.

Interbatch variability

Based on the results obtained with the LQ DAT, we selected 33 positive, 53 negative and 19 borderline, samples among the filter papers collected in Sudan and Nepal. For the positive filter papers, preference was given to samples with titres <1 : 204 800 to allow for the detection of differences in end titre between the different batches of antigen. The test was performed by one person on four different batches of FD antigen (071002, 03E1, 03G1 and 04B1), but always using the same batches of diluent, buffer and disposable materials. Consequently, the only varying factor in this experiment was the batch of FD antigen.

For practical reasons, the interbatch variability experiment was only performed on blood-impregnated filter papers. Eluates from two (5 mm) discs, punched out of each filter paper, were prepared in 250 μl of buffer. This quantity was sufficient to test the four different batches of FD antigen. Tests were performed as described earlier.

Stability tests

Direct agglutination test FD reagents (antigen, buffer, diluent, positive control and negative control) were stored for up to 2 years at three different temperatures (+4, +37 and +45 °C). The reagents were tested against a panel of reference sera after 6, 12, 18 and 24 months. The FD antigen was reconstituted and tested the same day. The reconstituted antigen and other reagents were stored at +4 °C and tested over a period of 4 weeks. Simultaneously, a limited stability test under fluctuating temperatures at weekly intervals over a period of 6 months was performed.

© 2006 Blackwell Publishing Ltd

Data analysis

Between-format and between-laboratory agreement was assessed by comparing the LQ and FD DAT results obtained at the local laboratories and at the reference lab (ITMA) using the samples collected in Sudan, Nepal and India (local laboratory only). DAT serial titre dilutions were transformed to a discrete numerical scale coded 0–10, with a titre <1 : 200 corresponding to 0 and a titre ⩾1 : 102 400 corresponding to 10. To assess the agreement on a numerical scale, data were plotted as titre difference $(x - y)$ against the mean titre $(x + y)/2$ according to Bland and Altman (1986). Systematic error was estimated by the mean difference of all paired titre differences $x_i - y_i$. Random error, or the variability in the differences between the paired observations, was estimated by the limits of agreement, indicating the range where 95% of all inter-observer differences can be expected. The limits of agreement are computed as the mean difference ± 1.96 SD, assuming differences are normally distributed. Bland and Altman plots for assessment of the between-format agreement are presented for the Sudanese samples. Figures for the Indian and Nepalese samples and for assessment of between-laboratory agreement are available on request from the corresponding author.

Direct agglutination test results were categorized as 'positive' and 'negative' based on a cut-off titre of 1 : 6400. Agreement on this binary scale was assessed using Cohen's kappa coefficient (Cohen 1968). Kappa coefficients were interpreted following Landis and Koch (Landis & Koch 1977): 1.00–0.81 excellent, 0.80–0.61 good, 0.60–0.41 moderate, 0.40–0.21 weak and 0.20–0.00 negligible agreement.

For assessment of the interbatch variability, 105 samples were tested using four different batches of the DAT FD antigen at ITMA. Results were expressed as DAT serial titre dilutions on a discrete numerical scale coded 0–11, with a titre <1 : 200 corresponding to 0 and a titre ⩾1 : 204 800 corresponding to 11.

In an extension to the methods of Bland and Altman (Bland & Altman 1986), the deviations from the mean (over the four batches) were plotted against this overall mean (analysis not shown). In addition, these deviations from the overall mean were presented using box plots. For each batch, the systematic error was estimated by the mean deviation from the overall mean over the four batches. Limits of agreement *vs.* the overall mean were estimated for each batch. These limits of agreement *vs.* the overall mean indicate the range from the overall mean, in which 95% of the data using an individual batch can be expected. These limits of agreement *vs.* the overall mean correspond to the limits of agreement between two different batches of FD antigen divided by the square root of 2, which is equal to 1.4.

To assess the impact of the variability in test results on clinical diagnosis of VL, the percentage of samples testing positive or negative were estimated for cut-off titres varying from 1 : 200 to 1 : 102 400. Agreement between the batches on a binary scale based on a cut-off titre of 1 : 6400 was assessed using Cohen's kappa, as described earlier. In addition, the average agreement over the four batches was estimated using Light's kappa (Conger 1980).

Confidence intervals for Cohen's kappa and Light's kappa were calculated using the bias-corrected and accelerated bootstrap method (Efron & Tibshirani 1993) using 9999 replications.

All data were entered in a Microsoft Excel format and analysed with R 2.1.0 (R Development Core Team 2005).

Ethical aspects

The protocol for this study was approved by the ethical committee of the ITMA, Belgium. For Nepal, ethical clearance was obtained from the ethical review board at BPKIHS. In Sudan, clearance was given by the health authorities of the region (Dr Bellario, Secretariat of Health). Informed consent was asked from each person before inclusion in the study. Kala-azar cases were treated under the responsibility of the clinicians in charge on the basis of parasitology and/or local LQ DAT results, in accordance with the current WHO guidelines. All other patients had access to standard care at the hospital facility. For the Indian study, ethical clearance was obtained from the ethical committee of the IMSBHU, Varanasi, India.

Results

Comparison of liquid and freeze-dried antigens

Agreement between results obtained with FD DAT and LQ DAT was generally good, both at the reference laboratory and at the local laboratories. Identical titres were obtained with FD DAT and LQ DAT for 69% of samples at the reference laboratory and for 77% of samples at the country laboratories. Less than 10% of samples gave results differing by more than two titration steps.

Examining the difference between FD DAT and LQ DAT at ITMA, the limits of agreement (Bland & Altman 1986) were −2.9 to +1.7 titration steps for the Sudanese samples (Figure 1a) and −1.6 to +1.9 for the Nepalese samples. At the country laboratories, the limits of agreement were −2.2 to +2.0 for the Sudanese samples (Figure 1b) and −3.7 to +3.3 for the Indian samples. Limits of agreement from the local laboratory analysis of

 © 2006 Blackwell Publishing Ltd

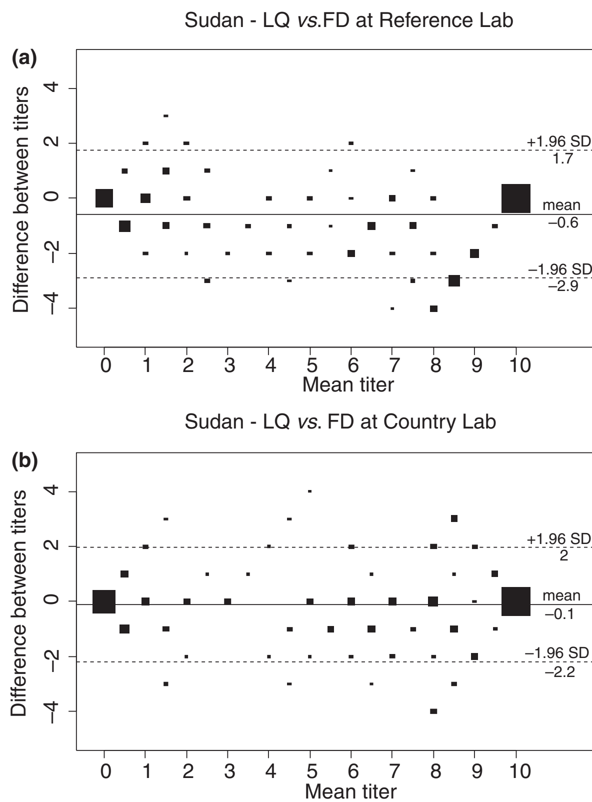

Figure 1 Differences of the titres against their average for the FD DAT and LQ DAT results for the Sudanese samples. (a) Evaluated at the reference laboratory (ITMA). (b) Evaluated at the country laboratory (Sudan).

Nepalese data were not calculated, because borderline samples were not included in the series tested at the country lab.

The size of each box indicates the number of observations represented. Titres are on a 0 (dilution <1 : 200) to 10 (dilution ⩾1 : 102 400) scale. A horizontal line indicates the average difference in titres. Dotted lines indicate limits of agreement.

When assessed on a binary scale (test positive/negative) using the recommended cut-off titre of 1 : 6400, the agreement between the FD and LQ formats was excellent, both at the reference and local laboratories (Table 1). Kappa values for comparison between FD and LQ formats at the ITMA were 91.1% for the Sudanese samples and 99.3% for the Nepalese samples. At the local laboratories, kappa values were 91.1% in Sudan, 100% in Nepal and 84.6% in India.

Interbatch variability

A total of 105 serum samples were selected for inclusion in the assessment of interbatch variability of the FD antigen. For 30 samples (29%), the resulting titres using the four batches agreed exactly. For 46 samples (46%), individual titres were all within one titre of the overall mean result. For the remaining 27 samples (26%), at least one result differed by more than one titre from the overall mean for the sample.

How the four different batches compare to the average titre of the four batches is shown in Figure 2. Batch 04B1 showed generally titres above the overall mean, while batch 071002 gave results generally below the overall mean. The limits of agreement of an individual batch and

Table 1 Agreement between FD DAT and LQ DAT on a binary scale (test positive/negative) using a cut-off titre of 1 : 6400

	No. of samples	No. of patients†				Kappa	95% CI
		FD/LQ (+/+)	FD/LQ (+/−)	FD/LQ (−/+)	FD/LQ (−/−)		
Reference lab							
Sudan	280	161	10	2	107	91.1	84.8–94.9
Nepal	271	139	1	0	131	99.3	96.3–100.0
Total	551	300	11	2	238		
Country lab							
Sudan	280	161	8	4	107	91.1	84.9–95.0
Nepal‡	222	114	0	0	108	100.0	100.0–100.0
India	518	161	23	13	321	84.6	79.2–88.9
Total	1020	436	31	17	536		

FD, freeze-dried; LQ, liquid; DAT, direct agglutination test.
†Number of patients with a positive (+) or negative (−) test based on FD or LQ antigen.
‡Incomplete data set: borderline sera were not included.

© 2006 Blackwell Publishing Ltd

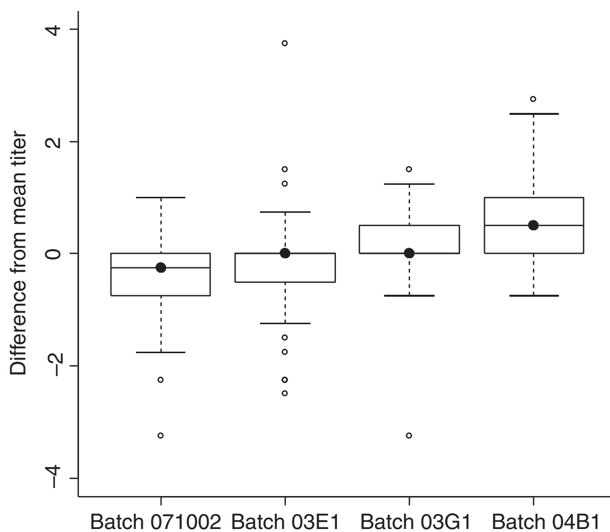

Figure 2 Box plot of the differences of the titres using four different batches of FD antigen compared with their average.

the overall mean over the four batches were ±1.4 titres, indicating that 95% of the test results using a specific batch are expected to lie within 1.4 titres of the overall mean (Table 2). This corresponds to the limits of agreement between two different batches of DAT FD antigen of ±2.0 titres. The box extends from the first to the third quartile and contains 50% of the values. The median is indicated by the filled circle. The whiskers extend to ±1.58 interquartile range divided by square root of n (McGill *et al.* 1978). Pairwise agreement between the four batches of antigen at the recommended dilution cut-off of 1 : 6400 was good to excellent (Table 3). The overall agreement over the four batches was excellent at 81%.

Between-centre comparison of LQ DAT and FD DAT results

For LQ DAT as well as for FD DAT, agreement between results obtained at the country and reference laboratories was assessed for the samples collected in Sudan.

Table 2 Mean difference between titres and their overall mean (n = 105)

	Mean (SD) difference from overall mean	95% limits of agreement *vs.* overall mean
Batch 071002	−0.37 (0.61)	−1.56 to 0.83
Batch 03E1	−0.24 (0.72)	−1.65 to 1.17
Batch 03G1	0.12 (0.54)	−0.93 to 1.17
Batch 04B1	0.49 (0.64)	−0.77 to 1.75
Pooled	0 (0.71)	−1.40 to 1.40

Identical titres were obtained at local and reference laboratories for 54% of samples using LQ antigen and for 57% of samples using FD antigen. Using LQ antigen, 10.7% of samples gave results differing by more than two titration steps between local and reference laboratories *vs.* 8.2% of samples using FD antigen.

Limits of agreement were −3.0 to 2.9 for the LQ antigen and −2.2 to 3.1 for the FD antigen.

When assessed on a binary scale (test positive/negative) using the recommended cut-off titre of 1 : 6400, the agreement between the local and reference laboratories was excellent, both for the LQ format, kappa = 91.2%, and FD format, kappa = 91.0% (Table 4).

Stability tests

The FD antigen, the FD control sera and the buffer/diluent solutions were tested after 6, 12, 18 and 24 months storage at +4, +37, +45 °C and after 6 months storage at temperatures fluctuating between +4 and +45 °C. At +4 °C, all reagents remained stable for 24 months. At 37 °C, the FD antigen and the buffer/diluent solution remained stable for 24 and 12 months, respectively. However, the quality of the control sera decreased within 6 months. At 45 °C, the pellet of the FD antigen was reduced in size, but reactivity was not affected after 24 months, provided that buffer/diluent stored at +4 °C was used for testing. The quality of the control sera and the buffer/diluent decreased within 6 months. Storage at fluctuating temperatures for 6 months resulted in a decrease of 2 titres for high-reactive samples (end titre >1 : 204 800) and of 1 titre for low-reactive or borderline samples.

Discussion

This study shows high agreement between results obtained with LQ DAT and FD DAT, both at the reference laboratory and at the country laboratories. When evaluating test results over the full scale of dilutions, identical titres were obtained with FD DAT and LQ DAT for 69% of the samples at the reference laboratory and 77% of the samples at the country laboratories, whereas less than 10% gave results differing by more than 2 titres. However, in clinical practice, the test is read in a binary way, and reproducibility should therefore be assessed in a 2 × 2 contingency table. At a cut-off titre of 1 : 6400, the agreement between the LQ format and the FD format was excellent. Kappa values for comparison between the LQ and FD formats at the reference laboratory were 91.1% for the Sudanese samples and 99.3% for the Nepalese samples. The higher agreement in Nepalese samples can be explained by the fact that the Sudanese series contained a

 © 2006 Blackwell Publishing Ltd

Table 3 Agreement between results on a binary scale (test positive/negative) for four batches of freeze-dried antigen using a cut-off titre of 1 : 6400 ($n = 105$)

	No. of samples (%)		Pairwise kappa		
	Negative	Positive	Batch 03E1	Batch 03G1	Batch 04B1
Batch 071002	52 (50%)	53 (50%)	86.7%	88.6%	71.3%
Batch 03E1	51 (49%)	54 (51%)		86.6%	73.1%
Batch 03G1	46 (44%)	59 (56%)			78.3%
Batch 04B1	37 (35%)	68 (65%)			
	Overall kappa: 80.8% (95% CI: 71.8–87.7)				

Table 4 Agreement between country and reference results of FD DAT and LQ DAT on a binary scale (test positive/negative) using a cut-off titre of 1 : 6400 (Sudan)

		No. of patients†					
	No. of samples	Cntry/Ref (+/+)	Cntry/Ref (+/−)	Cntry/Ref (−/+)	Cntry/Ref (−/−)	Kappa	95% CI
LQ antigen	280	158	5	7	110	91.2	85.0–95.0
FD antigen	280	164	7	5	104	91.0	84.7–95.1

FD, freeze-dried; LQ, liquid; DAT, direct agglutination test.
†Number of patients with a positive (+) or negative (−) test based on FD or LQ antigen.

higher proportion of borderline titres (50%) compared with the Nepalese (22%). At the country laboratories, kappa values were 91.1% in Sudan, 100% in Nepal (borderline results are missing) and 84.6% in India for the comparison of LQ DAT and FD DAT.

A previous study using LQ antigen (Boelaert *et al.* 1999) showed poor reproducibility between the local laboratory in Sudan and the reference laboratory, partly because of deterioration of the antigen during transport and storage. In the present study, using FD antigen, the agreement between the local laboratory (Sudan) and the reference laboratory is excellent. Using FD DAT clearly solves the problem of uncertain quality due to transport and storage conditions.

The variability between four different batches of FD antigen indicates that 95% of the test results are situated within 1.4 titres of the overall mean, indicating that reproducibility between different FD batches is excellent.

When stored at 45 °C, the FD antigen remains stable for a period of 24 months. Even if the pellet of the FD antigen was reduced in size, reactivity was not affected. The buffer/diluent and the control sera remain stable for, respectively, 12 and 6 months at 37 °C.

The relative limited stability of the latter reagents means no constraint on their applicability under field conditions. In practice, mobile teams go in the field to collect blood samples on filter paper and bring them to a central laboratory for testing. In case the DAT has to be performed under field conditions, reagents can be dispatched to the field from a central storage facility, where they are kept below 10 °C to guarantee maximal stability.

Acknowledgements

We thank Dr K. Schenkel for his contribution in collecting part of the Nepal samples and Mr Ranjith Kumar for testing the Nepalese samples. Our sincere thanks also to Ms S. De Wit for her appreciated assistance in the Sudanese part of the study.

References

Bland JM & Altman GA (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.

Boelaert M, El-Safi S, Mousa H *et al.* (1999) Multi-centre evaluation of repeatability and reproducibility of the direct agglutination test for visceral leishmaniasis. *Tropical Medicine and International Health* **4**, 31–37.

Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* **70**, 213–220.

Conger AJ (1980) Integration and generalisation of kappa's for multiple raters. *Psychological Bulletin* **88**, 322–328.

Efron B & Tibshirani R (1993) *An Introduction to the Bootstrap*. Chapman & Hall, Boca Raton, FL, USA.

Harith AE, Kolk AHJ, Kager PA *et al.* (1986) A simple and economical direct agglutination test for serodiagnosis and sero-epidemiological studies of visceral leishmanisis. *Transactions of*

© 2006 Blackwell Publishing Ltd

the Royal Society of Tropical Medicine and Hygiene **80**, 583–587.

Harith AE, Kolk AHJ, Leeuwenburg J *et al.* (1988) Improvement of a direct agglutination test for field studies of visceral leishmaniasis. *Journal of Clinical Microbiology* **26**, 1321–1325.

Landis JR & Koch JJ (1977) The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174.

Marlet MV, Wuillaume F, Jacquet D, Quispe KW, Dujardin JC & Boelaert M (2003) A neglected disease of humans: a new focus of visceral leishmaniasis in Bakool, Somalia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **97**, 667–671.

McGill R, Tukey JW & Larsen WA (1978) Variations of box plots. *The American Statistician* **32**, 12–16.

Meredith SEO, Kroon NCM, Sondorp E *et al.* (1995) Leish-KIT, a stable direct agglutination test based on freeze dried antigen for serodiagnosis of visceral leishmaniasis. *Journal of Clinical Microbiology* **33**, 1742–1745.

Oskam L, Nieuwenhuijs L & Hailu A (1999) Evaluation of the direct agglutination test (DAT) using freeze-dried antigen for the detection of anti-Leishmania antibodies in stored sera from various patient groups in Ethiopia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **93**, 275–277.

R Development Core Team (2005) *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.r-project.org.

Seaman J, Mercer AJ & Sondorp E (1996) The epidemic of visceral leishmaniasis in western Upper Nile, southern Sudan: course and impact from 1984 to 1994. *International Journal of Epidemiology* **25**, 862–871.

Zijlstra EE, Siddig MA, El-Hassan AM *et al.* (1992) Kala-azar: a comparative study of parasitological methods and the direct agglutination test in diagnosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **86**, 505–507.

Corresponding Author Diane Jacquet, Instituut voor Tropische Geneeskunde 'Prins Leopold', Eenheid Toegepaste Technologie & Productie, Nationalestraat 155, 2000 Antwerp, Belgium. Tel.: +32 3 24 76 351; Fax: +32 3 24 76 373; E-mail: djacquet@itg.be

Evaluation comparative d'antigènes lyophilisés et liquides pour le diagnostic sérologique de la leishmaniose viscérale par le test d'agglutination directe

OBJECTIF Le test d'agglutination directe basé sur l'antigène liquide (LQ DAT) pour la leishmaniose viscérale (LV) est connu pour être modérément reproductible à cause de la variabilité inter observateur et de lots à lots, ainsi qu'une sensibilité à la température et à l'agitation lors du transport. Nous avons mené une évaluation détaillée d'une version lyophilisée du test (FD DAT) et les concordances avec le test (LQ DAT).
MÉTHODE Des échantillons cliniques de sang suspects LV et de contrôles sains de milieu endémique ont été collectés au Soudan, au Népal et en Inde. Les tests FD DAT et LQ DAT ont été appliqués en duplicats dans les pays d'origine et dans le laboratoire de référence. Des tests sur la variabilité inter lots et sur la stabilité ont été menés et les concordances ont été analysées au sein des sites et entre eux, sur une échelle dichotomique de la valeur du Kappa de Cohen ainsi que sur une échelle continue par le diagramme de Bland–Altman.
RÉSULTATS L'antigène lyophilisé demeure encore totalement actif même une conservation à 45 °C pendant 24 mois. Une excellente concordance entre la version lyophilisée et la version liquide du test a été obtenue pour un titre seuil fixé à 1/6400.
CONCLUSION L'avantage principal de l'antigène lyophilisé est sa stabilité meilleure à des températures plus élevées et sa plus longue durée de vie, qui la rendent plus appropriée que la version liquide pour une utilisation sur le terrain.

mots clés Leishmaniose viscérale, test d'agglutination directe, antigène lyophilisé, antigène liquide

Evaluación comparativa de antígeno liofilizado y antígeno líquido en la Prueba de Aglutinación Directa para el serodiagnóstico de la Leishmaniasis Visceral

OBJETIVO Se sabe que la Prueba de Aglutinación Directa (PAD) para la Leishmaniasis Visceral (LV) con antígeno líquido (PAD ALQ) es solo moderadamente reproducible debido a la variación inter-observador y a la variablidad entre lotes, así como por lo sensible que es a la temperatura y a la agitación durante el transporte. Hemos conducido una evaluación exhaustiva de una versión de la PAD con antígeno liofilizado (PAD ALF) y su concordancia con PAD ALQ.
MÉTODO Se recogieron muestras de sangre de pacientes con sospecha clínica de LV y de controles sanos de áreas endémicas, en Sudán, Nepal e India. Tanto la PAD ALF como la PAD ALQ fueron realizadas por duplicado en los países respectivos y en un laboratorio de referencia. Se miraron la variabilidad entre lotes, las pruebas de estabilidad y la concordancia entre y dentro de los centros en una escala dicotómica utilizando la kappa de Cohen, así como en una escala continua mediante gráficos de Bland–Altman.
RESULTADOS El antígeno congelado estaba totalmente activo incluso después de guardarlo a 45 °C durante 24 meses. Utilizando como corte un título de 1 : 6400, la concordancia entre el antígeno congelado y el antígeno líquido fue excelente.
CONCLUSIÓN La principal ventaja del antígeno liofilizado es su mayor estabilidad a temperaturas más altas y una vida media más larga, que lo hace mucho más apropiado que la versión líquida para su uso en el campo.

palabras clave Leishmaniasis Visceral, Prueba de Aglutinación Directa, antígeno congelado antígeno liofilizado, antígeno líquido

© 2006 Blackwell Publishing Ltd