Transactions of the Royal Society of Tropical Medicine and Hygiene (2008) **102**, 32—40

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevierhealth.com/journals/trst

Diagnostic tests for kala-azar: a multi-centre study of the freeze-dried DAT, rK39 strip test and KAtex in East Africa and the Indian subcontinent

M. Boelaert[a,*], S. El-Safi[b], A. Hailu[c], M. Mukhtar[b], S. Rijal[d], S. Sundar[e], M. Wasunna[f], A. Aseffa[g], J. Mbui[f], J. Menten[a], P. Desjeux[h], R.W. Peeling[i]

[a] *Epidemiology and Disease Control Unit, Institute of Tropical Medicine, Nationalestraat 155, 2000 Antwerp, Belgium*
[b] *Department of Microbiology and Parasitology, Faculty of Medical Laboratory Sciences, Khartoum University, P.O. Box 102, Khartoum, Sudan*
[c] *Department of Microbiology, Immunology and Parasitology, Faculty of Medicine, Addis Ababa University, P.O. Box 9086, Addis Ababa, Ethiopia*
[d] *Department of Medicine, BP Koirala Institute of Health Sciences, Dharan, Nepal*
[e] *Institute of Medical Sciences, Banaras Hindu University, Varanasi, India*
[f] *Kenya Medical Research Institute, Centre for Clinical Research, P.O. Box 20778 00202, Nairobi, Kenya*
[g] *Armauer Hansen Research Institute, P.O. Box 1005, Addis Ababa, Ethiopia*
[h] *Institute for One World Health (IOWH), 899 rue Jean de Gingins, 01220 Divonne, France*
[i] *Diagnostics R&D, PDE, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR), World Health Organization, 20, Avenue Appia, 1211 Geneva 27, Switzerland*

Received 27 March 2007; received in revised form 4 September 2007; accepted 4 September 2007
Available online 22 October 2007

KEYWORDS
Visceral leishmaniasis;
Sensitivity and specificity;
Diagnostic tests;
Africa;
India;
Nepal

Summary Three diagnostic tests for visceral leishmaniasis (VL), the freeze-dried direct agglutination test (FD-DAT), the rK39 dipstick and a urine latex antigen test (KAtex), were evaluated for use in primary care in East Africa and the Indian subcontinent. Clinical suspects were prospectively recruited and tissue, blood and urine samples were taken. Direct microscopic examination of tissue smear, and FD-DAT, rK39 and KAtex were performed. Sensitivity and specificity with 95% credible intervals were estimated using Bayesian latent class analysis. On the Indian subcontinent both the FD-DAT and the rK39 strip test exceeded the 95% sensitivity and 90% specificity target, but not so in East Africa. Sensitivity of the FD-DAT was high in Ethiopia and Kenya but lower in Sudan, while its specificity was below 90% in Kenya. Sensitivity of the rK39 was below 80% in the three countries, and its specificity was only 70% in Ethiopia. KAtex showed moderate to very low sensitivity in all countries. FD-DAT and rK39 can be recommended

* Corresponding author. Tel.: +32 3 247 62 83; fax: +32 3 247 66 58.
E-mail address: mboelaert@itg.be (M. Boelaert).

doi:10.1016/j.trstmh.2007.09.003

for clinical practice on the Indian subcontinent. In East Africa, their clinical use should be carefully monitored. More work is needed to improve existing formats, and to develop better VL diagnostics.

1. Introduction

Visceral leishmaniasis (VL) is a deadly parasitic disease, causing an estimated 500 000 new cases a year, mainly on the Indian subcontinent and in East Africa. In both regions the main challenge for control is to increase access to treatment, as many VL patients go undetected and untreated (Gerstl et al., 2006; Singh et al., 2006). VL patient care should be decentralized from tertiary care and research centres to district hospitals and primary care settings, but this requires safe and effective drugs and simple and robust diagnostic tests.

The current reference standard for VL diagnosis is demonstration of amastigote stages of the *Leishmania* parasite in tissue smears and/or promastigote stages in culture from the same tissues. These parasitological techniques are highly specific, but their sensitivity varies depending on the type of tissue aspirate. Sensitivity of spleen aspirates approaches 95% (Bryceson, 1996; Thakur, 1997; Zijlstra et al., 1992); bone marrow or lymph node aspirates have a much lower sensitivity: 60—80% and 50—60%, respectively. The Ministry of Health in Sudan recommends lymph node aspirates, but Sudan is the only country where lymphadenopathy is a common presentation of VL (Zijlstra et al., 1992). Splenic aspiration is invasive and contraindicated in patients with severe anaemia and bleeding tendency and in those without palpable spleen or in restless children. After the procedure the patient has to be observed in a facility where blood transfusion and surgery are available. Even with the safer procedure that was developed in Kenya (Kager et al., 1983), a small risk of fatal haemorrhage exists (Chulay and Bryceson, 1983), and several authors have reported iatrogenic morbidity and mortality (Boussery et al., 2001; Sundar et al., 2002b). One death was observed in a series of 671 splenic aspirates in Kenya, and three in 3000 in India. The Kenyan Ministry of Health adopted splenic aspirate as the basis for VL diagnosis whenever it can be implemented under controlled conditions. The experience of non-governmental organizations (NGOs) operating in Southern Sudan shows that specialized VL programmes can safely practice splenic aspirates in a precarious environment (de Beer et al., 1991), but their level of staffing, resources and supervision is hard to reproduce in primary health care settings that often lack basic laboratory facilities.

In summary, parasite detection techniques require clinical and laboratory expertise, which is usually limited to referral hospitals in endemic areas, and better and more robust VL tests are needed. Therefore, one of the priorities of WHO/TDR over the past decade has been the development of a simple, highly sensitive, specific, reliable and affordable diagnostic test for VL that can be used in first-line health services in endemic countries. The target product profile aims at >95% sensitivity and >90% specificity in the clinical setting (WHO/TDR, unpublished).

Several leishmanial antibody-detection tests have been developed over the past decades, although few of them can be carried out without sophisticated laboratory equipment. These serological tests all share non-invasiveness and high sensitivity, but their main limitation is that they are not specific for active VL disease. They stay positive for months after cure and are also positive in healthy persons with asymptomatic leishmanial infection. Therefore, a combination of the WHO clinical case definition and a positive antibody test in patients with a first VL episode was taken as the basis for the decision to treat in several reported studies (Boelaert et al., 2004; Sundar et al., 2002b, 2006).

The first serological test to be proposed for field use was the direct agglutination test (DAT) (El Harith et al., 1986). The initial 1985 study by El Harith et al. (1986) reported high sensitivity and specificity of the DAT, values that were corroborated in other laboratory-based studies (el Safi and Evans, 1989; Hailu, 1990; Singla et al., 1993; Sinha and Sehgal, 1994). As the ongoing VL epidemic in Sudan (de Beer et al., 1991) created a pressing demand, the DAT test was rapidly taken to the field. A multi-centre study reported low reproducibility due to reading problems and heat and shock sensitivity of the liquid DAT antigen (Boelaert et al., 1999), and a freeze-dried version of the test was proposed to circumvent the latter problem (Meredith et al., 1995; Zijlstra et al., 1997). Over the past 20 years, the DAT was extensively used in kala-azar programmes run by NGOs in war-stricken Southern Sudan (Ritmeijer and Davidson, 2003).

More recently, high sensitivity and specificity was reported for an immunochromatographic test (ICT), the rK39 strip test, which is based on a 39-amino-acid-repeat recombinant leishmanial antigen from *Leishmania chagasi* (Sundar et al., 1998). The rK39 was first introduced in an ELISA (Badaro et al., 1996; Zijlstra et al., 1998) and, only later, in a dipstick format (Sundar et al., 1998). The latter format is very easy to use in the field and the initial study showed 100% sensitivity and 98% specificity (Sundar et al., 1998), but this particular format (Arista Biologicals, Allentown, PA, USA) is no longer commercially available. An evaluation in Sudan of an ICT from the same producer showed only 67% sensitivity (Zijlstra et al., 2001). An ICT produced by another company (INBIOS, Seattle, WA, USA) proved to be a good diagnostic guide in suspected cases of Indian kala-azar (Sundar et al., 2002a). Sarker et al. (2003) found excellent accuracy with the same ICT in Bangladesh. In Nepal, an early prototype showed a specificity of only 71% in controls with clinical signs (Chappuis et al., 2003); however, more encouraging results were obtained with later generations (Bern et al., 2000; Boelaert et al., 2004; Chappuis et al., 2006a), as well as with the IT-Leish format produced by another manufacturer (DiaMed-IT Leish®; DiaMed AG, Cressier sur Morat, Switzerland) (Ritmeijer et al., 2006).

Preliminary data are also available on a urinary antigen detection test (Attar et al., 2001). Sarkari et al.

(2002) described a low-molecular-weight, heat-stable carbohydrate in the urine of VL patients. An agglutination test format for the detection of this antigen, KAtex®, showed 100% specificity and a sensitivity between 64 and 100% in laboratory studies (Attar et al., 2001). However, the sensitivity of the KAtex was disappointingly low in clinically suspect patients in Nepal (Rijal et al., 2004). Further work seems warranted, as this technique holds promise for the monitoring of prognosis, which none of the serological tests has so far accomplished.

In 2001, WHO/TDR decided to evaluate the value of the three diagnostic tests that were considered candidates for use in peripheral health services — the freeze-dried version of the DAT antigen, the rK39 dipstick test and a urine latex antigen test — in the two most affected regions, East Africa and the Indian subcontinent.

2. Materials and methods

This diagnostic evaluation prospectively recruited all patients presenting with clinical symptoms of VL, the group on which the diagnostics should be applied in the future (Peeling et al., 2006; Zhou et al., 2002).

2.1. Study site

The study was conducted between 2003 and 2006. Prospective recruitment was done for all patients presenting at health facilities in the five countries with a compatible clinical presentation (see below). In Ethiopia, patients were recruited at the Kahsay Abera Hospital (KAH) in Humera district and at the Gondar College of Medical Sciences (GCMS) for patients in Gondar and the surrounding lowlands. In Kenya, patients were recruited at the Kabarnet hospital in Baringo district. In Sudan, patients were recruited at the Umalkhair and Tabarakelleh health centres in the Gedaref endemic area. In India, patients were recruited at the Kala-Azar Medical Research and Training Centre (KAMRC) in Muzaffarpur, Bihar state, and in Nepal, in the BP Koirala institute of Health Sciences, in Dharan. Each enrolled patient who gave informed consent received a unique identifier code number, and a clinical case report form with clinical and epidemiological data was filled out.

2.2. Sample size

In each country, the aim was to enrol at least 100 true VL cases (i.e. parasitologically positive) and at least 100 true negatives (i.e. patients with similar signs and symptoms, but without active VL disease), for achieving adequate precision (min. 10%) for the sensitivity and specificity estimates. Sample size computation was based on an expected prior probability of 33% true VL cases in the group of clinical VL suspects who present to the health services, with an addition of $n = 10$ subjects to compensate for any loss of information. The total sample size was thus fixed at 310 clinical suspects to enrol in each country. Patients were to be consecutively enrolled until the required sample size was achieved.

2.3. Inclusion criteria

All clinically suspect adult and paediatric patients who presented to the respective health facilities from the starting date of the study were enrolled until completion of sample size, on the condition of informed consent of the patient or his/her guardian. All sites considered every patient presenting with a history of fever of 2 weeks duration or more and with splenomegaly and/or lymphadenopathy as 'clinically suspect for VL', except for the Indian site, which also included patients with fever of lesser duration. Patients younger than 2 years old and pregnant women were excluded from the study. In the East African sites, patients with a thick film-positive malaria episode were excluded, in accordance with the WHO operational case definition (WHO, 1999).

2.4. Laboratory procedures

The following biological samples were collected from patients: tissue aspirates (lymph node, bone marrow or spleen) for direct microscopic examination and/or culture, a 7.5 ml venous blood sample for serological tests (5 ml in children) and a 5 ml urine sample. The direct microscopic examination of the tissue aspirate, the DAT (freeze-dried and liquid antigens), the rK39 dipstick and the urine test were performed at the peripheral study site under field conditions. Serum (1.5 ml) was collected in a 1.8 ml cryogenic vial (Corning®) and stored at −20 °C until transport to the central laboratory. The freeze-dried DAT and the rK39 dipstick were repeated twice and independently on each patient sample in the standardized conditions of the central laboratory, by a senior laboratory technician who was blinded to all the previously obtained data on the patients.

2.5. Parasitology

Direct semi-quantified microscopic examination of five methanol-fixed, five Jenner-stained and fifteen Giemsa-stained smears of lymph node or bone marrow aspirate were performed at the field laboratory (WHO, 1990). Whenever both aspirations were negative, spleen aspiration was performed if the safety of the procedure could be guaranteed (trained personnel, haemoglobin and platelet count within acceptable limits, blood for transfusion and surgical facility available). In the Ethiopian, Kenyan and Indian sites, splenic aspiration was performed on all VL suspects. Smears were graded on a scale from 0 to 6+ according to WHO criteria (WHO, 1990). In Ethiopia, NNN medium was inoculated with aspirate material for parasite isolation.

2.6. Freeze-dried DAT (FD-DAT)

DAT was performed on serum as described by Meredith et al. (1995) using a freeze-dried antigen supplied by the Koninklijk Instituut voor de Tropen (KIT), Amsterdam, The Netherlands (Batch number 0415). A sample was considered positive if it had a titre $\geq$1:3200. The same freeze-dried serum controls (supplied by KIT, Amsterdam) were used by all countries for quality assurance.

2.7. rK39 dipstick test

The Kalazar Detect Rapid Test™ (rK39 dipstick) was obtained from InBios International, Inc., Seattle, USA (Lot # DE1093). Procedures of the dipstick assay involve loading 20 μl serum samples obtained from freshly collected whole blood onto test strips, followed by addition of two to three drops of the buffer supplied in the kit. Test results were read after 10 min and recorded qualitatively as positive or negative reactions depending on the presence or absence of a line in the test area of the strip.

2.8. Latex antigen agglutination test in urine

KAtex® was supplied by Kalon Biological Ltd, Guildford, UK (Cat no. L3-040 and L3-041; Lot # K14-89). The urine latex agglutination tests were performed as described by the manufacturers.

2.9. Other tests

Haemoglobin, thick film and thin smear for malaria, as well as acid-fast stain of sputum or X-ray were done on all patients. Other diagnostic tests were done at the discretion of the clinician in charge of the patient. An HIV test subject to specific informed consent was offered by the Kenyan centre.

2.10. Data analysis

Estimates of sensitivity and specificity for each test in each centre were obtained by latent class analysis (LCA). LCA is a mathematical technique that models associations between observed variables that imperfectly measure a non-observable (latent) variable. The true disease status of a group of people is considered a latent variable with two mutually exclusive categories, 'diseased' and 'non-diseased'. If results of several diagnostic tests are observed for a group of individuals with unknown disease status, LCA will model the probability of each combination of test results conditional on their disease status (or latent class membership). The LCA model yields an estimate of disease prevalence and of sensitivity and specificity of all the diagnostic tests (Goetghebeur et al., 2000). In a basic latent class model, the observed variables have to be conditionally independent (conditional on the latent class), i.e. there should be no associations between the diagnostic tests within each category of the latent variable. More advanced models exist, in which this condition is relaxed. We used a Bayesian approach to model conditional dependence between multiple diagnostic tests (Black and Craig, 2002; Branscum et al., 2005).

A priori, a positive correlation of test outcomes in VL subjects was expected between DAT and rK39, related to the strength of immunoresponse to leishmanial infection, and between the urine KAtex test and parasitology results, related to the parasite load. Exploration of the data indicated that the local independence assumption was not supported. Different dependence models were assessed by a Bayesian *P*-value and the deviance information criterion (Berkvens et al., 2006). Including pairwise correlations between DAT and rK39 results and between urine KAtex test and parasitology results in VL subjects improved the fit significantly, and this model showed a satisfactory fit to the data. This model was used for the estimation of diagnostic test sensitivities and specificities and 95% credible intervals. If a correct statistical model is used, the 95% credible intervals contain the true sensitivity or specificity with 95% probability (Gelman et al., 2005). This corresponds to a classical 95% confidence interval. More complicated dependence structures were explored, but did not show a consistently better fit to the data across the five countries.

Agreement studies were carried out in several, but not all, centres, due to budget constraints. Repeatability was assessed by comparing the results of two repeats on the same sample for each test in Nepal as well as India. Reproducibility was assessed by comparing a test result obtained in field conditions with that of the same test carried out in a reference (central) laboratory in Ethiopia, India, Kenya and Sudan. Both tests were performed on the same serum or urine samples. Repeatability and reproducibility were summarized using Cohen's kappa coefficient (Cohen, 1968). Kappa coefficients were interpreted following Landis and Koch (1977): 1.00—0.81 excellent, 0.80—0.61 good, 0.60—0.41 moderate, 0.21—0.40 weak, 0.20—0.00 negligible agreement. Confidence intervals for Cohen's kappa were calculated using the bootstrap method (Efron and Tibshirani, 1993).

Statistical analyses were performed using R 2.2 (R Development Core Team and R Foundation for Statistical Computing, 2005) and WinBugs 1.4.1 (Spiegelhalter et al., 1996, 2004).

2.11. Ethical considerations

All patients with positive parasitology and/or positive DAT were treated in accordance with country policy and WHO guidelines. Data were handled on the basis of confidentiality.

3. Results

Patient recruitment was completed in Sudan between March and May 2003, in Ethiopia between January and March 2004, in Kenya between mid-August 2004 and September 2005, in India between May and August 2005 and in Nepal between May 2005 and March 2006.

The analysis was performed on all patients clinically suspect for VL that were recruited in the study following the protocol-defined inclusion criteria ($n = 1150$). Subjects identified in active screening (Ethiopia, $n = 63$) or healthy endemic controls (India, $n = 100$; Nepal, $n = 50$) were excluded from the analysis. Per protocol, subjects with malaria were excluded from the analysis in East Africa (Kenya, $n = 2$; Sudan, $n = 16$). The patient characteristics are described in Table 1. Over the five countries combined, 65% of subjects were male and 36% were children aged 12 or younger. The majority of subjects (58%) had fever for more than 1 month.

LCA was performed on the 1143 subjects that had data on all four diagnostic tests. Only seven subjects had data

Table 1 Characteristics of study patients with clinically suspected visceral leishmaniasis

	East Africa			Indian subcontinent	
	Ethiopia	Kenya	Sudan	India	Nepal
	(n = 38)	(n = 308)	(n = 294)	(n = 352)[a]	(n = 158)
Gender, % male	95	65	64	55	59
Age					
% ≤12 years	0	45	36	44	25
In years: median (range)	25 (16—49)	13 (1—75)	20 (2—72)	15 (2—65)	23 (2—68)
Duration of fever					
% >1 month	65	54	20	0.4	58
In weeks: mean (SD)	10.4 (9.2)	9.7 (16.1)	4.1 (2.1)	2.0[b] (0.7)	9.9 (9.7)

[a] For India, demographics computed for 260 of 352 enrolled patients only (data for one sub-centre lacking).
[b] In India, 35% of subjects presented with fever of less than 2 weeks duration.

missing for any of the four tests (Ethiopia, n = 3; Kenya, n = 1; Sudan, n = 3) and were excluded from the LCA analysis. The sensitivity and specificity estimates are summarized in Figure 1 and Table 2.

The repeatability of the three diagnostic tests was assessed on the Indian subcontinent. Repeatability of FD-DAT, rK39 dipstick and KAtex was excellent both in India and Nepal (data not shown).

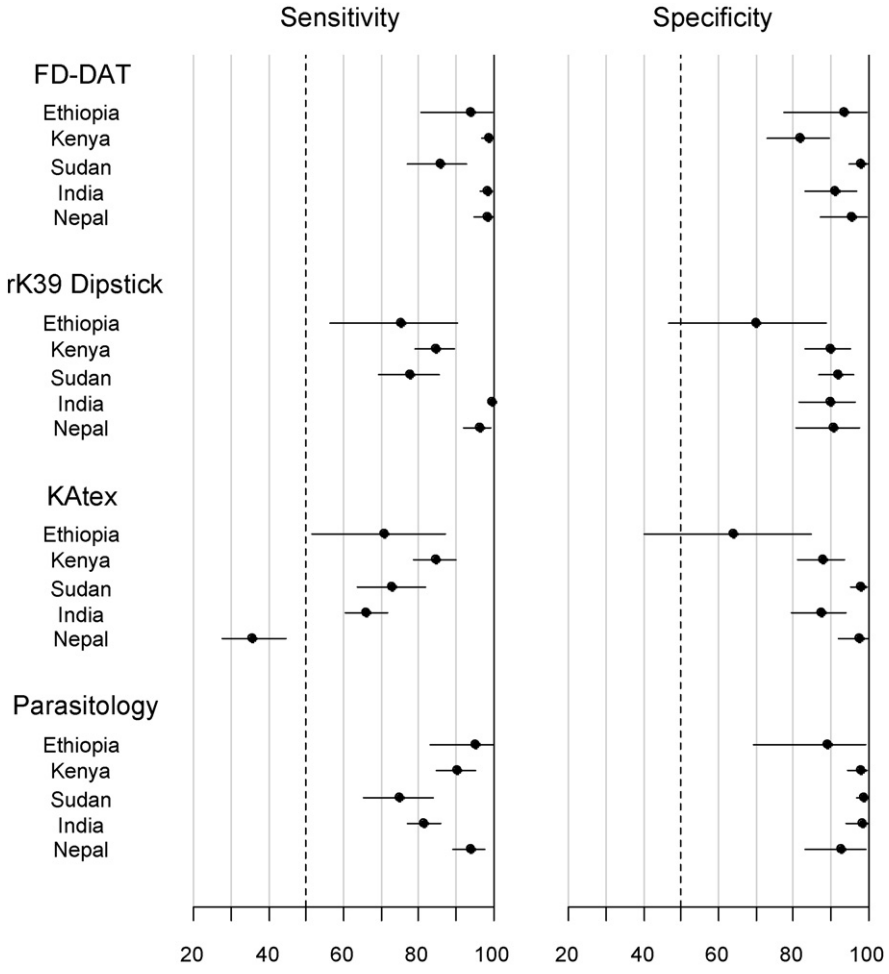

Figure 1 Sensitivity and specificity (estimate and 95% credible interval obtained by latent class analysis) of freeze-dried direct agglutination test (FD-DAT), rK39 dipstick, KAtex and parasitological examination used for the diagnosis of visceral leishmaniasis in field conditions on clinical suspect patients in five countries in East Africa and the Indian subcontinent.

Table 2 Prevalence in the test sample and sensitivity and specificity of freeze-dried direct agglutination test (FD-DAT), rK39 dipstick, KAtex and parasitological examination used for the diagnosis of visceral leishmaniasis (VL) in five countries in East Africa and the Indian subcontinent

Parameter	East Africa			Indian subcontinent	
	Ethiopia	Kenya	Sudan	India	Nepal
	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)
Prevalence	57.2 (40.5—73.4)	60.9 (54.7—66.7)	37.0 (31.0—43.2)	79.6 (75.1—83.7)	71.0 (63.5—77.9)
FD-DAT					
Sensitivity	94.0 (80.0—99.8)	98.8 (96.6—99.9)	85.7 (77.0—92.7)	98.3 (96.2—99.6)	98.5 (94.8—100)
Specificity	93.6 (77.4—99.8)	81.9 (73.2—89.8)	98.2 (94.8—99.9)	91.0 (83.0—96.6)	95.4 (87.1—99.6)
rK39 dipstick					
Sensitivity	75.4 (55.9—90.5)	84.7 (78.6—89.8)	77.9 (69.2—85.6)	99.6 (98.4—100)	96.5 (92.1—99.2)
Specificity	70.0 (46.3—88.9)	89.9 (83.2—95.1)	91.8 (86.7—96.2)	90.0 (81.2—96.4)	90.9 (80.8—97.5)
KAtex					
Sensitivity	71.0 (52.0—87.1)	84.5 (78.6—89.8)	72.9 (63.4—81.7)	66.1 (60.4—71.6)	35.8 (27.5—44.7)
Specificity	64.0 (40.0—84.6)	87.8 (80.8—93.5)	98.3 (95.3—99.8)	87.6 (79.1—94.1)	97.8 (92.1—100)
Parasitology					
Sensitivity	95.2 (82.5—99.9)	90.2 (84.5—95.1)	74.9 (65.0—83.8)	81.7 (76.8—86.0)	94.0 (88.9—97.7)
Specificity	89.2 (68.9—99.4)	98.0 (94.5—99.8)	98.8 (96.6—99.9)	98.3 (93.9—100)	92.7 (83.0—99.2)

95% CI: 95% credible interval.
Obtained from Bayesian latent class model with pairwise correlated outcomes for FD-DAT and rK39 and for KAtex and parasitology in VL subjects and uniform priors for all model parameters.

Table 3 Reproducibility of freeze-dried direct agglutination test (FD-DAT) and rK39 dipstick used for the diagnosis of visceral leishmaniasis in field conditions on clinical suspect patients in three countries in East Africa and India

	K (95% CI)
FD-DAT	
Ethiopia	0.74 (0.43—0.93)
Kenya	0.86 (0.79—0.91)
Sudan	0.78 (0.70—0.84)
India	0.94 (0.87—0.97)
Pooled	0.87 (0.83—0.90)
rK39 dipstick	
Ethiopia	0.88 (0.59—1.00)
Kenya	0.80 (0.73—0.86)
Sudan	0.85 (0.77—0.91)
India	0.95 (0.89—0.98)
Pooled	0.88 (0.85—0.91)

The reproducibility of the diagnostic tests between field and central laboratory settings was assessed in Ethiopia, Sudan, Kenya and India. Reproducibility was good to excellent for FD-DAT, and excellent for rK39 dipstick. Reproducibility assessment of KAtex was not included, as the test needs to be performed on fresh urine (Table 3).

4. Discussion

This study showed that two tests, FD-DAT and the rK39 strip test, reached the sensitivity (>95%) and specificity (>90%) target on the Indian subcontinent. In East Africa their performance was not up to this standard, except for the FD-DAT in Ethiopia. The sensitivity of the rK39 strip test in East African countries was in the moderate range of 75 to 85%, while its specificity was moderate to good. The FD-DAT in Kenya had excellent sensitivity but moderate specificity, while in Sudan the opposite was observed: moderate sensitivity but excellent specificity. KAtex failed the sensitivity target by a large amount in all five countries. The various parasitological algorithms showed a clear deficit in sensitivity in Sudan as well as in India, but matched the specificity criteria of >90%.

This evaluation focused on the estimation of sensitivity and specificity of serological assays when used in combination with a clinical case definition in the clinical setting. No conclusions can be drawn about the performance of these tests when used for screening at population level. For this reason, a number of cases in Ethiopia that were recruited through active screening were excluded from the analysis.

Our study was done with particular brands and products, and some variance in reported estimates could be related to occasional poor batch performance of DAT (Jacquet et al., 2006), or brand-related features of rK39. Our results are consistent with those of a recent meta-analysis that also showed the excellent performance of rK39 and DAT in the clinical setting on the Indian subcontinent, but the lower sensitivity of rK39 strip test in the East African setting (Chappuis et al., 2006b). The latter could be due to lower levels of specific anti-rK39 antibodies in East African kala-azar patients that might be age- or stage-related. However, a study conducted in Uganda by Chappuis et al. (2005) with another brand of rK39 rapid test actually demonstrated a significantly higher sensitivity (97%) than that of the dip-

stick (82%) used in the present study, and should therefore be evaluated in other East African endemic foci.

Another factor of variance in diagnostic test performance could be related to the technical complexity and ease of reading of the test. Simple test formats such as rK39 will yield more reproducible results than complex titration procedures such as the FD-DAT. Reproducibility of tests was good to excellent for rK39 as well as for FD-DAT. The reproducibility assessment done in our study may yield a too optimistic view of the reality, as the repeat executions were mostly performed by technicians belonging to the same research team, or at least working in facilities where a high turnover of tests would maintain a sufficient level of skills. This cannot be readily transposed to the performance of laboratory technicians at the first level of the health service. Therefore, the simpler format of the rK39 strip test suggests clear advantages for use in peripheral health services compared with the current FD-DAT format.

This study included the antigen detection assay KAtex, although the test was clearly not at the same level of development as the rK39 or DAT, and probably not yet ready for field evaluation. Although the technical principle of the KAtex test is promising, its current format, requiring the boiling of urine, does not allow recommendation for clinical practice. Further development should therefore be encouraged, given other potential advantages of this test. KAtex showed good performance in HIV-positive patients, and showed potential as a prognostic marker (Riera et al., 2004).

Although this study was set up as a multi-centre study with a common protocol, we had to respect existing clinical practice guidelines regarding tissue aspiration and could therefore not adopt splenic aspirate as reference standard in all countries.

Also, the parasitological reference is known to be suboptimal in its sensitivity, and can lead to substantial misclassification. Therefore we used a statistical method, LCA, for case classification, and this yielded estimates of sensitivity and specificity of all tests that can be easily compared across countries. We refer to the country-specific papers that include sensitivity and specificity estimates of serology obtained in the comparison with parasitology; these were broadly consistent with the above estimates (Sundar et al., 2007).

This evaluation showed that none of the three candidate field tests, nor parasitology, meets the target product profile in all countries. Moreover, the antibody detection tests have known limitations that were not addressed in this evaluation: they are not specific for acute disease and are positive in asymptomatic carriers as well as in past but cured VL cases. This means that they are useless for diagnosing relapse, in which case parasitology should still bring the answer. They are equally useless in establishing a proof of cure in a treated patient. However, this does not preclude the use of rK39 or FD-DAT in clinical practice in first level health centres, if they are embedded in a rational approach to diagnosis. Each country or region will have to adopt diagnostic algorithms on the basis of known test characteristics, and the solution might lie in a combination of tests. Researchers in the VL elimination programme on the Indian subcontinent decided recently to rely on serological tests in case management algorithms. This study also showed that the quest for better VL diagnostics has to continue. More work needs to be done on the improvement of existing test formats, and there is still ample scope for radical innovation.

Authors' contributions: The protocol for this study was developed by MB, SES, AH, MM, MW and PD. This protocol was later adapted for use on the Indian subcontinent by SS and SR. SES, AH, MM, SR, SS and MW were the principal investigators at the respective study sites; AA and JMb had substantial input in data collection, monitoring and interpretation of the data; JMe did the statistical analysis for the study; all authors participated in the analysis, the interpretation of results, and the drafting and reviewing of the manuscript, and read and approved the final version. MB is guarantor of the paper.

Acknowledgements: Prof. A.E. Harith is gratefully acknowledged for his support to this study, as well as the several co-investigators and the laboratory technicians in each study centre.

Funding: UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR). Authors had full control over the conduct of the research, the primary data and the decision to publish.

Conflicts of interest: None declared.

Ethical approval: The protocol for this study was approved by the Institutional Review Boards of the Institute of Tropical Medicine, Antwerp, Belgium; Khartoum University, Khartoum, Sudan; Addis Ababa University, Addis Ababa, Ethiopia; BP Koirala Institute of Health Sciences, Dharan, Nepal; Banaras Hindu University, Varanasi, India; Kenya Medical Research Institute, Nairobi, Kenya; Armauer Hansen Research Institute, Addis Ababa, Ethiopia; Institute for One World Health, Divonne, France; and by the WHO Research Ethics Review Committee, Geneva, Switzerland.

References

Attar, Z.J., Chance, M.L., El Safi, S., Carney, J., Azazy, A., El Hadi, M., Dourado, C., Hommel, M., 2001. Latex agglutination test for the detection of urinary antigens in visceral leishmaniasis. Acta Trop. 78, 11—16.

Badaro, R., Benson, D., Eulalio, M.C., Freire, M., Cunha, S., Netto, E.M., Pedral-Sampaio, D., Madureira, C., Burns, J.M., Houghton, R.L., David, J.R., Reed, S.G., 1996. rK39: a cloned antigen of *Leishmania chagasi* that predicts active visceral leishmaniasis. J. Infect. Dis. 173, 758—761.

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

Bern, C., Jha, S.N., Joshi, A.B., Thakur, G.D., Bista, M.B., 2000. Use of the recombinant K39 dipstick test and the direct agglutination test in a setting endemic for visceral leismaniasis in Nepal. Am. J. Trop. Med. Hyg. 63, 153—157.

Black, M.A., Craig, B.A., 2002. Estimating disease prevalence in the absence of a gold standard. Stat. Med. 21, 2653—2669.

Boelaert, M., el Safi, S.H., Mousa, H., Githure, J., Mbati, P.A., Gurubacharya, V.L., Shrestha, J., Jacquet, D., De Muynck, A., Le Ray, D., Van der Stuyft, P., 1999. Multi-centre evaluation of

repeatability and reproducibility of the direct agglutination test for visceral leishmaniasis. Trop. Med. Int. Health 4, 31—37.

Boelaert, M., Rijal, S., Regmi, S., Singh, R., Karki, B., Jacquet, D., Chappuis, F., Campino, L., Desjeux, P., Le Ray, D., Koirala, S., Van der Stuyft, P., 2004. A comparative study of the effectiveness of diagnostic tests for visceral leishmaniasis. Am. J. Trop. Med. Hyg. 70, 72—77.

Boussery, G., Boelaert, M., van Peteghem, J., Ejikon, P., Henckaerts, K., 2001. Visceral leishmaniasis (kala-azar) outbreak in Somali refugees and Kenyan shepherds, Kenya. Emerg. Infect. Dis. 7, 603—604.

Branscum, A.J., Gardner, I.A., Johnson, W.O., 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. Prev. Vet. Med. 68, 145—163.

Bryceson, A., 1996. Leishmaniasis, in: Cook, G.C. (Ed), Manson's Tropical Diseases. WB Saunders, London, pp. 1213—1243.

Chappuis, F., Rijal, S., Singh, R., Acharya, P., Karki, B.M., Das, M.L., Bovier, P.A., Desjeux, P., Le Ray, D., Koirala, S., Loutan, L., 2003. Prospective evaluation and comparison of the direct agglutination test and an rK39-antigen-based dipstick test for the diagnosis of suspected kala-azar in Nepal. Trop. Med. Int. Health 8, 277—285.

Chappuis, F., Mueller, Y., Nguimfack, A., Rwakimari, J.B., Couffignal, S., Boelaert, M., Cavailler, P., Loutan, L., Piola, P., 2005. Diagnostic accuracy of two rK39 antigen-based dipsticks and the formol gel test for rapid diagnosis of visceral leishmaniasis in northeastern Uganda. J. Clin. Microbiol. 43, 5973—5977.

Chappuis, F., Rijal, S., Jha, U.K., Desjeux, P., Karki, B.M., Koirala, S., Loutan, L., Boelaert, M., 2006a. Field validity, reproducibility and feasibility of diagnostic tests for visceral leishmaniasis in rural Nepal. Trop. Med. Int. Health 11, 31—40.

Chappuis, F., Rijal, S., Soto, A., Menten, J., Boelaert, M., 2006b. A meta-analysis of the diagnostic performance of the direct agglutination testand rK39 dipstick for visceral leishmaniasis. BMJ 333, 723.

Chulay, J.D., Bryceson, A.D., 1983. Quantitation of amastigotes of *Leishmania donovani* in smears of splenic aspirates from patients with visceral leishmaniasis. Am. J. Trop. Med. Hyg. 32, 475—479.

Cohen, J., 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol. Bull. 70, 213—220.

de Beer, P., El Harith, A., Deng, L.L., Semiao Santos, S.J., Chantal, B., van Grootheest, M., 1991. A killing disease epidemic among displaced Sudanese population identified as visceral leishmaniasis. Am. J. Trop. Med. Hyg. 44, 283—289.

Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap. Chapman and Hall, Florida, USA.

El Harith, A., Kolk, A.H., Kager, P.A., Leeuwenburg, J., Muigai, R., Kiugu, S., Laarman, J.J., 1986. A simple and economical direct agglutination test for serodiagnosis and sero-epidemiological studies of visceral leishmaniasis. Trans. R. Soc. Trop. Med. Hyg. 80, 583—587.

el Safi, S.H., Evans, D.A., 1989. A comparison of the direct agglutination test and enzyme-linked immunosorbent assay in the sero-diagnosis of leishmaniasis in the Sudan. Trans. R. Soc. Trop. Med. Hyg. 83, 334—337.

Gelman, A., Van Mechelen, I., Verbeke, G., Heitjan, D.F., Meulders, M., 2005. Multiple imputation for model checking: completed-data plots with missing and latent data. Biometrics 61, 74—85.

Gerstl, S., Amsalu, R., Ritmeijer, K., 2006. Accessibility of diagnostic and treatment centres for visceral leishmaniasis in Gedaref State, northern Sudan. Trop. Med. Int. Health 11, 167—175.

Goetghebeur, E., Liinev, J., Boelaert, M., Van der Stuyft, P., 2000. Diagnostic test analyses in search of their gold standard: latent class analyses with random effects. Stat. Methods Med. Res. 9, 231—248.

Hailu, A., 1990. Pre- and post-treatment antibody levels in visceral leishmaniasis. Trans. R. Soc. Trop. Med. Hyg. 84, 673—675.

Jacquet, D., Boelaert, M., Seaman, J., Rijal, S., Sundar, S., Menten, J., Magnus, E., 2006. Comparative evaluation of freeze-dried and liquid antigens in the direct agglutination test for serodiagnosis of visceral leishmaniasis (ITMA-DAT/VL). Trop. Med. Int. Health 11, 1777—1784.

Kager, P.A., Rees, P.H., Manguyu, F.M., Bhatt, K.M., Bhatt, S.M., 1983. Splenic aspiration: experience in Kenya. Trop. Geogr. Med. 35, 125—131.

Landis, J.R., Koch, J.J., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159—174.

Meredith, S.E., Kroon, N.C., Sondorp, E., Seaman, J., Goris, M.G., van Ingen, C.W., Oosting, H., Schoone, G.J., Terpstra, W.J., Oskam, L., 1995. Leish-KIT, a stable direct agglutination test based on freeze-dried antigen for serodiagnosis of visceral leishmaniasis. J. Clin. Microbiol. 33, 1742—1745.

Peeling, R.W., Smith, P.G., Bossuyt, P.M., 2006. A guide for diagnostic evaluations. Nat. Rev. Microbiol. 4 (12 Suppl.), S2—S6.

R Development Core Team, R Foundation for Statistical Computing, V.A., 2005. A Language and Environment for Statistical Computing. Elsevier Science, Amsterdam, The Netherlands.

Riera, C., Fisa, R., Lopez, P., Ribera, E., Carrio, J., Falco, V., Molina, I., Gallego, M., Portus, M., 2004. Evaluation of a latex agglutination test (KAtex) for detection of *Leishmania* antigen in urine of patients with HIV-*Leishmania* coinfection: value in diagnosis and post-treatment follow-up. Eur. J. Clin. Microbiol. Infect. Dis. 23, 899—904.

Rijal, S., Boelaert, M., Regmi, S., Karki, B.M., Jacquet, D., Singh, R., Chance, M.L., Chappuis, F., Hommel, M., Desjeux, P., Van der Stuyft, P., Le Ray, D., Koirala, S., 2004. Evaluation of a urinary antigen-based latex agglutination test in the diagnosis of kala-azar in eastern Nepal. Trop. Med. Int. Health 9, 724—729.

Ritmeijer, K., Davidson, R.N., 2003. Royal Society of Tropical Medicine and Hygiene joint meeting with Médecins Sans Frontières at Manson House, London, 20 March 2003: field research in humanitarian medical programmes. Médecins Sans Frontières interventions against kala-azar in the Sudan, 1989—2003. Trans. R. Soc. Trop. Med. Hyg. 97, 609—613.

Ritmeijer, K., Melaku, Y., Mueller, M., Kipngetich, S., O'keeffe, C., Davidson, R.N., 2006. Evaluation of a new recombinant K39 rapid diagnostic test for Sudanese visceral leishmaniasis. Am. J. Trop. Med. Hyg. 74, 76—80.

Sarkari, B., Chance, M., Hommel, M., 2002. Antigenuria in visceral leishmaniasis: detection and partial characterisation of a carbohydrate antigen. Acta Trop. 82, 339—348.

Sarker, C.B., Momen, A., Jamal, M.F., Siddiqui, N.I., Siddiqui, F.M., Chowdhury, K.S., Rahman, S., Talukder, S.I., 2003. Immunochromatographic (rK39) strip test in the diagnosis of visceral leishmaniasis in Bangladesh. Mymensingh. Med. J. 12, 93—97.

Singh, S., Reddy, D., Mishra, R., Sundar, S., 2006. Knowledge, attitude and practices related to Kala-Azar in a rural area of Bihar state, India. Am. J. Trop. Med. Hyg. 75, 505—508.

Singla, N., Singh, G.S., Sundar, S., Vinayak, V.K., 1993. Evaluation of the direct agglutination test as an immunodiagnostic tool for kala-azar in India. Trans. R. Soc. Trop. Med. Hyg. 87, 276—278.

Sinha, R., Sehgal, S., 1994. Comparative evaluation of serological tests in Indian kala-azar. J. Trop. Med. Hyg. 97, 333—340.

Spiegelhalter, D., Thomas, A., Best, N., Gilks, W., 1996. BUGS 0. 5 — Bayesian Inference Using Gibbs Sampling Manual (version ii). MRC Biostatistics Unit, Cambridge, UK.

Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2004. WinBUGS with DoodleBUGS, Version 1.4.1. Imperial College and Medical Research Council, London.

Sundar, S., Reed, S.G., Singh, V.P., Kumar, P.C., Murray, H.W., 1998. Rapid accurate field diagnosis of Indian visceral leishmaniasis. Lancet 351, 563—565.

Sundar, S., Pai, K., Sahu, M., Kumar, V., Murray, H.W., 2002a. Immunochromatographic strip-test detection of anti-K39 anti-

body in Indian visceral leishmaniasis. Ann. Trop. Med. Parasitol. 96, 19—23.

Sundar, S., Sahu, M., Mehta, H., Gupta, A., Kohli, U., Rai, M., Berman, J.D., Murray, H.W., 2002b. Noninvasive management of Indian visceral leishmaniasis: clinical application of diagnosis by K39 antigen strip testing at a kala-azar referral unit. Clin. Infect. Dis. 35, 581—586.

Sundar, S., Maurya, R., Singh, R.K., Bharti, K., Chakravarty, J., Parekh, A., Rai, M., Kumar, K., Murray, H.W., 2006. Rapid, non-invasive diagnosis of visceral leishmaniasis in India: comparison of two immunochromatographic strip tests for detection of anti-K39 antibody. J. Clin. Microbiol. 44, 251—253.

Sundar, S., Singh, R.K., Bimal, S.K., Gidwani, K., Mishra, A., Maurya, R., Singh, S.K., Manandhar, K.D., Boelaert, M., Rai, M., 2007. Comparative evaluation of parasitology and serological tests in the diagnosis of visceral leishmaniasis in India: a phase III diagnostic accuracy study. Trop. Med. Int. Health 12, 284—289.

Thakur, C.P., 1997. A comparison of intercostal and abdominal routes of splenic aspiration and bone marrow aspiration in the diagnosis of visceral leishmaniasis. Trans. R. Soc. Trop. Med. Hyg. 91, 668—670.

WHO, 1990. Control of the Leishmaniasis. Report of a WHO Expert Committee. World Health Organization, Geneva, Technical Report Series No. 793.

WHO, 1999. WHO Recommended Surveillance Standards, second ed. World Health Organization, Geneva, WHO/CDS/CSR/ISR/99.2.

Zhou, X.H., McClish, D.K., Obuchowski, N.A., 2002. Statistical Methods in Diagnostic Medicine. John Wiley, New York.

Zijlstra, E.E., Ali, M.S., el Hassan, A.M., el Toum, I.A., Satti, M., Ghalib, H.W., Kager, P.A., 1992. Kala-azar: a comparative study of parasitological methods and the direct agglutination test in diagnosis. Trans. R. Soc. Trop. Med. Hyg. 86, 505—507.

Zijlstra, E.E., Osman, O.F., Hofland, H.W., Oskam, L., Ghalib, H.W., el Hassan, A.M., Kager, P.A., Meredith, S.E., 1997. The direct agglutination test for diagnosis of visceral leishmaniasis under field conditions in Sudan: comparison of aqueous and freeze-dried antigens. Trans. R. Soc. Trop. Med. Hyg. 91, 671—673.

Zijlstra, E.E., Daifalla, N.S., Kager, P.A., Khalil, E.A., el Hassan, A.M., Reed, S.G., Ghalib, H.W., 1998. rK39 enzyme-linked immunosorbent assay for diagnosis of *Leishmania donovani* infection. Clin. Diagn. Lab. Immunol. 5, 717—720.

Zijlstra, E.E., Nur, Y., Desjeux, P., Khalil, E.A., el Hassan, A.M., Groen, J., 2001. Diagnosing visceral leishmaniasis with the recombinant K39 strip test: experience from the Sudan. Trop. Med. Int. Health 6, 108—113.