LSHTM Research Online

https://researchonline.lshtm.ac.uk

A Global Comparative Evaluation of Commercial Immunochromatographic Rapid Diagnostic Tests for Visceral Leishmaniasis

Jane Cunningham,[1,a] Epco Hasker,[2,a] Pradeep Das,[3] Sayda El Safi,[4] Hiro Goto,[5] Dinesh Mondal,[6] Margaret Mbuchi,[7] Maowia Mukhtar,[8] Ana Rabello,[9] Suman Rijal,[10] Shyam Sundar,[11] Monique Wasunna,[7] Emily Adams,[12] Joris Menten,[2] Rosanna Peeling,[13] and Marleen Boelaert[2] for the WHO/TDR Visceral Leishmaniasis Laboratory Network[b]

[1]UNICEF/UNDP/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases, Geneva, Switzerland; [2]Institute of Tropical Medicine, Antwerp, Belgium; [3]Rajendra Memorial Research Institute of Medical Science, Patna, India; [4]Department of Medical Microbiology and Parasitology, University of Khartoum, Sudan; [5]Laboratório de Soroepidemiologia e Imunobiologia, Instituto de Medicina Tropical de São Paulo, Brazil; [6]International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh; [7]Kenya Medical Research Institute, Nairobi; [8]Institute Endemic Diseases, University of Khartoum, Sudan; [9]Centro de Pesquisas René Rachou-FIOCRUZ, Belo Horizonte, Brazil; [10]B. P. Koirala Institute of Health Sciences, Dharan, Nepal; [11]Banaras Hindu University, Varanasi, India; [12]Royal Tropical Institute/Koninklijk Instituut voor de Tropen, Amsterdam, Netherlands and [13]London School of Hygiene and Tropical Medicine, London, United Kingdom

Background. Poor access to diagnosis stymies control of visceral leishmaniasis (VL). Antibody-detecting rapid diagnostic tests (RDTs) can be performed in peripheral health settings. However, there are many brands available and published reports of variable accuracy.

Methods. Commercial VL RDTs containing bound rK39 or rKE16 antigen were evaluated using archived human sera from confirmed VL cases (n = 750) and endemic non-VL controls (n = 754) in the Indian subcontinent (ISC), Brazil, and East Africa to assess sensitivity and specificity with 95% confidence intervals. A subset of RDTs were also evaluated after 60 days' heat incubation (37°C, 45°C). Interlot and interobserver variability was assessed.

Results. All test brands performed well against ISC panels (sensitivity range, 92.8%–100%; specificity range, 96%–100%); however, sensitivity was lower against Brazil and East African panels (61.5%–91% and 36.8%–87.2%, respectively). Specificity was consistently > 95% in Brazil and ranged between 90.8% and 98% in East Africa. Performance of some products was adversely affected by high temperatures. Agreement between lots and readers was good to excellent ($\kappa > 0.73$–0.99).

Conclusions. Diagnostic accuracy of VL RDTs varies between the major endemic regions. Many tests performed well and showed good heat stability in the ISC; however, reduced sensitivity against Brazilian and East African panels suggests that in these regions, used alone, several RDTs are inadequate for excluding a VL diagnosis. More research is needed to assess ease of use and to compare performance using whole blood instead of serum and in patients coinfected with human immunodeficiency virus.

Visceral leishmaniasis (VL) is a parasitic disease transmitted through the bite of an infected phlebotomine sandfly [1]. The clinical syndrome is characterized by fever, weight loss, splenomegaly, and pancytopenia and is nearly always fatal if left untreated. Though visceral leishmaniasis is endemic in >60 countries, 90% of the 200 000–400 000 annual cases occur in just 6 countries: Bangladesh, Brazil, Ethiopia, India, Nepal, and Sudan [2].

Parasitological confirmation remains the reference standard for diagnosis but is not very sensitive unless

a spleen puncture is performed. The invasiveness and potentially fatal complications associated with splenic aspiration has motivated the development of noninvasive serological tests such as direct agglutination test (DAT) [3] and lateral flow immunochromatographic tests (ICT), commonly referred to as rapid diagnostic tests (RDTs). To be useful, VL RDTs must have adequate (1) sensitivity to detect a high proportion of clinical cases, (2) specificity to accurately discriminate VL from other relevant disease conditions, (3) thermal stability for accuracy to be maintained after transport and storage in ambient conditions, and (4) ease of use to allow the correct interpretation of results. A meta-analysis [4] and a multicenter evaluation [5] corroborated earlier findings of high diagnostic accuracy of the rK39 ICT and led to its adoption as a diagnostic test in the Indian subcontinent VL Elimination Initiative. The enthusiasm and rapid uptake of RDTs for VL in the Asian region has prompted a surge of commercial tests targeting serum antibodies to rK39 and other antigens (eg, rKE16) [6]. However, in other endemic regions such as East Africa, reports of lower test sensitivity [7–9] have left the role of RDTs less clear. Moreover, there are few, if any, reports of diagnostic accuracy in the peer-reviewed literature for tests other than the Kalazar Detect (Inbios International) and DiaMed-IT LEISH (Bio-Rad Laboratories) and equally few head-to-head comparisons. Essential characteristics as heat stability are rarely assessed. As independent data on how well these assays meet criteria are lacking in countries without regulation by national testing authorities, the UNICEF/World Bank/United Nations Development Programme/World Health Organization (WHO) Special Programme for Research and Training in Tropical Diseases (TDR) coordinated a multiregional head-to-head laboratory-based evaluation of 4 commercially available RDTs in 3 global regions of VL endemicity using well-characterized panels of human sera; a fifth RDT was included in the Indian subcontinent.

METHODS

Site Selection

In 2008, TDR issued a public request for applications for any laboratory interested in becoming a member of a VL laboratory network to perform the VL RDT evaluation. Before implementing the evaluation, each laboratory had DAT refresher training and passed a DAT proficiency assessment. All technicians underwent training in good clinical laboratory practices and sites were independently monitored.

Assay Selection

An extensive search of VL RDTs available was performed using both generic (Google) and scientific (PubMed) Internet-based search engines. In March 2009, companies selling VL RDTs were contacted by email and the evaluation was publically announced through an open call for expression of interest for tests that fit the following inclusion criteria: (1) rapid—test result available in 15 minutes; (2) simple—test can be performed with minimal equipment and training; and (3) easy to interpret—cassette or strip format with visual read-out. Manufacturers who responded were required to (1) provide a certificate of quality manufacturing, either ISO 13485:2003 or US Food and Drug Administration Title 21 CFR certification; (2) supply sufficient quantities of products (lot 1: 1849; lot 2: 1794 [exception for OnSite Leishmania Ab Rapid Test Strip, CTK Biotech, which provided lot 1: 623 and lot 2: 598]); and (3) a signed confidentiality agreement with the WHO permitting the publication of results in the public domain.

Evaluation Panel

Each laboratory assembled a performance panel using locally archived and depersonalized sera. Prior to serum banking, VL disease was confirmed parasitologically (from spleen, bone marrow, or lymph node) by microscopy and/or culture. Archived sera were characterized by DAT (KIT, lot 0904) (titer between 1:100 and 1:102 400 antigen dilution) as low, medium, or high antibody titer. All healthy endemic control samples and potentially cross-reactive samples including malaria, Chagas disease, tuberculosis, and cutaneous leishmaniasis were DAT negative (<1:3200 antigen dilution) (Table 1).

Samples were thawed once and aliquoted in volumes required to evaluate all RDTs included in the evaluation and were then stored at −70°C until testing.

Sample Blinding

Daily, samples (in sets of 10–20) were randomized and relabeled with study codes by a study team member not involved in the assessment of the RDTs at any stage. The study code key was kept in a secure location and provided at the completion of testing.

Sample Size

We aimed for a sensitivity and specificity estimate pooled at regional level. Assuming a true sensitivity of 95%, a sample size of 250 VL cases in each endemic region is required to assure with a power of 80% and a confidence level of 95% that the lower margin of the confidence interval (CI) is at least 90%. The same applies to specificity; in this case the sample consists of nondiseased controls. We therefore selected, per region, 250 cases and 250 controls; included among the latter were samples of 210 healthy endemic controls and of 40 patients with potentially cross-reactive, endemic disease conditions (Table 1, Supplementary Table 1).

Table 1. Performance Evaluation Panel Characteristics

| | Sample Size | | | DAT Titer, Cases (%) | | | Average Age of Patients (y) | | HIV Status[a] (%) | | |
| | | | | | | | | | | Known | |
	Cases	HEC	Non-VL Disease Condition Controls	Low (≤100)	Medium (1:200–1:6400)	High (>1:6400)	Case	Controls	Unknown	Positive	Negative
East Africa	250	210	40	0.4	8.8	90.8	14	20	47.2	0	52.8
Brazil	237–250	206–209	42–45	0.0	11.6	88.4	19	27	98.6	0	1.4
ISC	250	210	39	0.0	3.2	96.8	24	31	50.8	0.5	48.7

Abbreviations: DAT, direct agglutination test; HEC, healthy endemic control; HIV, human immunodeficiency virus; ISC, Indian subcontinent; VL, visceral leishmaniasis.

[a] Includes cases and controls combined.

RDT Testing Methodologies

Daily sample sets were removed from the −70°C freezer and brought to room temperature. Each sample was tested once against each product according to manufacturers' instructions. RDT envelopes were opened immediately before use. The specified volume of serum was dispensed onto the RDT by micropipette. The buffer was applied using the dropper provided. Results were read and recorded on a standardized form by a first technician at the minimum reading time and within 30 minutes the test was read and recorded by a second technician, blinded to the reading of the first. Results of test and control lines were recorded as positive or negative by each technician. More specifically, band intensity was recorded for each test, based on standardized charts. If the control line was recorded as absent by either technician, the same sample was tested once against a new RDT. If the control line was still absent, the test result was recorded as invalid.

In each of the 9 evaluation centers, for each test kit, the same lot was used. All RDTs from the first lot were tested before 25% of panel samples, including all 3 infection status categories, were retested against a second lot of the same test kit, to assess lot-to-lot variability.

Thermal Stability

In 1 laboratory per region, each RDT, excluding 1 product (Signal KA) not recommended for room temperature storage, was tested against a panel of 24 serum samples (20 VL cases and 4 negative controls) at day 0 as a baseline measure; then stored in original packaging in calibrated incubators of 4°C, 37°C, and 45°C and tested (against the same panel) on day 60.

Ethics

Each laboratory participating in the evaluation obtained approval from the local ethics board and the WHO Research Ethics Committee.

Data Management and Analysis

Data were entered into an Epi Info database using a double data entry procedure. Data files were compared to identify typing errors. For data analysis we used Stata/IC version 10.1 (StataCorp, College Station, Texas). Diagnostic accuracy was calculated using RDT results from the first reading at the minimum reading time. We calculated proportions with 95% CIs [10, 11]. As a measure for reproducibility, we used Cohen κ coefficient [11]. These were interpreted following Landis and Koch: 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 [12]. Thermal stability results were reported as proportions of positive results for case samples and negative results for control samples, as were proportions of invalid results.

RESULTS

Site Selection

Twenty-two applications were received and 9 laboratories in the Indian subcontinent (n = 4), South America (n = 2), and Eastern Africa (n = 3) were chosen based on several criteria including access to patients, geographical location, laboratory facilities, expertise, and experience with RDTs. Laboratories were Rajendra Memorial Research Institute of Medical Sciences, India; Institute of Medical Sciences, Banaras Hindu University, India; Parasitology Laboratory, International Centre for Diarrhoeal Disease Research, Bangladesh; B. P. Koirala Institute of Health Sciences, Nepal; Kenya Medical Research Institute, Kenya; Faculty of Medicine, University of Khartoum, Sudan; Institute Endemic Diseases, University of Khartoum, Sudan; Laboratório de Soroepidemiologia e Imunobiologia Instituto de Medicina Tropical de São Paulo, Brazil; and Centro de Pesquisas René Rachou, Fundação Oswaldo Cruz, Fiocruz, Brazil. The Institute of Tropical Medicine, Belgium (ITM), was contracted as an independent partner to coordinate logistics of RDT supplies, as well as proficiency testing in the DAT

for VL laboratories alongside the Royal Tropical Institute, Netherlands.

Assay Selection
The survey identified 3 suppliers, and a fourth (CTK Biotech) responded independently to the public expression of interest after the deadline. In total, 5 commercial RDTs that met the inclusion criteria were included in the evaluation (Table 2).

RDT Shipments
All manufacturers shipped tests from 2 lots to ITM. Here, the RDTs were repackaged for courier shipment to each laboratory with at least 1 temperature monitoring device and, if applicable, cooling agents and insulating packaging.

A shipment destined to 1 evaluation center in India was delayed and temperature log data revealed that all RDTs (except OnSite Leishmania Ab Rapid Test Strip [CTK Biotech], which was shipped separately on different dates) in the shipment were subjected to temperatures exceeding manufacturers' recommendations (>30°C) for a period of 4 weeks. An ad hoc meeting of the VL Network recommended the exclusion of Signal-KA (Span Diagnostics) from the Banaras Hindu University center because of its requirement for storage between 2°C and 8°C. All other tests were included at the center. Upon arrival, tests were immediately stored according to the manufacturers' instructions.

Diagnostic Accuracy
Sensitivity and specificity based on pooled data per region per test are presented in Table 3. Results were highly comparable intraregionally (data not shown) but variable between regions. On the Indian subcontinent all tests performed well, with high sensitivity, which ranged from 92.8% (95% CI, 88.9%–95.4%; CrystalKA) to 100% (95% CI, 97.9%–100%; Signal KA) and high specificity, which ranged from 99.2% (CI, 97.1%–99.8%; CrystalKA) to 100% (95% CI, 97.8%–100%; Signal KA). However, in East Africa and Brazil lower sensitivity was observed, which ranged from 36.8% (95% CI, 31.1%–42.9%; CrystalKA) to 92% (95% CI, 87.8%–94.8%; IT LEISH). The sensitivity of IT LEISH was significantly better than any other products evaluated in East Africa (P < .0001) and in Brazil (P = .013).

Furthermore, in these 2 regions specificity was generally higher than sensitivity and more consistent, ranging from 90.8% (95% CI, 86.6%–93.8%; Kalazar Detect) to 98.8% (95% CI, 96.6%–99.6%; Signal KA).

There were no major differences in specificity between samples from healthy endemic controls and those of potentially cross-reactive controls in any of the regions. We therefore present specificity results based on all control samples combined.

Lot-to-Lot and Interobserver Variability
In general, agreement of test results between lots and readers was high; with a few exceptions, points ranging from even the lower margins of the CIs can still be considered "good" agreement (Supplementary Tables 2 and 3).

Thermal Stability
In the Indian subcontinent all tests assessed showed excellent baseline performance, which was maintained after incubation for 60 days at 4°C and 37°C. However, one product's performance (IT LEISH) was seriously affected by post–60 day incubation at 45°C, returning 100% invalid results (n = 24). In East Africa and Brazil, suboptimal baseline performance of 2 tests suggests samples may be at the limit of detection of some products and therefore variations at baseline probably reflect intertest variation. Nonetheless, in Brazil 2 products appear to be stable and the other (IT LEISH), despite the best baseline performance, showed reductions in performance at 35°C and 45°C but no invalid results as reported in both the Indian subcontinent and East Africa (Supplementary Table 4). In East Africa, test performance appeared to improve for one product (80% detection at baseline to 100% detection after 60 days at 45°C).

DISCUSSION

This study is the first global head-to-head comparison of performance of VL RDTs in 3 endemic regions. In the ISC, our results reflect previous reports of high accuracy of rK39 RDTs; however, our study extends to rKE16-based products and an rK39 commercial test not previously independently evaluated. Overall, our findings illustrate that in the ISC several RDTs demonstrated high sensitivity and specificity (Table 3). Furthermore, performance of these products did not seem to be significantly affected by heat stress induced, inadvertently, during transport to one center in India. This is a reassuring finding, as these circumstances can be expected to occur regularly in routine settings. However, in East Africa and Brazil, the tests performed with variable sensitivity but high specificity (3 of 4 tests, >95%). In Brazil and in 2 sites in East Africa, the rKE16-based products appeared to perform less well than rK39 products. This may be partially explained by the fact that rKE16 antibody-detecting tests are based on a recombinant antigen (Ld-rKE16) from a newly isolated Indian strain of *Leishmania donovani* (MHOM/IN/KE16/1998), whereas rK39 is based on *Leishmania infantum*, the species causing VL in Brazil [13].

Differences in product performance between regions is likely attributable to parasite diversity and/or differences in antibody concentrations which may in turn be linked to different age patterns, immune response, and nutritional status

Table 2. Characteristics of Visceral Leishmania Rapid Diagnostic Tests Included in the Evaluation

Product	Manufacturer	Catalog No.	Lateral Flow Format	Lot No.	Storage Temperature	Shelf Life	Control Line	Test Line, Bound Antigen	Sample Volume (μL)	Buffer Quantity (Drops)	Minimum Reading Time (min)	Maximum Reading Time
CrystalKA	Span Diagnostics, Ltd	56FT101	Dipstick	R2009004	4°C–30°C	18 mo	Yes	rKE16	20	5	15	30
				R2009003								
DiaMed-IT LEISH	Bio-Rad Laboratories	710124	Cassette	46240.27.01	2°C–30°C	16 mo	Yes	rK39	8–12 (10 μL mark on pipette)	Step 1: 1 (conjugate well), Step 2: 4 (wash well)	5	Not specified
				46240.28.01								
Kalazar Detect	InBios International, Inc	INS105	Dipstick	KE 2108	RT (20°C–28°C)	24 mo	Yes	rK39	20	3	10	10
				KE 1047								
Signal–KA	Span Diagnostics Ltd	56FT100	Cassette	4000002561	2°C–8°C	12 mo	Yes	rKE16	Mix 1 (20 μL) part serum with 4 parts normal saline, then use 50 μL.	Step 1: 2 Step 2: 2 Step 3: 3	2	10
				4000002604								
OnSite Leishmania Ab Rapid Test	CTK Biotech Inc	RO122S	Dipstick	F0317G2	2°C–30°C	18 mo	Yes	rK39	5	2	15	15
				F0318G2								

Table 3. Sensitivity and Specificity of Rapid Diagnostic Tests, All Regions

		East Africa		Brazil		Indian Subcontinent	
Product	Manufacturer	Sensitivity (95% CI) (n = 250)	Specificity (95% CI) (n = 250)	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI) (n = 250)	Specificity (95% CI) (n = 249)
CrystalKA	Span Diagnostics	36.8% (31.1%–42.9%)	98.0% (95.4%–99.1%)	61.5% (55.2%–67.4%)[a]	98.4% (95.9%–99.4%)[b]	92.8% (88.9%–95.4%)	99.2% (97.1%–99.8%)
DiaMed-IT LEISH	Bio-Rad Laboratories	87.2% (82.5%–90.8%)	96.4% (93.3%–98.1%)	92.0% (87.8%–94.8%)[c]	95.6% (92.2%–97.5%)[d]	98.8% (96.5%–99.6%)	97.6% (94.8%–98.9%)
Kalazar Detect	InBios International,	67.6% (61.6%–73.1%)	90.8% (86.6%–93.8%)	84.7% (79.7%–88.7%)[e]	96.8% (93.9%–98.4%)[f]	99.6% (97.8%–99.9%)	96.0% (92.8%–97.8%)
Signal–KA	Span Diagnostics	73.2% (67.4%–78.3%)	96.4% (93.3%–98.1%)	79.2% (73.7%–83.8%)[g]	98.8% (96.6%–99.6%)[h]	100% (97.9%–100%)[i]	100% (97.8%–100%)[j]
OnSite Leishmania Ab Rapid	CTK Biotech	NA	NA	NA	NA	99.6% (97.8%–99.9%)	96.8% (93.8%–98.4%)

Abbreviations: CI, confidence interval; NA, not applicable.

[a] n = 244.
[b] n = 249.
[c] n = 237.
[d] n = 248.
[e] n = 249.
[f] n = 252.
[g] n = 250.
[h] n = 254.
[i] n = 175.
[j] n = 170.

of patient. The average patient age of VL cases in each region was 24 years, 14 years, and 19 years in the Indian subcontinent, East Africa, and Brazil, respectively. Furthermore, the proportion of DAT titers >1:102 400, representing very high antibody response in the Indian subcontinent, was 58%, dropping to 40% and 27% in East Africa and Brazil, respectively; and may highlight differential antibody response/production to infection in different geographical areas.

In addition, subgroup analysis based on human immunodeficiency virus (HIV) status could not be completed owing to limited availability of information on HIV status of patients. HIV and VL coinfections are important to consider in VL endemic areas and test reliability is known to fall in some of these cases [14].

In all regions, agreement between lots or batches of the same product was good to excellent ($\kappa = 0.73$–0.98) according to Landis and Koch [12], and agreement between readers (second reading within 30 minutes of the first) was excellent ($\kappa > 0.9$). All manufacturers in this evaluation have current ISO 13485:2003 certification, a standard designed to give assurance of consistency of quality of final product; however, it cannot be guaranteed that the results here will predict results from different RDT lots. Ideally, quality control materials should be developed so that manufacturers and procurers alike can assess test lots prior to purchase to ensure that expected performance is maintained. Furthermore, training and supervision of operators must be implemented on a programmatic level to ensure the quality of testing (preparation and interpretation) at the point of care.

VL is endemic in regions where daytime temperatures can regularly exceed 30°C or 40°C, so it is likely that RDTs will be exposed to temperatures above the manufacturers' recommendations (usually 30°C) during transport, storage, or use in field settings. Heat is known to diminish the performance of some malaria RDTs [15] but there are no published reports on the thermal stability of VL RDTs. Our assessment does not mimic the fluctuating heat and humidity conditions in real-life settings, nor does it necessarily predict long-term stability in field conditions. However, results are useful to highlight potential losses to test sensitivity should similar conditions be encountered. Although RDTs evaluated at the site in India were exposed to high temperatures during transport, it is quite clear from the parallel assessments in other regions that only 1 of the 4 heat-stressed products (IT LEISH) was less stable after 60 days at 45°C (Supplementary Table 4).

Archived samples were used to avoid the time and expense associated with prospective sample collection and the complexities of comparing several products simultaneously. Samples with the fewest freeze-thaw cycles were preferentially selected, and retested by DAT as an estimate of total antibody reactivity. Because of limited volumes of sera and restrictions in shipping of biological specimens internationally, each evaluation center assembled its own panel following proficiency testing and sample revalidation using study-specific standard operating procedures and materials. The absence of significant differences in test performance within regions supports the pooling of data and suggests that samples from each laboratory were representative of the patient population in each global region.

Samples were selected for inclusion if patients were parasitologically confirmed with VL using microscopy or culture of clinical material (spleen, lymph node, bone marrow). Unfortunately, this is not 100% sensitive and patients with high parasite burden may have been preferentially selected, which may have artificially enhanced clinical accuracy of the RDTs. Furthermore, the majority of controls were healthy individuals from endemic areas (with negative DAT results) and 10% were from patients with other disease conditions that mimic VL. This distribution does not represent the "VL suspect" population and therefore may overestimate RDT specificity. This underlines the importance in clinical practice of combining RDTs with the WHO clinical case definition [16] to avoid false-positive results.

CONCLUSIONS

The 5 RDTs tested in the ISC show high sensitivity and specificity and good lot and reader agreement, and most are heat stable. RDT sensitivity is more variable in East Africa and Brazil (Table 3); in Brazil and in 2 sites in East Africa, the rKE16-based products appeared to perform less well than rK39 products. Ultimately, outside the ISC, in clinical practice VL RDT positive results may be adequate to direct treatment (when combined with the clinical case definition) but should be interpreted with caution before excluding a diagnosis of VL. In all settings, RDTs should be implemented according to predefined acceptable limits of performance and within an appropriate diagnostic algorithm.

The results of this evaluation may be used to guide procurement and highlight the need for additional research into test performance among HIV- and VL-coinfected patients and when used on whole blood rather than serum. Furthermore, our results should be combined with a detailed ease-of-use assessment performed in clinical settings to best inform procurement decisions.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online (http://www.oxfordjournals.org/our_journals/cid/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all

supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

References

1. Matlashewski G, Arana B, Kroeger A, et al. Visceral leishmaniasis: elimination with existing interventions. Lancet Infect Dis **2011**; 11:322–5.
2. Alvar J, Velez ID, Bern C, et al. Leishmaniasis worldwide and global estimates of its incidence. PLoS One **2012**; 7:e35671.
3. Harith AE, Kolk AH, Kager PA, et al. A simple and economical direct agglutination test for serodiagnosis and sero-epidemiological studies of visceral leishmaniasis. Trans R Soc Trop Med Hyg **1986**; 80:583–6.
4. Chappuis F, Rijal S, Soto A, Menten J, Boelaert M. A meta-analysis of the diagnostic performance of the direct agglutination test and rK39 dipstick for visceral leishmaniasis. BMJ **2006**; 333:723.
5. Boelaert M, El-Safi S, Hailu A, et al. 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. Trans R Soc Trop Med Hyg **2008**; 102:32–40.
6. Barnett PG, Singh SP, Bern C, Hightower AW, Sundar S. Virgin soil: the spread of visceral leishmaniasis into Uttar Pradesh, India. Am J Trop Med Hyg **2005**; 73:720–5.
7. Zijlstra EE, Nur Y, Desjeux P, Khalil EA, el-Hassan AM, Groen J. Diagnosing visceral leishmaniasis with the recombinant K39 strip test: experience from the Sudan. Trop Med Int Health **2001**; 6:108–13.
8. Chappuis F, Mueller Y, Nguimfack A, et al. 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 **2005**; 43:5973–7.
9. Ritmeijer K, Melaku Y, Mueller M, Kipngetich S, O'keeffe C, Davidson RN. Evaluation of a new recombinant K39 rapid diagnostic test for Sudanese visceral leishmaniasis. Am J Trop Med Hyg **2006**; 74:76–80.
10. Wilson EB. Probable inference, the law of succession and statistical inference. J Am Stat Assoc **1927**; 22:209–12.
11. Cohen J. A coefficients of agreement for nominal scales. Edu and Psych Meas **1960**; 20:37–46.
12. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics **1977**; 33:363–74.
13. Sivakumar R, Dey A, Sharma P, Singh S. Expression and characterization of a recombinant kinesin antigen from an old Indian strain (DD8) of *Leishmania donovani* and comparing it with a commercially available antigen from a newly isolated (KE16) strain of *L. donovani*. Infect Genet Evol **2008**; 8:313–22.
14. ter Horst R, Tefera T, Assefa G, Ebrahim AZ, Davidson RN, Ritmeijer K. Field evaluation of rK39 test and direct agglutination test for diagnosis of visceral leishmaniasis in a population with high prevalence of human immunodeficiency virus in Ethiopia. Am J Trop Med Hyg **2009**; 80:929–34.
15. Chiodini PL, Bowers K, Jorgensen P, et al. The heat stability of *Plasmodium* lactate dehydrogenase-based and histidine-rich protein 2-based malaria rapid diagnostic tests. Trans R Soc Trop Med Hyg **2007**; 101:331–7.
16. World Health Organization. Control of the leishmaniasis: report of a meeting of the WHO Expert Committee on the Control of Leishmaniases. Geneva, Switzerland: WHO, **2010**.