Research

BMJ

A meta-analysis of the diagnostic performance of the direct agglutination test and rK39 dipstick for visceral leishmaniasis

François Chappuis, Suman Rijal, Alonso Soto, Joris Menten, Marleen Boelaert

Abstract

Objective To compare the performance of the direct agglutination test and rK39 dipstick for the diagnosis of visceral leishmaniasis.

Data sources Medline, citation tracking, January 1986 to December 2004.

Selection criteria Original studies evaluating the direct agglutination test or the rK39 dipstick with clinical visceral leishmaniasis as target condition; adequate reference classification; and absolute numbers of true positive, true negative, false positive, and false negative observations available or derivable from the data presented.

Results 30 studies evaluating the direct agglutination test and 13 studies evaluating the rK39 dipstick met the inclusion criteria. The combined sensitivity estimates of the direct agglutination test and the rK39 dipstick were 94.8% (95% confidence interval 92.7% to 96.4%) and 93.9% (87.7% to 97.1%), respectively. Sensitivity seemed higher and more homogenous in the studies carried out in South Asia. Specificity estimates were influenced by the type of controls. In phase III studies carried out on patients with clinically suspected disease, the estimated specificity of the direct agglutination test was 85.9% (72.3% to 93.4%) and of the rK39 dipstick was 90.6% (66.8% to 97.9%).

Conclusion The diagnostic performance of the direct agglutination test and the rK39 dipstick for visceral leishmaniasis is good to excellent and seem comparable.

Introduction

Visceral leishmaniasis, or kala-azar, is a disease caused by a bloodborne protozoan. The disease, fatal if untreated, affects up to 500 000 people yearly, mostly in poor rural areas of east Africa (Sudan, Ethiopia, Kenya, Uganda), South Asia (India, Bangladesh, Nepal), and Latin America (Brazil).[1] Pentavalent antimonials are the most widely used drugs for visceral leishmaniasis. Treatment, requiring painful intramuscular or intravenous injections for 30 days, carries a significant toxicity and is expensive as branded formulations. Accurate diagnosis of this condition is therefore vital to avoid under-diagnosis or over-diagnosis. Microscopical examination of spleen aspirates is sensitive and specific but requires expertise to carry out the aspiration safely and to read the slides accurately. Examination of bone marrow or lymph node aspirates is equally specific but less sensitive.[w1]

Serological tests have been developed to replace parasitological methods for the diagnosis of visceral leishmaniasis in the field. The direct agglutination test was first described by Allain

and Kagan and adapted by El Harith et al.[2] [w2] [w3] The test is semiquantitative and uses microplates with V shaped wells in which increasing dilutions of patient's serum or blood is mixed with stained killed promastigotes of *Leishmania donovani*. If antibodies to the protozoan are present, agglutination is visible with the naked eye. The test has been validated in several endemic areas and is being used for diagnosis of visceral leishmaniasis in countries such as Sudan.[w4] More recently a serological test was developed against a recombinant antigen derived from a 39 amino acid repeat in *Leishmania chagasi* (rK39). It proved to be accurate when used as an enzyme linked immunosorbent assay[3–5] and was later developed in a simpler dipstick format. Patient's serum or blood is added to a strip and migrates towards the fixed rK39 antigen. A reagent reveals the presence of specific antibodies on the rK39 antigen line. The results are qualitative (positive or negative) and are available within 20 minutes. Since 1998 several brands of dipsticks have been studied in different endemic areas using various study designs.

Several authors propose the direct agglutination test and the rK39 dipstick as the most appropriate serological tests for the diagnosis of visceral leishmaniasis in the field, but an independent and global assessment of their performance is needed. Also, as visceral leishmaniasis is an infectious disease presenting with substantial variability between regions, it is crucial to examine the extent and sources of heterogeneity in published results of serological test performance and the reasons for the heterogeneity. We carried out a meta-analysis of the diagnostic performance of the direct agglutination test and the rK39 dipstick.

Materials and methods

We searched Medline through PubMed for articles published from January 1986 to December 2004 that reported on the diagnostic accuracy of the direct agglutination test and the rK39 immunochromatographic assay used on serum or blood samples for visceral leishmaniasis. The search terms were "visceral leishmaniasis", "kala-azar", "*L.donovani*", "*L.infantum*", and "*L.chagasi*" combined with "diagnostic accuracy", "sensitivity", "specificity", "validation", and "diagnostic performance". This search generated 654 papers, which we subsequently combined with the search terms "Direct Agglutination Test or DAT" and, in a second step, with "rK39 and immunochromatographic test", "immunochromatographic assay", and "ICT or dipstick". We obtained additional articles by citation tracking of review articles and original articles.

P+ *Additional tables and references w1-w38 are on bmj.com*

We included original studies only. Other inclusion criteria were current clinical visceral leishmaniasis as the target condition (not leishmanial infection or past visceral leishmaniasis); human participants; the absolute numbers of true positive, false negative, true negative, and false positive observations available or derivable from the data presented; and the reference classification judged adequate to correctly classify the target condition. We included studies that complied with these four criteria, independent of other characteristics of design quality and reporting. We excluded studies that evaluated the tests in patients coinfected with HIV and studies of the rapid version of the direct agglutination test (the fast agglutination screening test).[6]

Data extraction

We used a standard form to extract data on type of study (phase I, II, or III trials), *Leishmania* species (*L donovani, L infantum,* or *L chagasi*), country in which the study was carried out, characteristics of the participants, study design, and test results. Data were entered in a spreadsheet (Microsoft Excel).

The number and type of participants was recorded and categorised as confirmed cases or controls. Several controls were distinguished: healthy non-endemic controls (healthy people living in an area with no transmission of *Leishmania* and no exposure by travel); healthy endemic controls (people living in an endemic region but without signs or symptoms of leishmaniasis); controls with potentially cross reacting diseases (patients with confirmed disease that might give a false positive reaction to a serological test); and controls (patients with the same clinical syndrome as confirmed cases with visceral leishmaniasis ruled out by a confirmatory test with a high negative predictive value such as on spleen aspirate).

In many articles the numbers of true positive, false negative, true negative, and false positive observations were available. If not, we derived the numbers from the marginal totals and the reported sensitivity and specificity.

Assessment of study quality

Two independent readers assessed the papers according to the quality assessment of studies of diagnostic accuracy approach.[7] Disagreements were resolved by a third reader who reassessed problematic papers. The quality assessment of studies of diagnostic accuracy tool lists 14 items that together assess the quality of a study (see tables A-C on bmj.com). Important criteria are a representative spectrum of patients, correct classification of disease status, complete description of reference test and index test, blinding of the readers of the serological test to the result of the reference test and vice versa, and a low potential for verification bias. Such bias may exist if the decision to carry out the standard examination depends on the results of the examination under investigation—that is, if verification with parasitology occurred more often in patients with positive serology results.

Data analysis

We used standard formulas to calculate sensitivity, specificity, and diagnostic odds ratio for each study. Wilson's score method was used to calculate confidence intervals for sensitivities and specificities,[8] and the normal approximation with continuity correction for the log of the odds ratio for confidence intervals for the diagnostic odds ratio.[9]

We explored the relation between sensitivity and specificity by plotting the log odds of a positive test result in diseased participants (cases) against the log odds of a negative test result in non-diseased participants (controls). Publication bias was assessed by graphing the estimated sensitivity and specificity against the number of cases and controls in each study.

Meta-analyses for sensitivity and specificity were carried out using logistic regression models accounting for over-dispersion.[10] This model weights each study proportional to its sample size while allowing for heterogeneity between studies by inclusion of a dispersion variable. The robustness of this statistical approach was explored by comparison with other statistical methods and simulation. The meta-analysis of the diagnostic odds ratio was carried out using the DerSimonian and Laird random effects model, with studies weighted using the Mantel-Haenszel method.[11] To assess the heterogeneity of studies we carried out separate meta-analyses in subgroups stratified by study phase, sample size, study quality, geographical region, species of *Leishmania*, type of direct agglutination test antigen, brand of dipstick, and type of controls.

Statistical analyses were carried out with the open source statistical language and environment R 2.0.1.[12] We used the generalised linear models module (procedure "glm" using the quasibinomial distribution) to fit over-dispersed logistic regression models and the R meta-analysis library (procedure "metabin") to fit the DerSimonian and Laird random effects model.

Results

Thirteen studies evaluating the rK39 dipstick and 30 evaluating the direct agglutination test were included in the meta-analysis.[w1-w38] The 13 dipstick studies included a total of 1119 patients with visceral leishmaniasis and 2676 controls, and the 30 studies of the direct agglutination test included 1698 patients with visceral leishmaniasis and 3876 controls. The ideal method for diagnosis of visceral leishmaniasis in all studies was a positive result on microscopical examination of lymph node, bone marrow, or spleen aspirate. Diagnosis was also confirmed by detection of parasites in culture (n=4 studies) or by polymerase chain reaction (n=1) or a good clinical response to antileishmanial drugs in patients with clinically suspected disease who had positive serology results (n=4). Results of the quality assessment of studies of diagnostic accuracy evaluation are on bmj.com. No evidence of publication bias was observed. Owing to the lack of correlation between the observed sensitivities and specificities (data not shown), the individual sensitivities and specificities for diagnostic accuracy were determined.

Sensitivity, specificity, and diagnostic odds ratio

The estimated sensitivity of both tests was relatively consistent across studies, indicating that a summary of the data in a single estimate with 95% confidence interval (obtained by a statistical method that allows for some variability in results across cases) is appropriate.

Figures 1 and 2 show the results of individual and combined sensitivity estimates for the tests: the estimated sensitivity for the rK39 dipstick was 93.9% (95% confidence interval 87.7% to 97.1%). Only one study seemed inconsistent with this estimate (67.2%, 54.1% to 78.2%).[w20] The estimated sensitivity for the direct agglutination test was 94.8% (92.7% to 96.4%). This estimate and confidence intervals were generally consistent with the individual study estimates.

Some heterogeneity in sensitivity of the tests seemed to be related to the geographical location of the study (tables 1 and 2). Sensitivity seemed higher and more homogenous in studies carried out in South Asia.

The estimated specificity of both tests was heterogeneous across studies (figs 1 and 2), indicating that a summary of the data in a single estimate with 95% confidence interval is not appropriate.[13] Specificity estimates were highly influenced by the
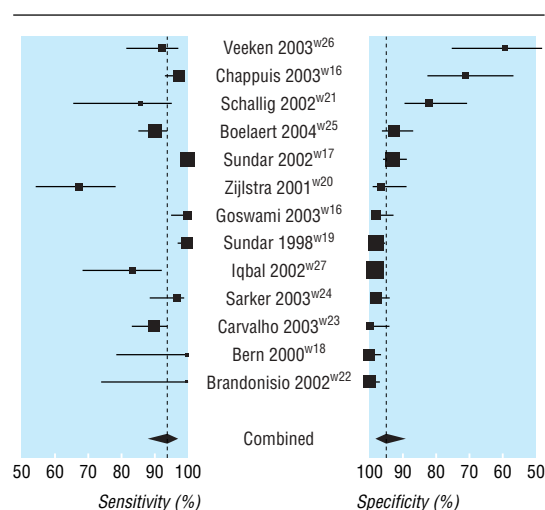
Fig 1 Estimates of sensitivity and specificity (95% confidence intervals) of rK39 dipstick

type of controls (tables 1 and 2). Studies using healthy controls showed high specificities. The estimated specificity in participants with clinically suspected disease in phase III studies of the rK39 dipstick was 90.6% (95% confidence interval, 66.8% to
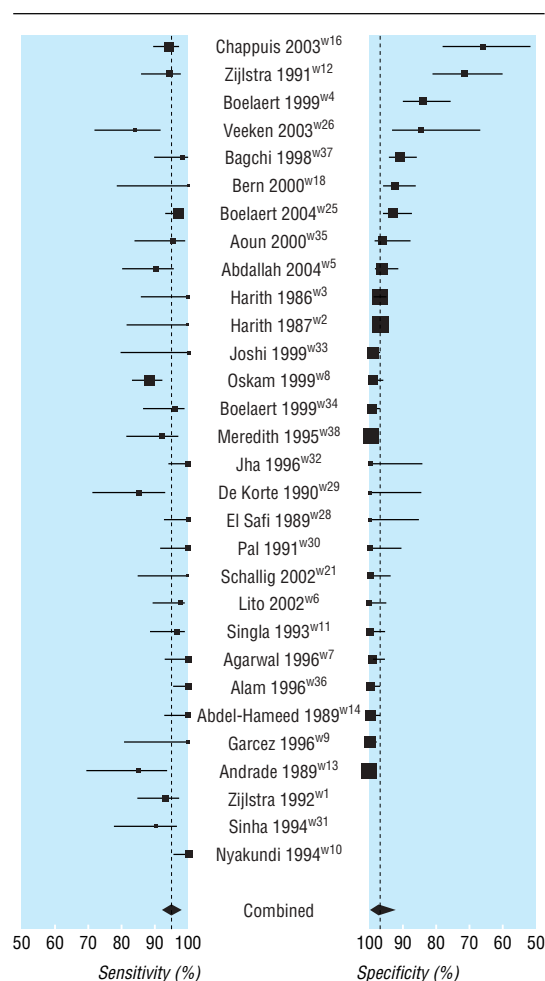

Fig 2 Estimates of sensitivity and specificity (95% confidence intervals) of direct agglutination test

Table 1 Sensitivity and specificity of rK39 dipsticks in various subgroups

Subgroups	No of studies	Sensitivity (95% CI)	No of studies	Specificity (95% CI)
All studies	13	93.9 (87.7 to 97.1)	13	95.3 (88.8 to 98.1)
Trial phase:				
I	5	86.0 (67.1 to 94.9)	5	96.9 (86.3 to 99.4)
II	4	96.5 (86.0 to 99.2)	4	96.8 (90.7 to 98.9)
III	4	94.8 (87.6 to 97.9)	4	91.2 (66.8 to 98.2)
Region:				
South Asia	7	97.1 (91.7 to 99.0)	7	95.3 (87.3 to 98.3)
East Africa	2	79.0 (46.7 to 94.2)	2	85.2 (28.2 to 98.8)
Elsewhere	4	88.8 (83.7 to 92.4)	4	97.0 (79.4 to 99.6)
Leishmania species:				
L donovani	10	94.6 (86.1 to 98.0)	10	95.2 (87.3 to 98.3)
Other	3	90.0 (84.9 to 93.5)	3	95.7 (62.7 to 99.7)
Brand:				
Kalazar Detect* (InBios International)	7	95.1 (89.5 to 97.8)	7	93.1 (84.1 to 97.2)
Other	6	91.2 (74.3 to 97.3)	6	97.2 (79.1 to 99.7)
Sample size:				
Cases or controls ≤100	8	89.6 (77.4 to 95.6)	6	89.0 (71.5 to 96.3)
Cases or controls >100	5	95.6 (88.2 to 98.5)	7	97.2 (94.1 to 98.7)
Quality assessment of studies of diagnostic accuracy:				
Score ≤7	4	92.7 (80.6 to 975)	4	93.9 (67.8 to 99.1)
Score >7	9	94.2 (85.6 to 97.8)	9	95.9 (89.2 to 98.5)
Health states of controls†:				
Healthy non-endemic	—	NA	0	NA
Healthy endemic	—	NA	10	95.9 (90.6 to 98.3)
Cross reacting diseases	—	NA	7	97.1 (88.5 to 99.3)
Clinically suspected disease	—	NA	7	93.0 (77.5 to 98.1)
Clinically suspected disease in phase III studies	—	NA	4	90.6 (66.8 to 97.9)

NA=not applicable.
*Previously known as Insure Rapid Test for Visceral Leishmaniasis.
†Several studies included more than one type of control.

97.9%) and of the direct agglutination test was 85.9% (72.3% to 93.4%).

For most studies the diagnostic odds ratio exceeded 100; in participants with clinically suspected disease the diagnostic odds ratio was generally lower than in other participants. In participants with clinically suspected disease from phase III studies, the combined estimate of the diagnostic odds ratio for the rK39 dipstick was 128 (26 to 629) and for the direct agglutination test was 76 (13 to 433).

Discussion

This meta-analysis showed good to excellent diagnostic performance of the rK39 dipstick and the direct agglutination test for visceral leishmaniasis in populations not known to be infected with HIV. The overall sensitivity of both tests was excellent but was lower in studies from East Africa than in those from South Asia. Several studies have already commented on the lower sensitivity of rK39 based dipsticks in Sudan.[14] [w20] For reasons that remain unclear, Sudanese patients seem to develop lower titres of antibodies against rK39 than do Indian patients (S G Reed, personal communication, 2006).

The estimated specificity of the rK39 dipstick and the direct agglutination test was too heterogeneous across studies to allow for a reliable measurement of a single estimate. We thus stratified the meta-analysis by different subgroups, showing that the type of controls significantly influenced the estimates for specificity.

Table 2 Sensitivity and specificity of direct agglutination test in various subgroups

Subgroups	No of studies	Sensitivity (95% CI)	No of studies	Specificity (95% CI)
All studies (n=30)	29	94.8 (92.7 to 96.4)	27	97.1 (93.9 to 98.7)
Trial phase:				
I	20	94.3 (91.5 to 96.2)	17	98.1 (94.2 to 99.4)
II	5	97.7 (87.4 to 99.6)	5	97.2 (92.5 to 99.0)
III	4	94.3 (87.9 to 97.4)	5	90.9 (75.9 to 96.9)
Region:				
South Asia	11	97.1 (94.9 to 98.4)	10	95.7 (88.1 to 98.5)
East Africa	11	93.2 (89.1 to 95.8)	10	96.1 (89.2 to 98.6)
Elsewhere	7	92.8 (86.8 to 96.2)	7	99.8 (97.5 to 100)
Leishmania species:				
L donovani	23	95.1 (92.7 to 96.7)	21	96.4 (92.5 to 98.4)
Other	6	93.0 (85.1 to 96.9)	6	99.7 (94.6 to 100)
Type of antigen:				
Freeze dried	4	89.0 (84.1 to 92.5)	4	99.1 (74.4 to 100)
Aqueous	25	96.2 (94.2 to 97.5)	23	96.7 (93.0 to 98.5)
Sample size:				
Cases or controls ≤100	26	95.7 (93.3 to 97.3)	12	91.3 (81.7 to 96.1)
Cases or controls >100	3	92.8 (85.6 to 96.5)	15	98.2 (96.4 to 99.1)
Quality assessment of studies of diagnostic accuracy:				
Score ≤7	19	96.0 (92.5 to 97.9)	17	98.6 (96.7 to 99.4)
Score >7	10	93.7 (90.7 to 95.8)	10	92.6 (83.7 to 96.9)
Health state of controls*:				
Healthy non-endemic		NA	8	100.0 (98.2 to 100)
Healthy endemic		NA	20	98.7 (97.1 to 99.5)
Cross reacting diseases		NA	16	98.8 (95.6 to 99.7)
Clinically suspected disease		NA	8	82.6 (70.4 to 90.4)
Clinically suspected disease in phase III studies		NA	4	85.9 (72.3 to 93.4)

NA=not applicable.

*Several studies included more than one type of control.

Studies that included healthy controls showed very high specificity. The estimate of the specificity of the rK39 dipstick and direct agglutination test that is most relevant for clinical practice is obtained in phase III studies—that is, those recruiting patients with clinically suspected disease consecutively and prospectively in a representative clinical setting. Our meta-analysis includes only four such studies for both the rK39 dipstick and the direct agglutination test, resulting in specificity estimates with wide confidence intervals: rK39 dipstick 90.6% (66.8% to 97.9%) and direct agglutination test 85.9% (72.3% to 93.4%).[w16 w19 w25 w26 w34] Some degree of underestimation of these specificity estimates might have occurred, as spleen puncture (the most sensitive reference standard) was not systematically carried out among patients with suspected disease in most of these phase III studies.[w4 w16 w19 w25 w34] Previous decision analysis has shown that a serological test with these high sensitivity and "acceptable" specificity levels is a better alternative for case detection algorithms than are tests with insufficient sensitivity, such as bone marrow or lymph node smears.[15]

By focusing the meta-analysis on sensitivities and specificities, we did not take into account the effect of heterogeneity introduced by variations in diagnostic thresholds. Different diagnostic thresholds to define positive and negative test results might have been used across the studies. Diagnostic thresholds might differ as a result of explicit factors such as differences in titration cut off points for the direct agglutination test or due to unknown factors such as differing reactivity of antigen batches or variations between observers.[16] We did not, however, detect any correlation between sensitivity and specificity, indicating that variations due to diagnostic thresholds were insignificant compared with other sources of heterogeneity between the studies.

Our meta-analysis supports the incorporation of these diagnostic tests in country policies. As the performance of the direct agglutination test and the rK39 dipstick is comparable, the choice should be made on the basis of other criteria such as cost, feasibility, and sustainability. rK39 dipsticks are cheaper and easier to use than the direct agglutination test, but a long term commitment of the current dipstick manufacturers is needed as two companies have already stopped production. The brand of dipstick did not significantly influence the pooled sensitivity and specificity estimates in our analysis. Nevertheless, dipsticks from different brands or from different generations of the same brand are likely to show variable diagnostic performance. Moreover, our meta-analysis did not include the results of recent studies carried out in East Africa and Asia with the DiaMed Opti-Leish IT (DiaMed AG, Switzerland), a newly manufactured rK39 dipstick that may perform better than other dipstick formats in East Africa.[14 17 18] Therefore we advise that each new brand of rK39 dipstick should be individually validated in each endemic region before being integrated in a diagnostic algorithm.

The development of the direct agglutination test and, more recently, the rK39 dipstick has brought a major improvement in the diagnosis of visceral leishmaniasis in the field. These tests urgently need to be more widely used in the field. Nevertheless, the limited sensitivity of the rK39 dipstick in Sudan and the suboptimal specificity of these serological tests when used for patients with clinically suspected visceral leishmaniasis necessitate the research and development of better diagnostic tools for this fatal disease.[19] Moreover, a systematic adoption of quality criteria for diagnostic accuracy studies,[8 20] would not only lead to more reliable results but also to quicker and more precise recommendations for policy makers.

What is already known on this topic

The direct agglutination test and rK39 dipsticks have been proposed as the most appropriate serological tests for diagnosis of visceral leishmaniasis in the field

A global evaluation of these tests, including an assessment of the heterogeneity of their performance, is lacking

What this study adds

The direct agglutination test and the rK39 dipstick showed similar diagnostic performance for visceral leishmaniasis, ranging from good to excellent

The performance of dipstick tests for diagnosis in Sudan needs to be improved

the final manuscript. AS and JM reviewed data analysis, did the statistical analysis, and reviewed the manuscript. MB supervised all phases of the study including the writing of the manuscript.

1 Desjeux P. Leishmaniasis. Public health aspects and control. *Clin Dermatol* 1996;14:417-23.
2 Allain DS, Kagan IG. A direct agglutination test for leishmaniasis. *Am J Trop Med Hyg* 1975;24:232-6.
3 Singh S, Gilman-Sachs A, Chang KP, Reed SG. Diagnostic and prognostic value of K39 recombinant antigen in Indian leishmaniasis. *J Parasitol* 1995;81:1000-3.
4 Badaro R, Benson D, Eulalio MC, Freire M, Cunha S, Netto EM, et al. rK39: a cloned antigen of Leishmania chagasi that predicts active visceral leishmaniasis. *J Infect Dis* 1996;173:758-61.
5 Zijlstra EE, Daifalla NS, Kager PA, Khalil EA, El-Hassan AM, Reed SG, et al. rK39 enzyme-linked immunosorbent assay for diagnosis of Leishmania donovani infection. *Clin Diagn Lab Immunol* 1998;5:717-20.
6 Hailu A, Kroon CC, Schoone GJ, Berhe N, Schallig HD, Kager PA. Sero-epidemiological assessment and diagnosis of visceral leishmaniasis in an endemic locality using fast agglutination screening test (FAST). *Acta Trop* 2002;83:93-101.
7 Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
8 Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;17:857-72.
9 Zhou X-H, Obuchowski NA, McClish DK. *Statistical methods in diagnostic medicine*. New York: Wiley, 2002.
10 Collet D. *Modelling binary data*. London: Chapman and Hall, 2003.
11 Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic reviews in health care*. London: BMJ books, 2001:285-311.
12 R Development Core Team. R: A language and environment for statistical computing. In: R foundation for statistical computing, Vienna, Austria, 2004:www.R-project.org (accessed 15 Nov 2005).
13 Deeks JJ. Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;323:157-62.
14 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.
15 Boelaert M, Lynen L, Desjeux P, Van der Stuyft P. Cost-effectiveness of competing diagnostic-therapeutic strategies for visceral leishmaniasis. *Bull World Health Organ* 1999;77:667-74.
16 Deeks JJ. Systematic reviews of evaluations of diagnostic and screening test. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic reviews in health care*. London: BMJ books, 2001:248-82.
17 Chappuis F, Mueller Y, Nguimfack A, Rwakimari JB, Couffignal S, Boelaert M, et al. Diagnostic accuracy of two rK39 antigen-based dipsticks and the Formol Gel test for the rapid diagnosis of visceral leishmaniasis in northeastern Uganda. *J Clin Microbiol* 2005;43(12):5973-7.
18 Sundar S, Maurya R, Singh RK, Bharti K, Chakravarty J, Parekh A, et al. Rapid, noninvasive diagnosis of visceral leishmaniasis in India: comparison of two immunochromatographic strip tests for detection of anti-K39 antibody. *J Clin Microbiol* 2006;44:251-3.
19 Mabey D, Peelin RW, Ustianowsk A, Perkin MD. Tropical infectious diseases: diagnostics for the developing world. *Nat Rev Microbiol* 2004;2:231-40.
20 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* 2003;49:7-18.

Geneva University Hospitals, Travel and Migration Medicine Unit, rue Micheli-du-Crest 24, 1211 Geneva 14, Switzerland
François Chappuis *lecturer*

BP Koirala Institute of Health Sciences, Dharan, Nepal
Suman Rijal *professor*

Hospital Nacional Hipolito Unanue, Universidad Peruana Cayetano Heredia, Lima, Peru
Alonso Soto *investigator*

Institute of Tropical Medicine, Antwerp, Belgium
Joris Menten *statistician*
Marleen Boelaert *professor*

Correspondence to: F Chappuis francois.chappuis@hcuge.ch