

Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease (Review)

Boelaert M, Verdonck K, Menten J, Sunyoto T, van Griensven J, Chappuis F, Rijal S



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2014, Issue 6

<http://www.thecochranelibrary.com>

WILEY

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	2
OBJECTIVES	4
METHODS	5
RESULTS	9
Figure 1.	10
Figure 2.	12
Figure 3.	13
Figure 4.	14
Figure 5.	15
Figure 6.	17
Figure 7.	18
Figure 8.	20
Figure 9.	22
Figure 10.	24
Figure 11.	25
Figure 12.	26
Figure 13.	27
Figure 14.	28
DISCUSSION	34
AUTHORS' CONCLUSIONS	35
ACKNOWLEDGEMENTS	36
REFERENCES	36
CHARACTERISTICS OF STUDIES	45
DATA	107
Test 1. rK39 immunochromatographic test.	107
Test 2. KAtex.	109
Test 3. FAST.	109
Test 4. rK26 immunochromatographic test.	110
Test 5. rK39 Primary Analysis.	110
Test 6. rKE16 immunochromatographic test.	111
ADDITIONAL TABLES	111
APPENDICES	114
CONTRIBUTIONS OF AUTHORS	118
DECLARATIONS OF INTEREST	118
SOURCES OF SUPPORT	119
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	119

[Diagnostic Test Accuracy Review]

Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease

Marleen Boelaert¹, Kristien Verdonck¹, Joris Menten¹, Temmy Sunyoto¹, Johan van Griensven¹, Francois Chappuis², Suman Rijal³

¹Institute of Tropical Medicine, Antwerp, Belgium. ²Division of International and Humanitarian Medicine, Geneva University Hospitals, Geneva, Switzerland. ³Department of Internal Medicine, BP Koirala Institute of Health Sciences, Sunsari, Nepal

Contact address: Marleen Boelaert, Institute of Tropical Medicine, Antwerp, Belgium. mboelaert@itg.be.

Editorial group: Cochrane Infectious Diseases Group.

Publication status and date: New, published in Issue 6, 2014.

Review content assessed as up-to-date: 3 December 2013.

Citation: Boelaert M, Verdonck K, Menten J, Sunyoto T, van Griensven J, Chappuis F, Rijal S. Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease. *Cochrane Database of Systematic Reviews* 2014, Issue 6. Art. No.: CD009135. DOI: 10.1002/14651858.CD009135.pub2.

Copyright © 2014 The Authors. The Cochrane Database of Systematic Reviews published by John Wiley & Sons, Ltd. on behalf of The Cochrane Collaboration. This is an open access article under the terms of the [Creative Commons Attribution-Non-Commercial](#) Licence, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

ABSTRACT

Background

The diagnosis of visceral leishmaniasis (VL) in patients with fever and a large spleen relies on showing *Leishmania* parasites in tissue samples and on serological tests. Parasitological techniques are invasive, require sophisticated laboratories, consume time, or lack accuracy. Recently, rapid diagnostic tests that are easy to perform have become available.

Objectives

To determine the diagnostic accuracy of rapid tests for diagnosing VL in patients with suspected disease presenting at health services in endemic areas.

Search methods

We searched MEDLINE, EMBASE, LILACS, CIDG SR, CENTRAL, SCI-expanded, Medion, Arif, CCT, and the WHO trials register on 3 December 2013, without applying language or date limits.

Selection criteria

This review includes original, phase III, diagnostic accuracy studies of rapid tests in patients clinically suspected to have VL. As reference standards, we accepted: (1) direct smear or culture of spleen aspirate; (2) composite reference standard based on one or more of the following: parasitology, serology, or response to treatment; and (3) latent class analysis.

Data collection and analysis

Two review authors independently extracted data and assessed quality of included studies using the QUADAS-2 tool. Discrepancies were resolved by a third author. We carried out a meta-analysis to estimate sensitivity and specificity of rapid tests, using a bivariate normal model with a complementary log-log link function. We analysed each index test separately. As possible sources of heterogeneity, we explored: geographical area, commercial brand of index test, type of reference standard, disease prevalence, study size, and risk of bias (QUADAS-2). We also undertook a sensitivity analysis to assess the influence of imperfect reference standards.

Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease (Review)

Copyright © 2014 The Authors. The Cochrane Database of Systematic Reviews published by John Wiley & Sons, Ltd. on behalf of The Cochrane Collaboration.

Main results

Twenty-four studies containing information about five index tests (rK39 immunochromatographic test (ICT), KAtex latex agglutination test in urine, FAST agglutination test, rK26 ICT, and rKE16 ICT) recruiting 4271 participants (2605 with VL) were included. We carried out a meta-analysis for the rK39 ICT (including 18 studies; 3622 participants) and the latex agglutination test (six studies; 1374 participants). The results showed considerable heterogeneity. For the rK39 ICT, the overall sensitivity was 91.9% (95% confidence interval (95% CI) 84.8 to 96.5) and the specificity 92.4% (95% CI 85.6 to 96.8). The sensitivity was lower in East Africa (85.3%; 95% CI 74.5 to 93.2) than in the Indian subcontinent (97.0%; 95% CI 90.0 to 99.5). For the latex agglutination test, overall sensitivity was 63.6% (95% CI 40.9 to 85.6) and specificity 92.9% (95% CI 76.7 to 99.2).

Authors' conclusions

The rK39 ICT shows high sensitivity and specificity for the diagnosis of visceral leishmaniasis in patients with febrile splenomegaly and no previous history of the disease, but the sensitivity is notably lower in east Africa than in the Indian subcontinent. Other rapid tests lack accuracy, validation, or both.

PLAIN LANGUAGE SUMMARY

Rapid diagnostic tests for visceral leishmaniasis

Visceral leishmaniasis (or kala-azar) is caused by a parasite, results in fever, a large spleen and other health problems, occurring in India, Bangladesh and Nepal, east Africa, the Mediterranean region and Brazil. Without treatment people die, and proper treatment can result in cure, so diagnosis is important. Many of the tests that are used to determine if a person has visceral leishmaniasis are complicated, costly, painful and sometimes dangerous for the patients. Now rapid diagnostic tests that are safe and easy to perform are available.

This Cochrane review describes how accurate these rapid diagnostic tests are for diagnosing visceral leishmaniasis. We summarize those studies that evaluated the rapid tests in people who, according to their physicians, could have the disease. We only included studies in which the researchers had used established methods to distinguish the people with visceral leishmaniasis from those who did not have the disease.

We found 24 studies, which contained information about five different rapid tests. A total of 4271 people participated in these studies. One of the rapid tests (called the rK39 immunochromatographic test) gave correct, positive results in 92% of the people with visceral leishmaniasis and it gave correct, negative results in 92% of the people who did not have the disease. This test worked better in India and Nepal than in east Africa. In India and Nepal, it gave correct, positive results in 97% of the people with the disease. In east Africa, it gave correct, positive results in only 85% of the people with the disease.

A second rapid test (called latex agglutination test) gave correct, positive results in 64% of the people with the disease and it gave correct, negative results in 93% of the people without the disease. For the other rapid tests evaluated, there are too few studies to know how accurate they are.

BACKGROUND

Target condition being diagnosed

Visceral leishmaniasis (VL), also known as kala-azar, is a life-threatening systemic disease caused by the obligate intracellular protozoan, *Leishmania*, and transmitted through the bites of phlebotomine sand flies (Herwaldt 1999). Leishmanial infection can

cause a diverse spectrum of diseases, among which VL is the most severe form and which is almost always fatal without adequate, timely treatment. The species that causes VL is *Leishmania donovani* in Asia and eastern Africa, and *L. infantum* in Europe, North Africa and Latin America (Boelaert 2000). The geographical distribution of VL is often limited to well-identified endemic foci, but it has also emerged as epidemics (Seaman 1996) and as one of the opportunistic infections in human immunodeficiency virus

(HIV)-positive patients (Alvar 2008; Pasquau 2005). VL is a neglected disease that affects the poorest and most vulnerable people in rural, remote settings where there is limited access to health care. The estimated incidence is 200,000 to 400,000 cases per year, of which more than 90% are reported from India, Bangladesh, Ethiopia, Sudan, South Sudan, and Brazil (Alvar 2012). The two main transmission modes of VL are anthroponotic and zoonotic. The human-to-human transmission is predominant in the Indian subcontinent and in eastern Africa, while the zoonotic transmission mode is found in the Mediterranean region and the Americas. The zoonotic transmission mode involves dogs as the main parasite reservoir (Chappuis 2007).

The pathogenesis of VL entails a complex interaction between the characteristics of the parasite and the host (Rittig 2000). The parasite life cycle involves the replication of the promastigote form in the gut of the female sand fly and is completed when she takes a blood meal on a non-immune host. In the human body, the *Leishmania* promastigote evolves into an amastigote form (without flagellum), which colonizes the macrophages in the liver, spleen and bone marrow. The control of the infection depends on the characteristics of the host, especially an intact cell-mediated immunity in the form of a T-helper type 1 response, involving the secretion of interleukin 12 and interferon-gamma. The failure of this cell-mediated immunity in some, but not all, of the infected individuals ultimately leads to the clinical manifestations of VL disease (Murray 2005).

The main clinical manifestations of VL are fever of insidious onset, weakness, loss of appetite, weight loss, abdominal distension due to enlargement of the spleen or the liver or both, lymph node enlargement, and a low blood cell count (pancytopenia). In children, other symptoms include diarrhoea, coughing, abdominal pain, and growth retardation. In any age group, if left untreated, the disease will progress with time, causing debilitation, bleeding, susceptibility to secondary infection and, eventually, death. Antileishmanial treatment is life-saving but non-response or relapse can occur. The pentavalent antimonials, sodium stibogluconate and meglumine antimoniate, despite their toxicity, have been the mainstay of VL treatment in many areas for decades. Alternatives include (liposomal) amphotericin B, paromomycin, and miltefosine. Combination therapy is the suggested way forward to increase treatment efficacy, prevent resistance, and reduce treatment duration and cost (van Griensven 2010). It is important that diagnostic tests for VL do not give false-negative results because VL may be fatal. Neither should they give false-positive results in order to prevent people without VL from receiving the toxic VL treatment. As stated above, asymptomatic or subclinical infections also occur, and in cross-sectional surveys in endemic areas, a significant proportion of the healthy individuals present with anti-leishmanial antibodies. In longitudinal studies, the ratio between incident clinical cases and incident asymptomatic infections ranged from 1:5 in Kenya (Ho 1982) to 1:8 in Brazil (Evans 1992), 1:9 in Ethiopia (Hailu 2009), 1:8.9 in India and Nepal (Ostyn 2011),

and 1:11 and 2.6:1 in Sudan (Zijlstra 1994). Furthermore, antibodies (seropositivity) in VL patients tend to persist for many years after cure (De Almeida 2006; Hailu 1990).

Index test(s)

Rapid diagnostic tests (RDTs) are defined as equipment-free diagnostic devices that do not require highly skilled laboratory staff. The results of an RDT can be read easily within minutes, or at most an hour or two (PATH 2008). Most RDTs work by capturing either an antigen or an antibody on a solid surface and then attaching molecules to them that allow detection by the naked eye. The technology used is mainly immunochromatography with a dipstick or lateral flow format. These immunochromatographic tests (ICTs) are used for VL diagnosis in dipstick or cassette format, using protein isolated from *Leishmania* sp as the antigen. The recombinant form of the 39 amino-acid-sequence from *L. chagasi* is the most widely used and is known as rK39. Other recombinant antigens such as rK9, rK16, rK26 and rK28 have also been evaluated. Currently, there are several on-going projects to develop a VL test based on antigen detection, for example, the latex agglutination test detecting a urinary leishmanial antigen.

The rK39 antigen was first used in an enzyme-linked immunosorbent assay (ELISA) format, but the newer dipstick or strip formats are increasingly being used in the field or away from the main centres. These rK39 ICTs give an immediate result (typically between 10 and 20 minutes) and give a binary reading (positive or negative). The test procedure involves adding the patient's blood or serum with diluent buffer on the strip. When present in the blood or serum, specific antibodies against rK39 antigens, which are bound to the strip, can be read with the naked eye in the test window. There is a limited number of commercially available RDTs for VL, such as the IT-LEISH® (DiaMed AG, Switzerland - now Biorad, France), Kalazar Detect® (InBios International, USA), Onsite Leishmania Ab Rapid Test (CTK Biotech, USA) as well as a number of prototypes under various formats (Cunningham 2012).

rK39 ICTs cannot be used to diagnose VL in people with a past history of VL due to the persistence of antibodies after cure. In addition, to avoid diagnosis of asymptomatic infections, these tests must be applied in patients suspected to have VL, ie patients with prolonged fever and splenomegaly (WHO 2010).

Clinical pathway

Prior test(s)

The RDTs are meant to be used in patients with a clinical suspicion of VL without previous history of VL. According to the case definition proposed by the World Health Organization (WHO),

VL is an illness with prolonged irregular fever, splenomegaly and weight loss as its main symptoms (WHO 2010).

Role of index test(s)

The RDTs were specifically developed for field use in VL-endemic areas. If they are sufficiently accurate, they could be used for the early diagnosis of VL at both peripheral and central levels. A positive RDT result would then confirm the diagnosis of VL in clinically suspected patients and allow the start of treatment (WHO 2010).

Alternative test(s)

A definite VL diagnosis is provided by the demonstration of the parasite through a microscopic visualization from spleen, bone marrow or lymph node aspirates. These invasive procedures are not always feasible in field settings, and using them is by no means free of risk: spleen aspirates are the most sensitive (93% to 99%) but the aspiration carries a rare but fatal risk of bleeding (Zijlstra 1992). Splenic aspiration should only be performed if there is rapid access to blood transfusion in case of bleeding and if there are no contraindications. Clinical contraindications include signs of active bleeding, jaundice, pregnancy, a barely palpable spleen and a bad general condition. Biological contraindications include severe anaemia, a prolonged prothrombin time and a low platelet count (WHO 2010). Bone marrow and lymph node aspiration are safer, but both have lower sensitivity (Sarker 2004; Babiker 2007). Parasitological confirmation through culture or molecular techniques such as a polymerase chain reaction (PCR) is possible, but their complexity restricts their use as a routine diagnosis method. In addition, in endemic areas, a substantial proportion of healthy individuals have parasite DNA in the blood, which is detected by PCR (Bhattarai 2009).

Serological tests based on indirect fluorescence antibody, ELISA and Western blot were developed but their use is limited in the first-line health services of endemic areas as they require a too sophisticated laboratory infrastructure. A more useful antibody detection test is the direct agglutination test (DAT), a semi-quantitative method based on visual agglutinations obtained by the increased dilution of blood or serum mixed with stained, killed parasites in V-shaped wells (Harith 1986). It has been extensively validated and used in the field. Nonetheless, it cannot be categorized as an RDT because it requires a degree of laboratory skills, facilities and equipment, and the result is ready only after overnight incubation (Boelaert 1999).

Rationale

Correctly diagnosing VL disease is crucial as the signs and symptoms of VL are not specific enough to differentiate the condition from chronic malaria, schistosomiasis or other systemic infec-

tions. VL should be suspected in patients presenting with fever and splenomegaly, but confirmation is needed. The classical method relying on microscopic smears from tissue aspirates (spleen, bone marrow, lymph node) are unsuitable for settings with limited resources. Other methods such as antibody detection are more feasible, most notably the DAT and the rK39 antigen-based ICT. Both have been specifically developed in the past two decades for use in such contexts and have shown high diagnostic accuracy in most endemic areas (Boelaert 2004; Chappuis 2006).

The major drawbacks of these serological methods are linked to the antibodies that remain detectable after a cure and those that are due to past or present asymptomatic infection present in a sizeable proportion of residents of endemic areas. The use of a clinical case definition and adequate medical history taking should help to avoid false positives, yet it underscores the need for better VL diagnosis tests that are specific to acute stage disease. Ideally, the test should be highly sensitive, as VL is potentially fatal, and it should also be specific, as presumptive treatment cannot be fully justified with the current treatment regimens (den Boer 2006).

The RDTs for VL, mainly but not exclusively the rK39-based ICTs, seem to be the current solution for field diagnosis in remote settings: their ease of use, convenience and cost make them potentially advantageous to increase patients' access to VL diagnosis and treatment. The current WHO recommendation for VL case management includes the use of the rK39-based ICT as the basis for initiating treatment and it has also been adopted as the diagnostic tool in first-line services in the VL elimination initiative in the Indian subcontinent (WHO 2005). In other areas, the RDTs have been used as the first test in diagnosis-treatment algorithms that often incorporate other test(s), such as the DAT or tissue aspiration (Raguenaud 2007; Veeken 2003;).

Despite its operational advantages, some regional variability in rK39-based ICT performance has been observed (Chappuis 2006). Other prototype tests have been proposed as RDTs, but their real value in clinical practice is unclear. Furthermore, in the published literature, many phase II studies (with a sub-optimal, case-control design) might overshadow the more limited information of better quality based on phase III studies (including consecutive patients suspected to have VL). Thus, it is of utmost importance to synthesize the available evidence in this rapidly evolving field. Knowing the diagnostic accuracy of these RDTs would help inform policy makers and stakeholders on how best to use them in VL control in diverse settings. The role of RDTs in the different epidemiological contexts could be better defined, giving clearer evidence on their accuracy, and this review aims to contribute to exactly this goal.

OBJECTIVES

To determine the diagnostic accuracy of rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease

presenting at health services in endemic areas.

Secondary objectives

Investigation of sources of heterogeneity

We investigated differences in diagnostic accuracy in relation to test conditions (index and reference: commercial brand of index test, type of reference standard), geographical region (the Indian subcontinent, eastern Africa, Latin America and the Mediterranean region), disease prevalence in the sample, and study quality.

METHODS

Criteria for considering studies for this review

Types of studies

Only original diagnostic accuracy studies of the phase III type (Zhou 2002) were included in the review, that is, prospective or retrospective cohort studies (meaning studies with patients who were consecutively enrolled, be it prospectively or retrospectively, and in which all patients were given the index test and reference standard), or studies that enrolled a random selection of patients from a series of patients. Randomized controlled trials in which patients were randomized to one of several index tests and all received the reference test could also be included.

Participants

Patients with a clinical suspicion of VL, that is, those who are febrile for more than two weeks, and present with splenomegaly; according to the WHO case definition (WHO 2010) presenting at health services in endemic areas.

We excluded studies with participants:

- who were previously treated for VL (non-responders or relapsed cases); and
- who had signs and symptoms of other forms of leishmaniasis, such as post kala-azar dermal leishmaniasis (PKDL).

Studies in which only a subgroup of participants was eligible for the review were included if it was possible to extract relevant data specific to that subgroup.

Studies of patients with HIV or other co-infections were eligible for this review.

Index tests

All types of RDTs for VL, with results that are read out within one hour, regardless of the manufacturer. The RDTs could be assessed alone or in comparison with other tests.

Target conditions

The target condition was restricted to current clinical VL, and did not include asymptomatic leishmanial infection or VL in the past.

Reference standards

We accepted the following reference standards for the diagnosis of active VL (adapted from Boelaert 2007):

- reference standard including direct smears or culture of splenic aspirate;
- composite reference standard (Sullivan 2003) based on a combination of several tests. Test algorithms could include smears or culture of tissue aspirate, serology (other than the index test), or clinical arguments; and
- latent class analysis (LCA) (Zhou 2002) based on one or more of the following: smears or culture of tissue aspirates, serology (other than the index test), clinical response to antimonial treatment; or specific clinical signs (pancytopenia, darkened skin). To be selected, the studies had to assess the conditional independence assumption between the tests, and, if conditional dependence was expected, they had to use appropriate statistical methods (Dendukuri 2001). More information about LCA is given in Appendix 1.

Search methods for identification of studies

We attempted to identify all relevant studies regardless of their publication status (published, unpublished, in press, ongoing). No language or date limits were applied. We limited our searches to studies conducted in humans.

Electronic searches

To identify all relevant studies, we searched the following databases using the search terms and strategy described in Appendix 2: MEDLINE, EMBASE, LILACS, CIDG SR, the Cochrane Central Register of Controlled Trials (CENTRAL) published in *The Cochrane Library*, ISI Web of Knowledge (Science Citation Index Expanded), Medion, Arif, and CCT.

We also searched the metaRegister of Controlled Trials (mRCT) and the WHO Clinical Trials Registry Platform using “leishman*” and “diagnos* OR RDT*” as search terms.

MeSH and other search terms included: visceral leishmaniasis, kala-azar, *Leishmania donovani*, *Leishmania infantum*, *Leishmania*

chagasi, K39, immunoassay, chromatography, dipstick, immunochromatography, rapid diagnostic tests; RDT; antibody detection, serology.

Searching other resources

We checked the reference lists of the studies identified by the above methods. We also checked conference proceedings and reports from the WHO.

Data collection and analysis

Selection of studies

Two review authors assessed the titles and abstracts identified through the search strategy. All potentially relevant articles were retrieved in full and two authors (MB and TS or MB and KV) independently examined them for inclusion in the present review using an assessment form. Any discrepancies were resolved by discussion with a third author (JvG).

Data extraction and management

We extracted a standard set of data from each study, using a specifically modified data extraction form. If only a subgroup of participants in a study was eligible, we extracted the data for that particular subgroup.

Two authors independently extracted the data (MB and JvG). The discrepancies were resolved by discussion with a third author (KV). For each study, data were extracted on:

Study ID	First author, year of publication.
Clinical features and settings	Presenting signs and symptoms, clinical setting.
Participants	Sample size, summary statistics of age (mean or median), sex (% male or female), HIV status (% positive, if available), country
Study design	Whether the sample was consecutive or random. Were consecutive patients enrolled retrospectively or prospectively? If the study evaluated more than one RDT, how were tests allocated to individuals, or did each individual receive all tests?
Target condition	<i>Leishmania</i> species, clinical features.
Reference standard	The reference standard test(s) used. Who performed the reference standard test(s) and where? How many observers or repeats were used? How were discrepancies between observers resolved? If not all patients received the reference tests, how many did not (and what proportion were they of the total)? If any participant received a different reference test, what reasons were stated for this, and how many participants were involved?
Index test	The commercial name, with batch number, if provided. Transport and storage conditions. Details of the test operators.
Index test results and reference standard results	Number of missing, not interpretable or doubtful results (for both index and reference tests) For each index test (and for each acceptable reference standard, if more than one reference standard was reported), the number of true and false positives, false and true negatives (2 x 2 contingency table); or for studies that used LCA as the reference standard: estimates and 95% confidence intervals (CI) for the prevalence and for

(Continued)

	the sensitivity and specificity of each index test
Notes	Anything else of relevance

Assessment of methodological quality

We first assessed the standard signalling questions of the QUADAS-2 tool ([Whiting 2011](#)) and decided to omit one question (“If a threshold was used, was it pre-specified?”) and not to add any specific signalling questions for this review.

The QUADAS-2 tool comprises four domains: patient selection, index test, reference standard, and flow and timing. We considered that the risk of bias for a certain domain was low if all the signalling questions for that domain indicated a low risk of bias. If the answer to at least one signalling question for a certain domain was ‘no’, the risk of bias for that domain was considered to be high. If none of the answers was ‘no’, but the answer to at least one question was ‘unclear’, the risk of bias for that domain was considered to be unclear.

Two review authors (JVG and MB) then independently applied the tool to the included records. For each record, they assessed risk of bias and concerns about applicability. Several of the original studies included in this review were conducted by the review team. In order to ensure an objective assessment of all the included records, the judgements about eligibility, bias and applicability were entirely based on the published documents and not on unpublished background information. In addition, two of the three review authors who made the judgements about bias and applicability (JVG and KV) had not been involved in any of the original studies. In case of discrepant judgements, KV took the final decision.

Statistical analysis and data synthesis

The aim of the analysis was to estimate the diagnostic accuracy, that is, the sensitivity (Se) and specificity (Sp), of the index tests (RDTs for VL). We analysed each RDT separately and did not plan to make comparisons between different index tests.

The analysis was performed using R ([R Development Team 2010](#)) and WinBugs ([Spiegelhalter 2003](#)) because of these programs’ flexibility in model fitting, in line with recommendations in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* - Chapter 10 Analysing and Presenting Results ([Macaskill 2010](#)). More technical detail on the statistical methods can be found in [Appendix 1](#) and [Menten 2013](#).

1) Outcome data

The statistical analysis was based on summary data as provided in the source publications: either the number of true and false positive or negative test results compared to the reference test, or the Sensitivity (Se) and Specificity (Sp) of the index tests and 95% CI or credible intervals obtained from latent class analysis.

For those studies that reported data using several reference standards ([Table 1](#)), we extracted and summarized all 2 x 2 tables or Se and Sp estimates. The main analysis used a single 2 x 2 table or set of Se and Sp estimates for each study. The primary data for each study was selected based on a predefined ranking of reference standards: (1) latent class analysis, (2) parasitology and serology, (3) parasitology including spleen aspirate without serology, and (4) parasitology not including spleen aspirate without serology. The influence of the possible selection of alternative data sets was explored in sensitivity analyses.

2) Basic data presentation

The basic study results were summarized using coupled forest plots, presenting the crude Se and Sp estimates and 95% CIs as presented in the original publications. A summary receiver operating characteristic (ROC) plot was presented to assess the need for a ROC-based analysis rather than the bivariate logistic normal model, planned and described below.

3) Basic statistical model

The primary analysis was performed using the bivariate model with a hierarchical approach ([Reitsma 2005](#)). The bivariate, rather than a ROC-based, approach was chosen based on the prior experience of the author team in a VL diagnostic meta-analysis ([Chappuis 2006](#)), where no threshold effects were observed. In addition, the rapid tests report the results as positive or negative, which implies a common cut-off point for test positivity and which indicates, on prior grounds, that the bivariate model is the most appropriate. We determined the most appropriate link function for the bivariate model using the deviance information criterion (DIC, [Verde 2010](#)) and graphical assessment of model fit. Based on this analysis, we selected the complementary log-log (cloglog) function (see [Appendix 1](#)) as providing the best and most parsimonious fit to the data. All results are presented on the probability scale allowing

direct comparison with results obtained using the logit link function.

For the studies estimating Se and Sp compared to a reference standard, at the lower level the cell counts in the 2 x 2 tables extracted from each study were modelled using binomial distributions (Macaskill 2010).

For the studies estimating Se and Sp through LCA, the cloglog transforms of the Se and Sp and their standard errors were derived from the estimates and CIs in the source publications. They were then entered in the lower level of the hierarchical model using the normal distribution.

The Se and Sp, irrespective of their source, were, in the higher level of the hierarchical model, presumed to come from the same bivariate normal, that is, the model proposed corresponds to a hybrid approach of the bivariate model of Reitsma (Reitsma 2005) and standard random effects meta-analyses (Whitehead 2002). This model was estimated using Markov Chain Monte Carlo (MCMC) methods as implemented by WinBugs with uniform $[0,1]$ priors for the index test Se and Sp. The WinBugs code for the model can be found in Appendix 1.

4) Model summary

Results from the models were provided as point estimates of test Se and Sp and joint 95% CI and prediction regions. The confidence regions and prediction regions were plotted on the Se versus 1-Sp diagram (that is, the ROC space), together with the crude or model-based point estimates of Se and Sp, or both, for each study. In tables, we presented marginal CIs obtained from these confidence regions.

Investigations of heterogeneity

The amount of between-study heterogeneity was graphically assessed in coupled forest plots.

To explain the between-study variability, we assessed the following possible sources of heterogeneity:

- geographical area: Indian subcontinent, eastern Africa, Latin America, Mediterranean region (note: this is related to the *Leishmania* species diversity and distribution: *L. donovani* complex in Asia and eastern Africa, and *L. infantum* in the Mediterranean region and Latin America);
- commercial brand or type of test: categories determined based on extracted data;
- type of reference standard: (a) parasitology including spleen aspirate - no serology; (b) parasitology not including spleen aspirate - no serology; (c) parasitology and serology; (d) latent class analysis;
- disease prevalence (in the sample): below versus equal to or above the median (65%) across all studies; and

- quality assessment: a first indicator of quality was based on the QUADAS-2 assessment. Here, we used an overall, study-level assessment of the risk of bias. If the risk of bias in a study was low in all four QUADAS-2 domains, we defined the overall risk of bias in that study as 'low'. If the risk of bias was high in at least one of the QUADAS-2 domains, we defined the overall risk of bias as 'high'. In all other cases, the overall risk of bias was labelled 'uncertain'. A second indicator of quality was the size of the study (below versus equal to or above median [250]), because in order to ensure reasonably precise estimates of sensitivity and specificity, investigators are expected to consider sample size issues during the planning of the study (Bachmann 2006; Peeling 2007). Larger studies may have been subjected to closer scrutiny and may have recruited a more representative sample of the clinical suspect population. Yet, the precision of Se and Sp estimates in original studies does not only depend on sample size, but also on disease prevalence (Deeks 2005).

We informally assessed the influence of these factors by presenting summary Sp and Se for subgroups of studies (categories of covariate). In a formal analysis, these sources were then individually added as fixed-effect predictors of Se and Sp in the statistical model described above. We assessed the estimates and 95% CI of the effects of each of the predictors on Se and Sp. If warranted by the results of these analyses and the amount of studies available, we planned to consider developing this model further using multiple predictors or assuming separate bivariate normal distributions in different subgroups of studies.

Sensitivity analyses

Analysis using secondary reference standards

As a sensitivity analysis to the choice of the set of estimates of a given study, we made an alternative selection of the reference standard for those studies presenting results for two reference standards (Table 1).

Analysis allowing for imperfect reference standards

The model described above estimated the Se and Sp by comparison to a reference standard presuming that this reference standard was perfect, that is, that its Se and Sp were both equal to 1. However, in VL, it can be expected that reference standards have less than perfect Se or Sp, or both.

To allow for possible imperfect reference standards, we used the following approach:

- during data extraction, each study was classified according to the type of reference standard: (a) parasitology including spleen aspirate - no serology; (b) parasitology not including spleen aspirate - no serology; (c) parasitology and serology; or (d) latent class analysis;

- for these reference standards, expert opinion for the Se and Sp was elicited from seven experts (three of the authors FC, MB and SR, and four from the WHO technical expert panel on leishmaniasis);
- the expert opinion on the reference test Se and Sp was then entered as priors, using the beta distribution, in an extension of the bivariate model described above as a Bayesian sensitivity analysis (Greenland 2006).

The bivariate model was extended using a multinomial distribution for the cell counts in the 2 x 2 tables, as in the latent class analysis of diagnostic tests (Black 2002). The Se and Sp of the index test were modelled through a bivariate normal distribution as in the basic model described above. The Se and Sp of the reference tests were assumed to be equal across studies using the same reference standard, with informative priors for the Se and Sp derived from expert opinion as described above. Prevalences were estimated for each study separately by providing a uniform]0,1[for the study-specific prevalences. Data from studies that performed LCA in the source publication were included as in the basic model, as they already allowed for uncertainty in the true disease status in the source publication. The model was fitted using MCMC methods. Care was taken in assessing the fit and identifiability of the model using posterior predictive checks (Gelman 1995). It could be expected that the model would remain identifiable through the use of informative priors on the Se and Sp of the reference test and the assumption that the Se and Sp of the index test across studies arose from the same normal bivariate distribution (Enoe 2000), but, if needed, further constraints on the model parameters could be added (for example, by constraining the Sp of some of the reference standards to 1).

The results from this model were compared with those obtained in the primary analysis, described above. We also contrasted the assumptions of our model against the assumption of perfect ascertainment underlying the primary analysis approach. The performance of the model was evaluated through simulation studies. Details are given in Appendix 1 and Menten 2013.

Assessment of structure of hierarchical model

We compared the primary results with those obtained using the standard logit link function for hierarchical model. In addition, we relaxed the assumption of normally distributed random-effects, by using the t-distribution instead of the normal distribution.

Assessment of reporting bias

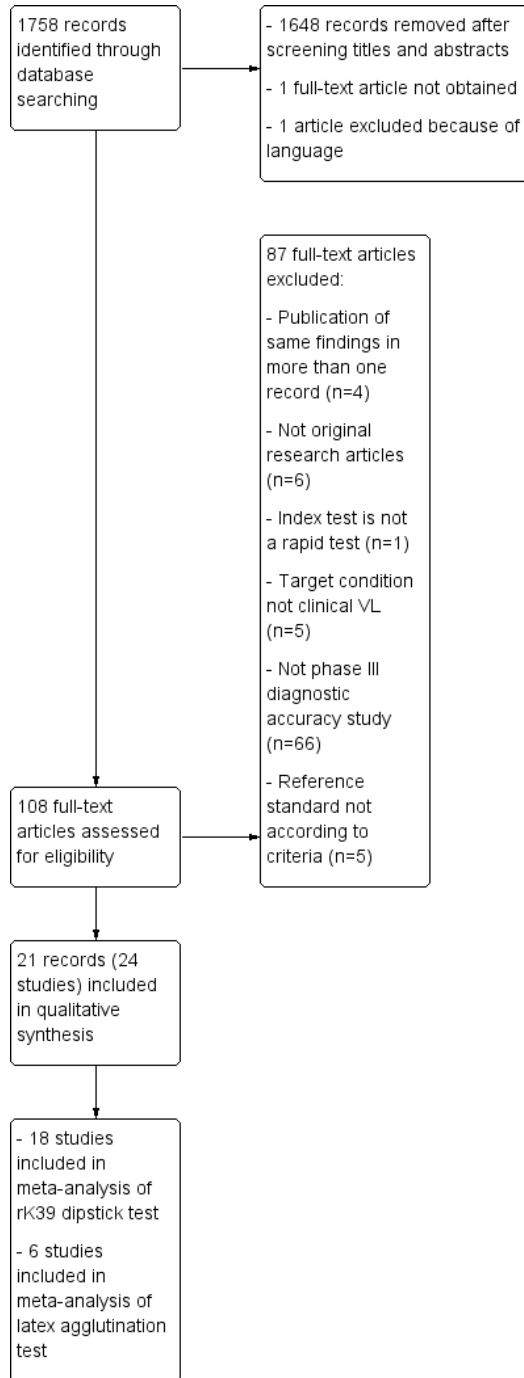
It is well known that reporting or publication bias in studies of diagnostic test accuracy may be difficult to detect and that formal tests of funnel plot asymmetry are biased (Macaskill 2010). We did not formally test for publication bias but explored the relationship between the logit Se, logit Sp, and log diagnostic odds ratio and the effective sample size (Deeks 2005) using funnel plots. However, the presence or absence of funnel plot asymmetry was not interpreted as definite proof of the presence or absence of publication bias.

RESULTS

Results of the search

The date of the search was 3 December 2013. The search identified 1758 records, each record corresponding to a published article (Figure 1). After screening titles and abstracts, 1648 irrelevant records were removed. We were unable to obtain the full-text article of one record and excluded one other article because it was written in Chinese language and no translators were available. We retrieved the full text of the remaining 108 records and excluded another 87 for the following reasons: publication of the same study in more than one record (we kept the record that was published most recently and excluded the others; n = 4); no original research (n = 6); index test not a rapid test (n = 1); target condition not clinical visceral leishmaniasis (n = 5); not a phase III diagnostic accuracy study (n = 66); and reference standard not according to criteria (n = 5; Figure 1 and Characteristics of excluded studies). The remaining 21 records were included in this review.

Figure 1. Study flow diagram showing the process of selection of records and studies for the review and for the meta-analyses



Several records reported multi-country, multicentre or stratified data. If a 'study' is defined as a phase III design leading to a sensitivity and specificity estimate of one or more RDTs in a specific patient population, then three of the 21 included records contained results of more than one unique study. One article described RDT performance in patient populations in five countries (Boelaert 2008, henceforth distinguished as [Boelaert 2008 - Ethiopia](#); [Boelaert 2008 - India](#); [Boelaert 2008 - Kenya](#); [Boelaert 2008 - Nepal](#); [Boelaert 2008 - Sudan](#)); Ter Horst et al. reported data separately for HIV-positive and HIV-negative patient populations ([ter Horst 2009 - HIV neg](#); [ter Horst 2009 - HIV pos](#)); and [Diro 2007](#) included data from clinically suspected patients presenting at the study site and from people identified through active case finding in the community. Furthermore, in one record, two different brands of the rK39 immunochromatographic test (ICT) were evaluated in the same population ([Chappuis 2005 - DiaMed](#); [Chappuis 2005 - InBios](#)), and we treated this record as two different studies. On the other hand, three patient populations were presented in more than one record. [Boelaert 2008 - Ethiopia](#) re-analysed the same patient population data as [Diro 2007](#), but using a different reference standard. Similarly, [Boelaert 2008 - India](#) described the same patient population as [Sundar 2007](#) but again, analysed it with a different reference standard. This also occurred in [Machado de Assis 2011](#) and [Machado de Assis 2012](#). We have treated these records as one single study in each country with a primary and a secondary analysis (see below). Altogether, the 21 included records contained information about 25 unique studies. We excluded the study based on active case finding (part of [Diro 2007](#)) from the review as it did not comply with the eligibility criteria of our protocol and, therefore, we finally included 24 studies. These 24 studies included 4271 participants of whom 2606 were classified as having VL. Four studies used a reference standard including direct smears or culture of splenic aspirate, or both, 11 used a composite reference standard, three used latent class analysis, and six presented two sets of accuracy estimates using two different reference standard categories ([Table 1](#)).

The 24 studies contained information about five index tests: the rK39 ICT, the latex agglutination test in urine, the FAST agglutination test, an rK26 ICT and an rKE16 ICT. Six studies assessed the accuracy of more than one type of index test in the same patient population: four studies evaluated the rK39 ICT and the latex agglutination test; one study evaluated the rK39 and the rKE16 ICT; and one study evaluated the rK39 ICT, the rK26 ICT, and the latex agglutination test ([Sundar 2007](#)). Overall, the rK39 ICT test was evaluated in 20 studies including a total of 3806 participants of whom 2370 had VL. The latex agglutination test was evaluated in seven studies, which corresponds to 1459 participants including 873 people with VL. The FAST agglutination screening test was

evaluated in two studies with a total of 148 participants including 69 with VL. The rK26 ICT test was assessed in one study with 352 participants of whom 282 had VL, and the rKE16 test was evaluated in one study with 219 participants of whom 131 had VL.

For the meta-analysis of the accuracy of the rK39 ICT test, 18 out of 20 studies were included ([Table 2](#)). At this stage, we excluded two studies because they only described HIV-positive patients ([ter Horst 2009 - HIV pos](#) and [Cota 2013](#)). In all the other studies using the rK39 ICT test, the included patients were HIV negative or the HIV status was unknown but considered of no importance in the study population. Six out of the 18 included studies generated more than one set of sensitivity and specificity estimates by using different reference standards in a primary and a secondary analysis. We selected one of the two estimates for the primary meta-analysis and explored the effects of this choice in the sensitivity analysis ([Table 1](#)).

For the meta-analysis of the accuracy of the latex agglutination test, we included six studies ([Table 2](#)). We excluded one study because it described mostly HIV-positive patients ([Vilaplana 2004](#)).

Methodological quality of included studies

The QUADAS-2 tool was used to assess the quality of the 21 included records in terms of the risk of bias and of concerns about applicability. [Figure 2](#) summarizes the overall methodological quality and [Figure 3](#) gives the ratings for each of the included records. In the domain of patient selection, we had high concerns about the risk of bias and the applicability of one record ([Veeken 2003 - composite](#)) because the inclusion of participants was conditional on the availability of stored serum samples and of diagnostic information from another serological test that could have correlated with the index test. In another record ([Cota 2013](#)), we had high concerns about the applicability to the review question because a considerable proportion of the study population (21%) had a previous history of VL. In the domains of the index test and the reference standard, the risk of bias was unclear for 14 of the 21 included records. The most frequent underlying reason was that the publication did not report whether the index test results were interpreted without knowledge of the reference standard and vice versa. We considered that there was a high risk of bias in those studies that used bone marrow aspirate smear tests and culture with or without clinical information as a reference standard due to the low sensitivity of these techniques ([Kilic 2008](#) and [Cota 2013 - composite 2](#)). In the domain of flow and timing, we considered that there was a high risk of bias in five records, because not all patients were included in the analysis ([Rijal 2004](#); [ter Horst 2009](#); [Veeken 2003](#); and [Peruhype-Magalhaes 2012](#)) or because not all patients received the same reference standard ([Sundar 1998](#)).

Figure 2. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studiesFootnoteFigure 2 represents studies but our quality assessment was done per record. Some of the records include more than one study. Figure 2 shows all the estimates used in this review, including the estimates for primary and sensitivity analyses of the same study.

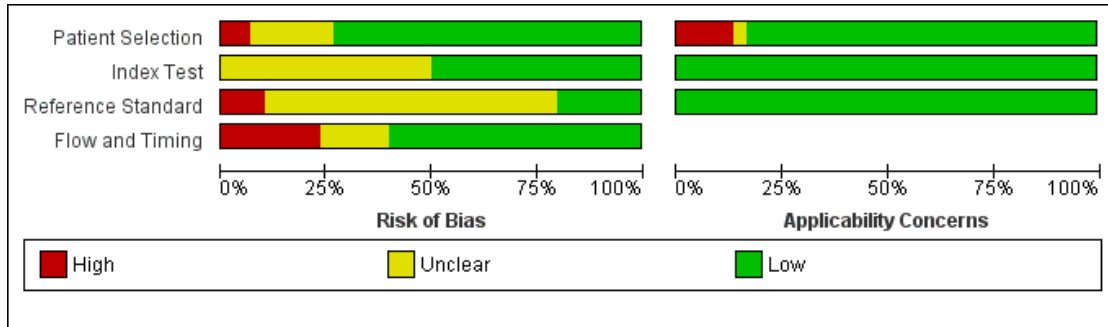


Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included studyFootnoteFigure 3 represents studies but our quality assessment was done per record. Some of the records include more than one study. Figure 3 shows all the estimates used in this review, including the estimates for primary and sensitivity analyses of the same study.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Boelaert 2004 - classic	+	?	?	+	+	+	+
Boelaert 2004 - LCA	+	?	?	+	+	+	+
Boelaert 2008 - Ethiopia	+	+	?	+	+	+	+
Boelaert 2008 - India	+	+	?	+	+	+	+
Boelaert 2008 - Kenya	+	+	?	+	+	+	+
Boelaert 2008 - Nepal	+	+	?	+	+	+	+
Boelaert 2008 - Sudan	+	+	?	+	+	+	+
Chappuis 2003	+	?	?	+	+	+	+
Chappuis 2005 - DiaMed	+	+	+	+	+	+	+
Chappuis 2005 - InBios	+	+	+	+	+	+	+
Chappuis 2006b	+	?	?	+	+	+	+
Cota 2013 - composite 1	+	+	+	+	+	+	+
Cota 2013 - composite 2	+	+	+	+	+	+	+
Diro 2007	+	?	?	+	+	+	+
Hailu 2006	+	?	?	+	+	+	+
Kilic 2008	?	?	+	?	+	+	+
Machado de Assis 2011	?	?	?	?	+	+	+
Machado de Assis 2012	?	?	?	?	+	+	+
Mbui 2013	+	+	+	+	+	+	+
Peruhype-Magalhaes 2012	?	?	?	+	+	+	+
Rijal 2004	+	+	+	+	+	+	+
Ritmeijer 2006	+	?	?	+	+	+	+
Sundar 1998	+	+	+	+	+	+	+
Sundar 2007	+	+	+	+	+	+	+
ter Horst 2009 - HIV neg	+	+	?	+	+	+	+
ter Horst 2009 - HIV pos	+	+	?	+	+	+	+
Toz 2004	?	?	?	?	+	+	+
Veeken 2003 - composite	+	?	?	+	+	+	+
Veeken 2003 - spleen	+	?	?	+	+	+	+
Vilaplana 2004	?	?	?	?	?	+	+

● High ? Unclear + Low

Findings

rK39-based rapid diagnostic tests

Data summary

Twenty studies reported Se and Sp estimates for the rK39 ICT. Six of these studies gave two sets of Se and Sp estimates, based on alternative reference standards (Table 1). A forest plot of the 26 available Se/Sp estimates is given in Figure 4. Figure 5 presents the available Se/Sp pairs according to the reference standard used; the results referring to the same data analysed with different reference standards are connected with a line.

Figure 4. rK39 ICT: forest plot of all the available estimates of sensitivity and specificity (n = 20 studies; 26 sets of estimates)FootnoteFor studies that use latent class analysis, the counts of true positive, false positive, false negative and true negative results are imputed from the Se and Sp estimates and the overall sample size. The estimates and confidence intervals are subsequently calculated from these imputed values.

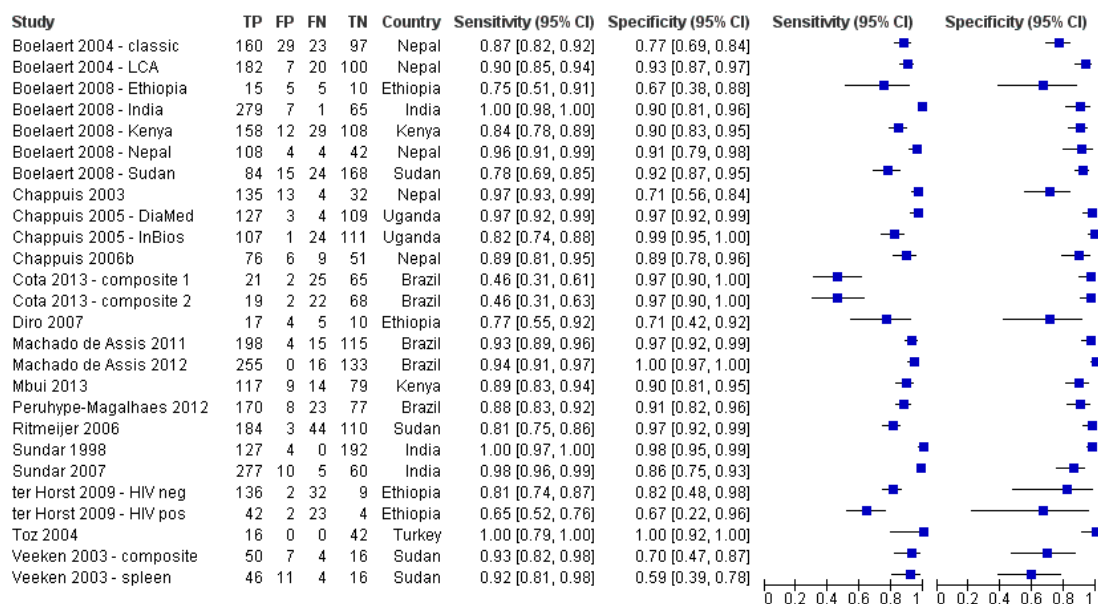
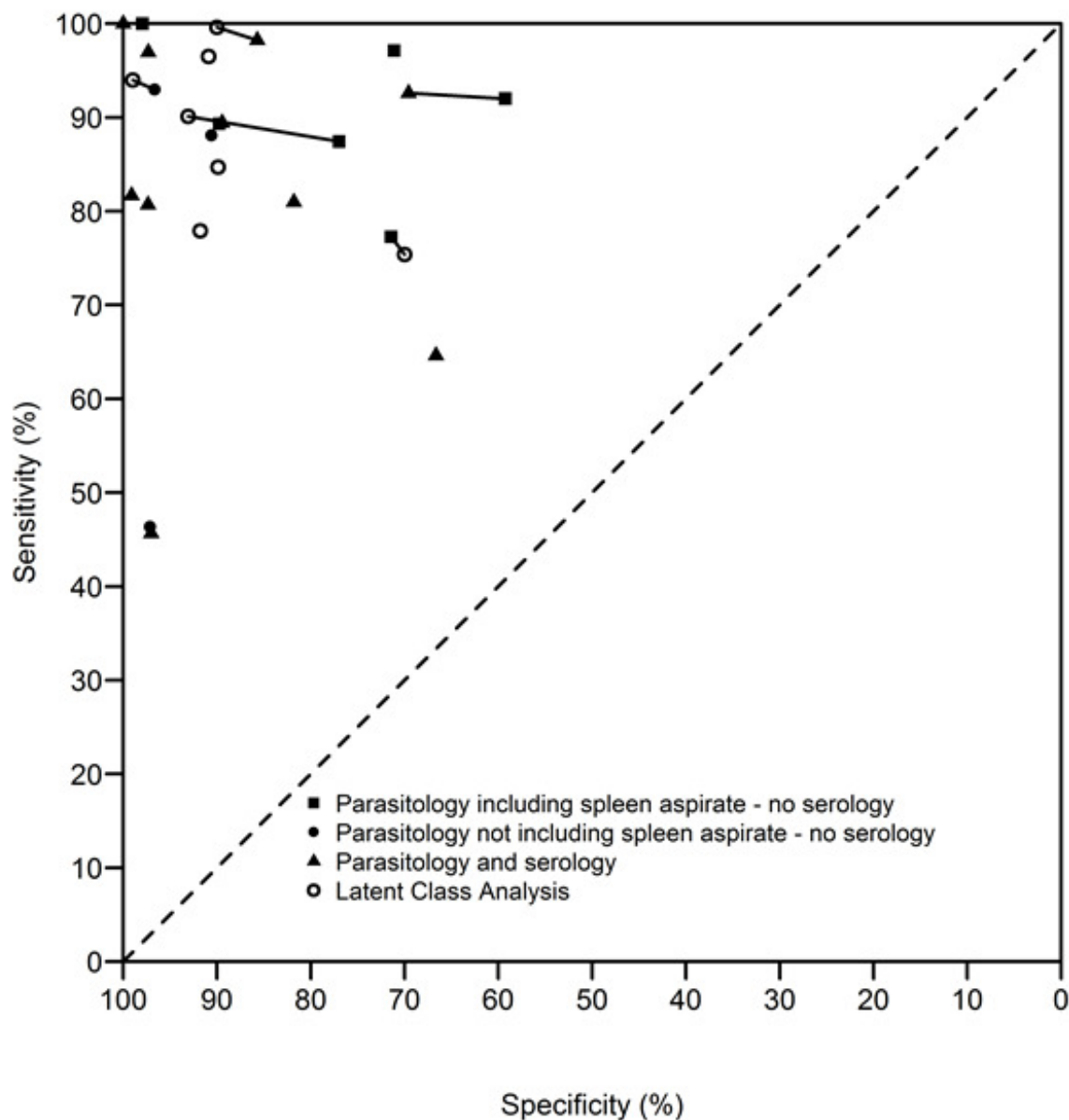


Figure 5. rK39 ICT: summary of sensitivity-specificity pairs according to the reference standard (n = 20 studies; 26 sets of estimates). The type of reference standard is classified as: (a) parasitology including spleen aspirate - no serology; (b) parasitology not including spleen aspirate - no serology; (c) parasitology and serology; or (d) latent class analysis. The data points that are connected with a line refer to the same data analysed with different reference standards.

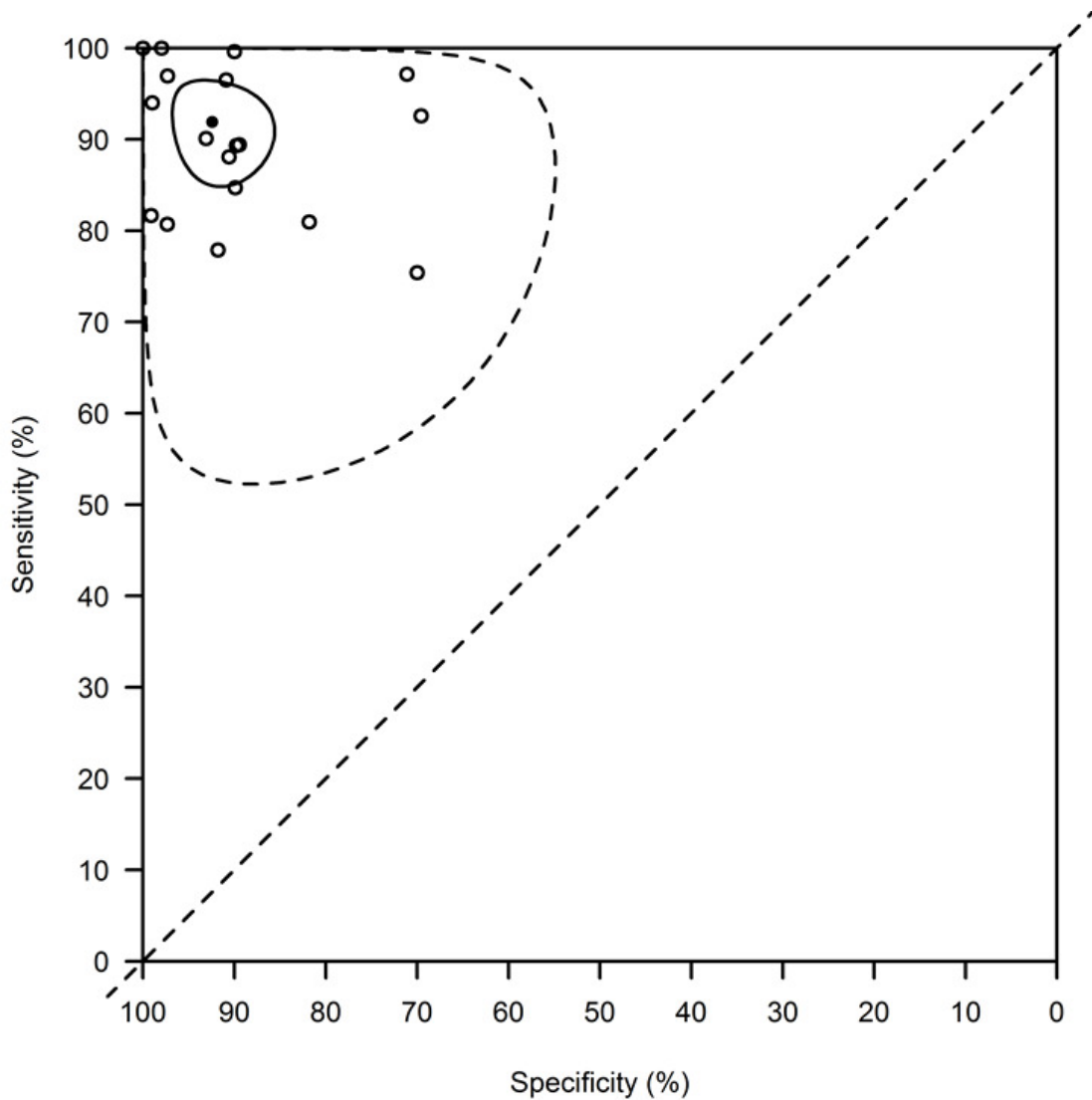


Meta-analysis

A formal meta-analysis was performed on 18 of the 26 available data points: studies including mainly HIV-infected patients were reported separately (ter Horst 2009 - HIV pos, Cota 2013 - composite 1 and Cota 2013 - composite 2) and for the five remaining studies with multiple Se and Sp estimates, we used one set of estimates for the primary analysis (Table 1). Combining data from all studies without accounting for possible covariates explaining heterogeneity and using the bivariate model (Reitsma 2005), the average (95% CI) Se was 91.9% (84.8 to 96.5) and the average Sp was 92.4% (85.6 to 96.8)

The 95% prediction interval (PI) for the diagnostic accuracy that would be observed in a new study was (52.3 to 100)% for the Se and (54.9 to 100)% for the Sp (Figure 6).

Figure 6. rK39 ICT: basic summary using the bivariate normal model with a complementary log-log link function. This analysis combines data from all studies without accounting for covariates. The average sensitivity is 91.9% (95% confidence interval: 84.8, to 96.5) and the average specificity is 92.4% (95% confidence interval: 85.6, to 96.8). The confidence region is indicated with a full line, the prediction region with a dotted line.

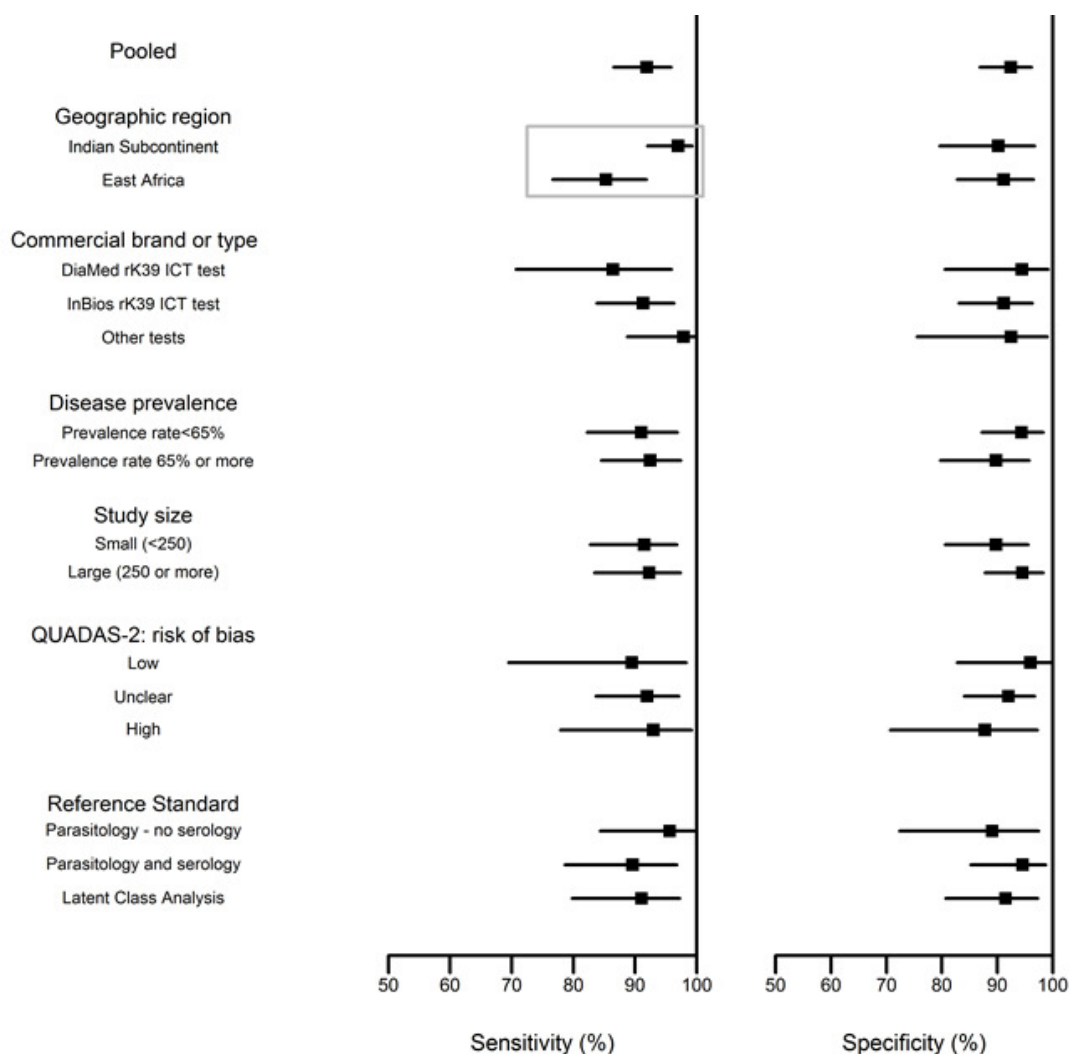


Of note is the fact that, in contrast to what is presumed in the HROC model, the estimated correlation between the transformed Se and Sp was positive: 0.16 (95% CI: -0.40 to 0.65). Given these observations, we only report results from the bivariate model using the complementary log-log, and not from the ROC-based analysis.

Assessment of heterogeneity

A summary of the heterogeneity assessment through meta-analysis can be found in [Figure 7](#) and [Table 3](#).

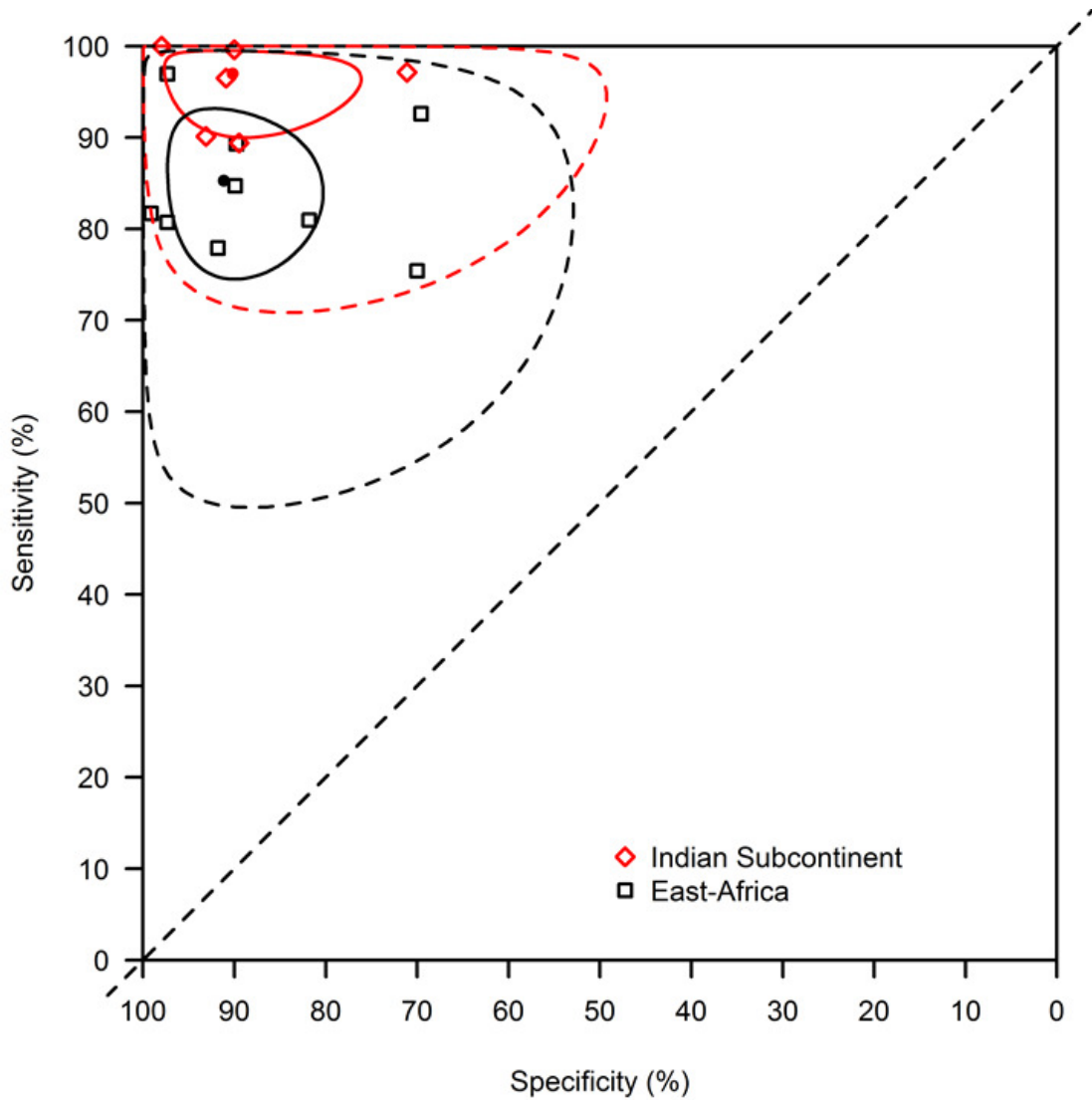
Figure 7. rK39 ICT; summary of the heterogeneity assessment through meta-analysis: sensitivity and specificity estimates by covariates. Rectangles indicate significant differences.



Geographic region

We assessed the difference in diagnostic accuracy effect between East Africa and the Indian subcontinent. There were nine data points from East Africa (two from Ethiopia, two from Kenya, three from Sudan, and two from Uganda), and six from the Indian subcontinent (two from India and four from Nepal). Three studies from other regions (two from Brazil and one from Turkey) were not considered in this analysis. The Se was significantly lower in East Africa (85.3%; 95% CI: 74.5 to 93.2) than in the Indian subcontinent (97.0%; 95% CI: 90.0 to 99.5). There was no significant difference in Sp: the Sp (95% CI) in east Africa was 91.1% (80.3 to 97.3) and in the Indian subcontinent 90.2% (76.1 to 97.7) ([Summary of findings 1](#) and [Summary of findings 2](#)). Confidence and prediction regions by geographic region are given in [Figure 8](#).

Figure 8. rK39 ICT: summary of the meta-regression model with effects for geographic region using the bivariate normal model with a complementary log-log link function. The sensitivity was significantly lower in East Africa (85.3%; 95% confidence interval: 74.5 to 93.2) than in the Indian subcontinent (97.0%; 95% confidence interval: 90.0 to 99.5). There was no significant difference in specificity: the specificity (95% confidence interval) in East Africa was 91.1% (80.3 to 97.3) and in the Indian subcontinent 90.2% (76.1 to 97.7). The confidence region is indicated with a full line, the prediction region with a dotted line.



Commercial brand or type of test

The majority of studies used the InBios rK39 ICT test (11 studies); four assessed the DiaMed rK39 ICT test; and three evaluated other tests (DiaMed DUAL-IT L/M ICT test, Amrad ICT rK39 ICT test, and Arista Biologicals rK39 ICT test). There were no significant differences in diagnostic accuracy between commercial brands.

Disease prevalence

We separated studies by estimated prevalence of VL in the study sample: nine studies reported prevalence rates < 65% and nine reported prevalence rates \geq 65%. There were no significant differences in diagnostic accuracy between studies with low and high prior probability of disease in the patient sample.

Quality assessment

We used study size and risk of bias according to QUADAS-2 as indicators of the quality of the primary studies.

Study size

We categorized studies according to the total number of study participants (true cases + non-cases) in small (< 250) and large (\geq 250) studies. There were 10 small and eight large studies. There were no significant differences in diagnostic accuracy between small and large studies, but there was a trend for larger studies to show a higher Sp estimate.

QUADAS-2: risk of bias

Studies were categorized according to the risk of bias assessed with QUADAS-2: low risk (three studies), high risk (four studies), and unclear risk of bias (11 studies). There were no significant differences among these categories, but there was a trend for studies with high risk of bias to show a low Sp and studies with low risk of bias to show a higher Sp.

Type of reference standard

Studies were categorized according to the reference standard used in the primary publication. There were no significant differences in accuracy among these categories. Further exploration of the influence of the reference standard is assessed in the sensitivity analysis below.

Diagnostic accuracy in HIV-positive participants

Two studies evaluated the rK39 ICT in HIV-positive participants and were not included in the meta-analysis.

The first study reported on the accuracy of the DiaMed rK39 ICT in 71 HIV-positive participants in Ethiopia ([ter Horst 2009 - HIV pos](#)). Based on a composite reference standard that combined direct parasitological examination of spleen aspirate with DAT serology, 65 participants had a diagnosis of VL. In this setting, the sensitivity of the rK39 ICT was 64.6% and the specificity 66.7%. The second study evaluated the InBios rK39 ICT in 113 HIV-infected people in Brazil. Data were extracted according to two different reference standards: (1) the decision of an adjudication committee after clinical follow-up and taking into account all available information (including parasitology and serology) ([Cota 2013 - composite 1](#); VL prevalence: 40.7%); and (2) direct smear and culture of bone marrow aspirate ([Cota 2013 - composite 2](#); VL prevalence: 36.9%). The choice of the reference standard was of little influence. The Se of the rK39 ICT was low: 45.7% with the first and 46.3% with the second reference standard; and the Sp was high: 97.0% with the first and 97.1% with the second reference standard.

Sensitivity analyses

We performed sensitivity analyses of the influence of the choice of reference standard and possible biases resulting from imperfect reference standards. Given the impact of geographic region on the diagnostic accuracy of rK39 ICT, we corrected the sensitivity analyses for region effects and limited the analyses to 15 studies performed in the Indian subcontinent and East Africa. An overview of the results of the sensitivity analyses is given in [Table 4](#).

Analysis using secondary reference standards

As a sensitivity analysis to the choice of the set of estimates of a given study, we made an alternative selection of the reference standard for those studies presenting results for two reference standards ([Table 1](#)). The estimated Sp was slightly lower compared to the main analysis set: 86.7% versus 90.2% for the Indian subcontinent and 90.1% versus 91.1% for east Africa. This may be related to a lower Se of the reference standards included in this analysis, which would result in an underestimation of the Sp of the index test. Estimated Se was similar in the sensitivity and main analysis.

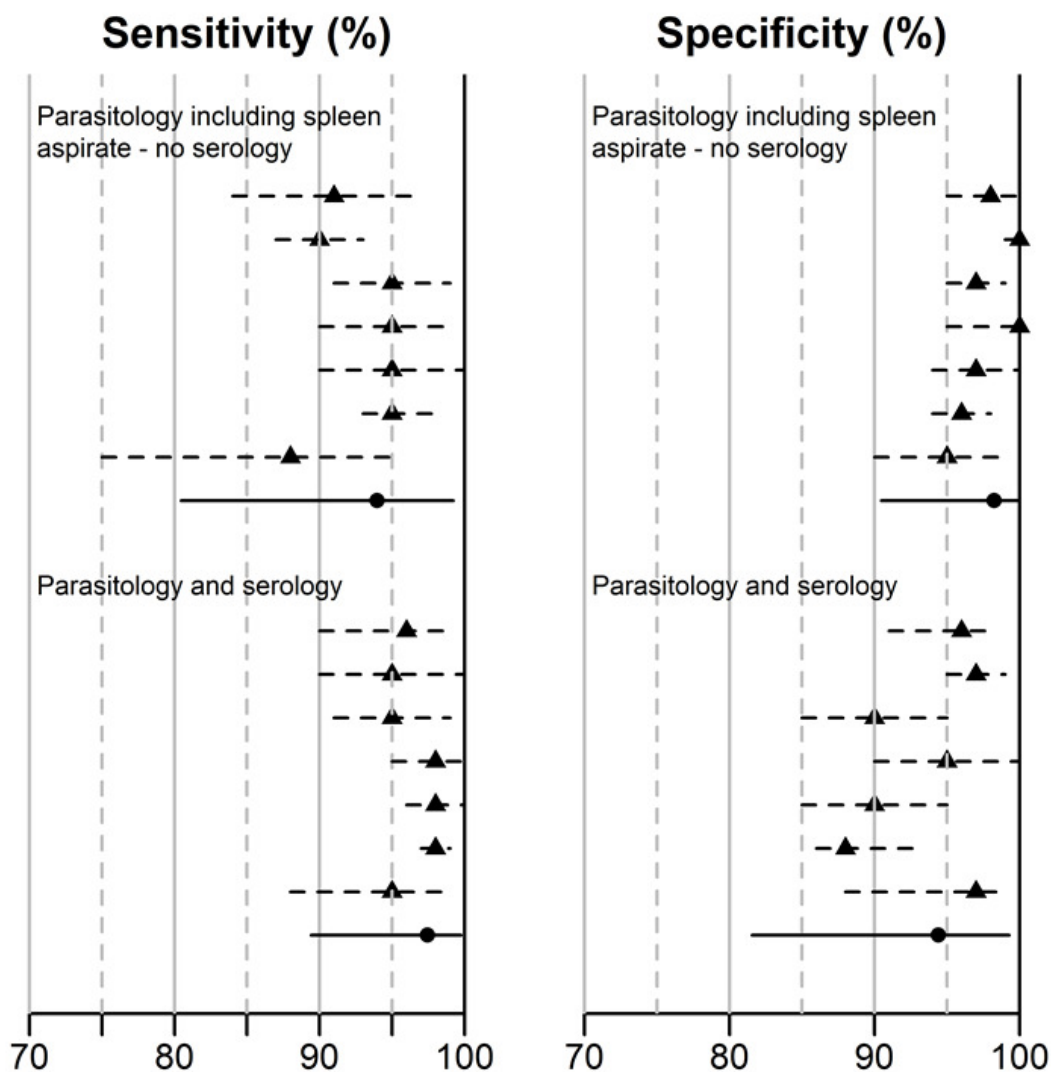
Analysis allowing for imperfect reference standards

In the studies available for this sensitivity analysis, two types of reference standards were used: parasitology including spleen aspirate - no serology; (three studies) and a combined reference standard

of parasitology and serology (six studies), in addition to the six studies analysed using latent class analysis. We elicited the opinion from seven VL experts on the diagnostic accuracy and possible variation across the two reference standards (Figure 9). We used these expert opinions as prior information in a Bayesian statistical model to estimate the diagnostic accuracy of the rK39 ICT. This analysis indicated that due to lack of sensitivity of the reference standard, the Sp of rK39 ICT may be somewhat underestimated. The estimated Sp correcting for this bias was 93.7% in the In-

dian subcontinent, and 95.7% in east Africa, compared to 90.2% and 91.1% assuming perfect reference standards. The estimates of Se did not change significantly when we corrected for imperfect reference standards. In this model, the Se of spleen aspirate was estimated to be 96.0% (95% CI: 92.5 to 98.8) and the Sp was estimated to be 99.4% (95% CI: 98.1 to 99.9). The estimated Se of the combined reference standard of spleen parasitology and a serological test was estimated to be 98.6% (95% CI: 97.1 to 99.4) and the estimated Sp was 96.5% (95% CI: 91.5 to 99.2).

Figure 9. Opinion from seven experts on visceral leishmaniasis regarding the diagnostic accuracy and possible variation of two reference standards: parasitology including spleen aspirate without serology and any combination of parasitology with serology. These expert opinions were used as prior information in a Bayesian statistical model to estimate the diagnostic accuracy of the rK39 ICT.

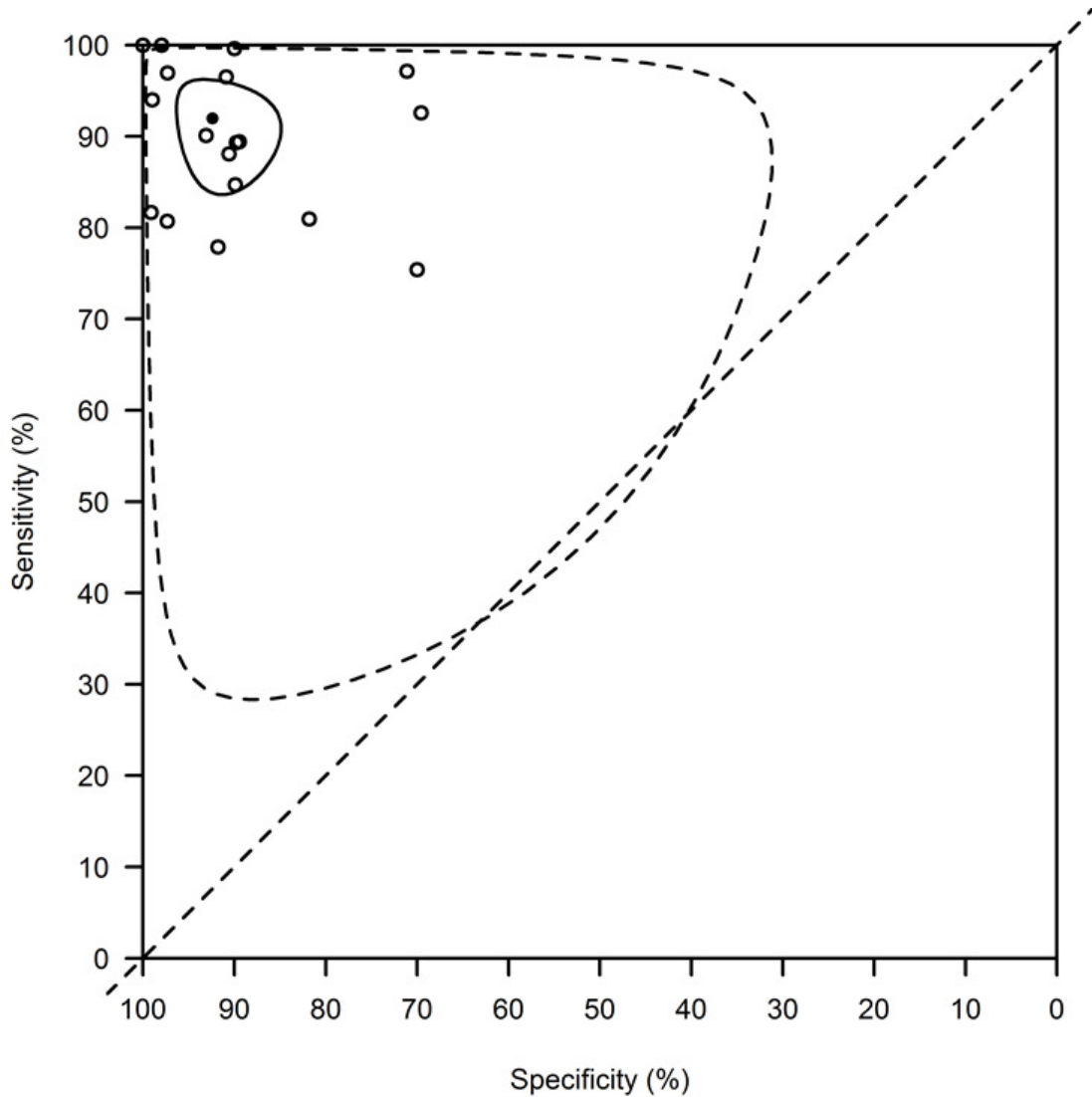


Similar results were obtained when we used non-informative priors instead of expert opinion to estimate the Bayesian model.

Assessment of structure of hierarchical model

Using the more standard logit link, in comparison to the better fitting complementary log-log link increased the DIC from 171.3 to 199.0. The logit link resulted in similar estimates of Se and Sp: the average (95% CI) Se was 91.9% (83.6 to 96.2) and the average Sp was 92.4% (84.9 to 96.3). However, the remaining heterogeneity was considerably larger resulting in wider prediction regions: 28.3 to 99.7% for Se and 31.1 to 99.7% for Sp (Figure 10) compared to the primary analysis.

Figure 10. rK39 ICT: basic model summary using the bivariate logistic normal model. This analysis combines data from all studies without accounting for covariates. The average sensitivity is 91.9% (95% confidence interval: 83.6, to 96.2) and the average specificity is 92.4% (95% confidence interval: 84.9, to 96.3). The confidence region is indicated with a full line, the prediction region with a dotted line. In this analysis, the prediction intervals are wide: 28.3 to 99.7% for the sensitivity and 31.1 to 99.7% for the specificity.

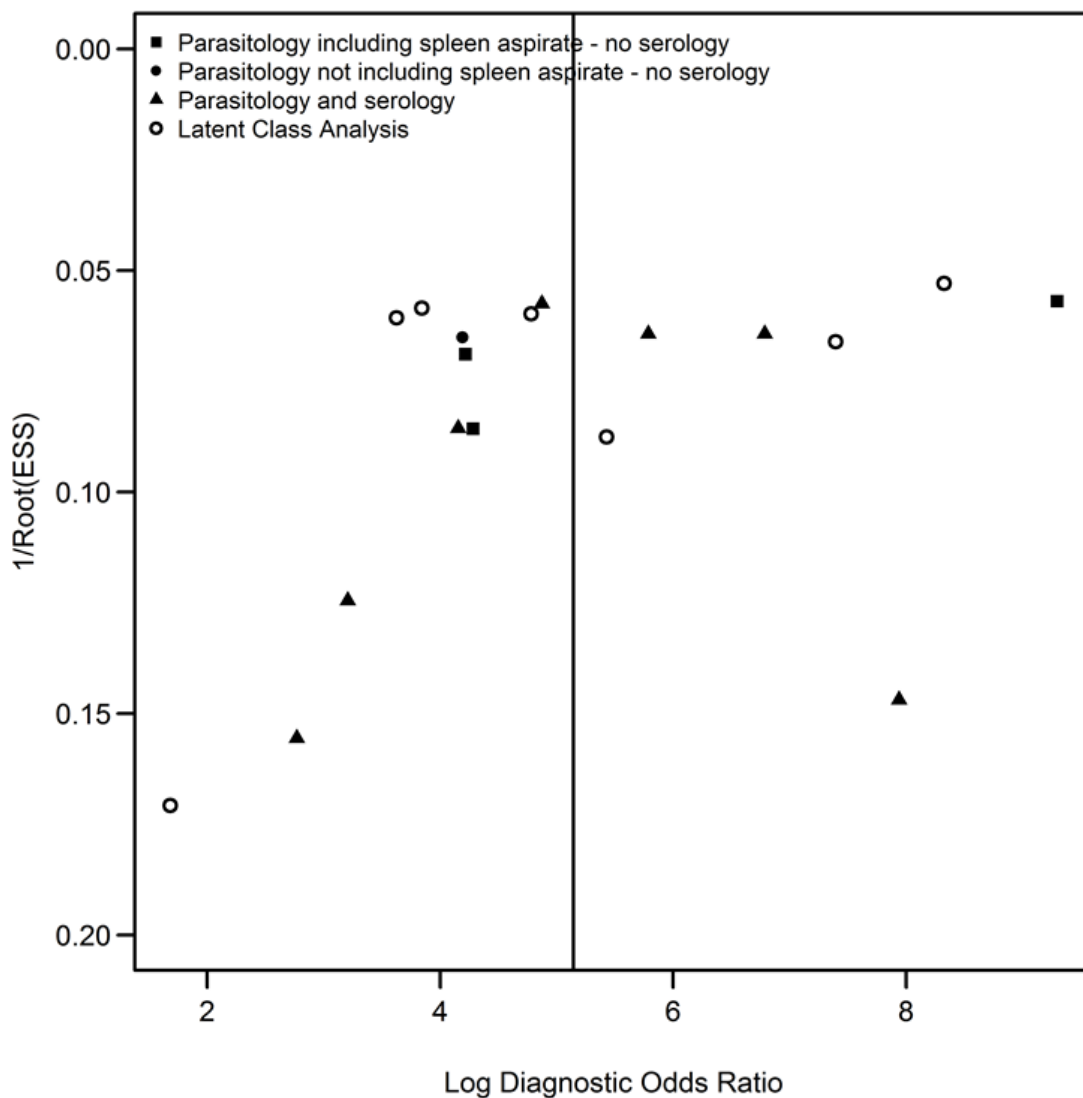


Using a bivariate t-distribution for the random-effects did not improve model fit compared to the normal model.

Assessment of reporting bias

Reporting bias was assessed through the robust funnel plot proposed by Deeks 2005. This robust funnel plot presents the log of the diagnostic odds ratio by a robust measure of the study size. No indication of reporting bias was observed in this plot (Figure 11).

Figure 11. rK39 ICT; assessment of reporting bias: funnel plot for log diagnostic odds ratio

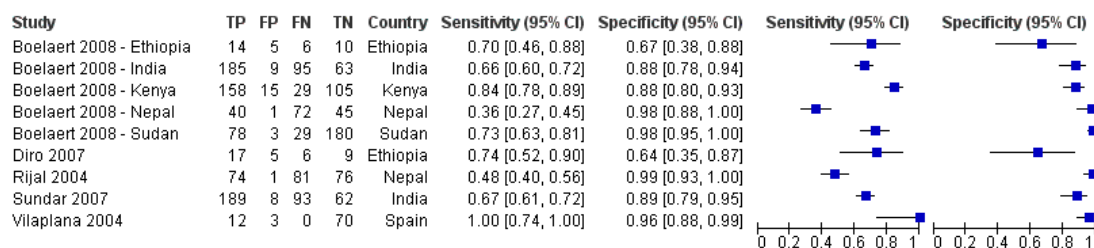


Latex agglutination test

Data summary

Seven studies were identified that reported Se or Sp results of the latex agglutination test in urine (KAtex). For two of these studies, two sets of Se and Sp estimates were available. A forest plot of the nine available Se/Sp estimates is given in [Figure 12](#).

Figure 12. Latex agglutination test: forest plot of all the available estimates of sensitivity and specificity (n = 7 studies; 9 sets of estimates).



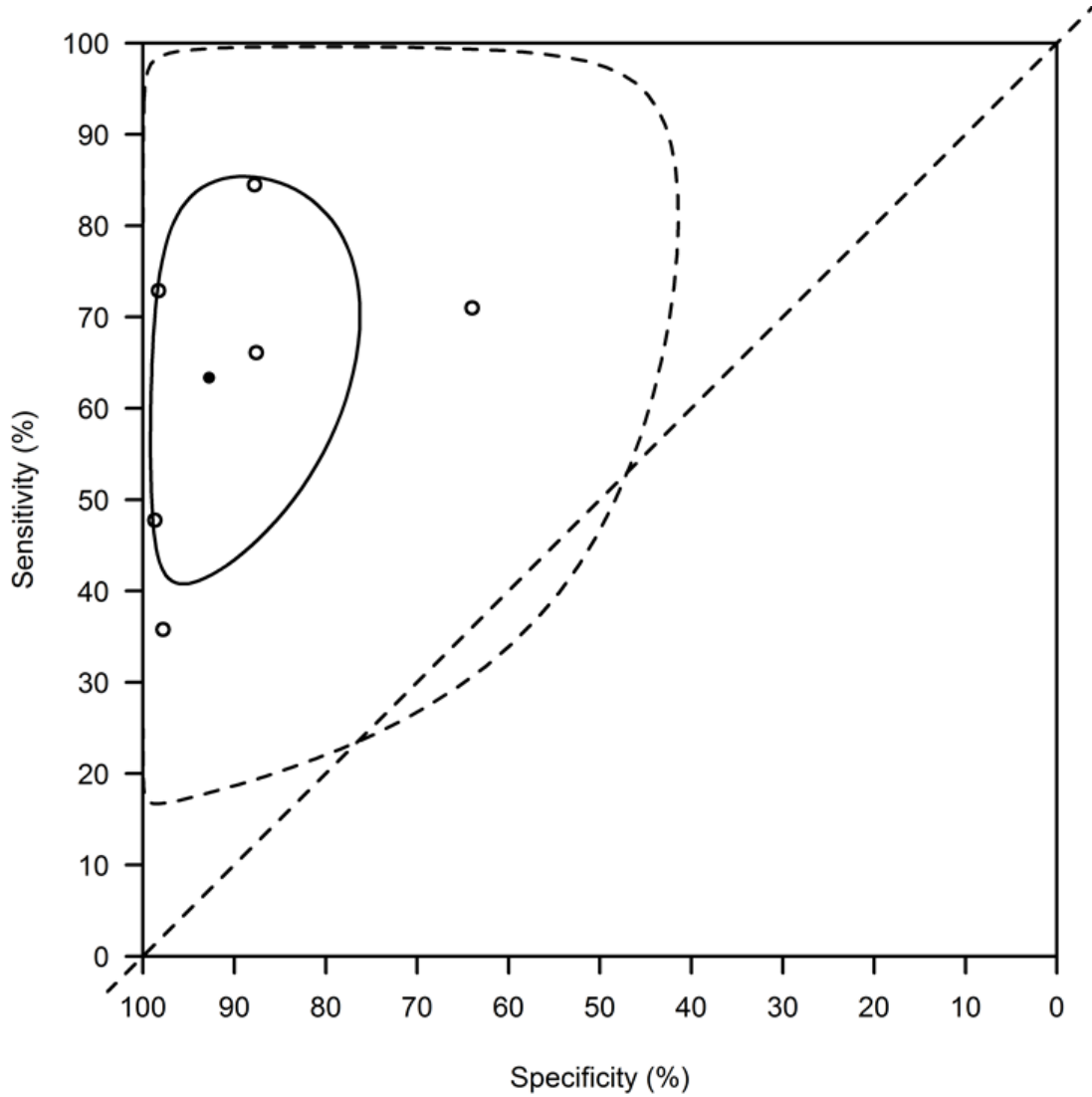
Meta-analysis

One study was not pooled with the others in the formal meta-analysis because, contrary to the other studies, this study included a majority of HIV-positive participants ([Vilaplana 2004](#)). For the two studies that presented Se/Sp estimates using LCA and a reference standard on the same data ([Boelaert 2008 - Ethiopia](#) and [Diro 2007](#), [Boelaert 2008 - India](#) and [Sundar 2007](#)), the results from the two analysis approaches were very similar. We used the

results from LCA for the primary meta-analysis.

Combining data from six studies without accounting for possible covariates explaining heterogeneity, and using the bivariate normal model with complementary log-log link, the average (95% CI) Se was 63.6% (40.9 to 85.6) and the average Sp was 92.9% (76.7 to 99.2) ([Figure 13](#) and [Summary of findings 3](#)). The 95% prediction intervals were very wide: (16.5 to 99.6)% for the Se and (41.3 to 100)% for the Sp ([Figure 13](#)). The estimated correlation between the transformed Se and Sp was -0.39 (95% CI: -0.90 to 0.70).

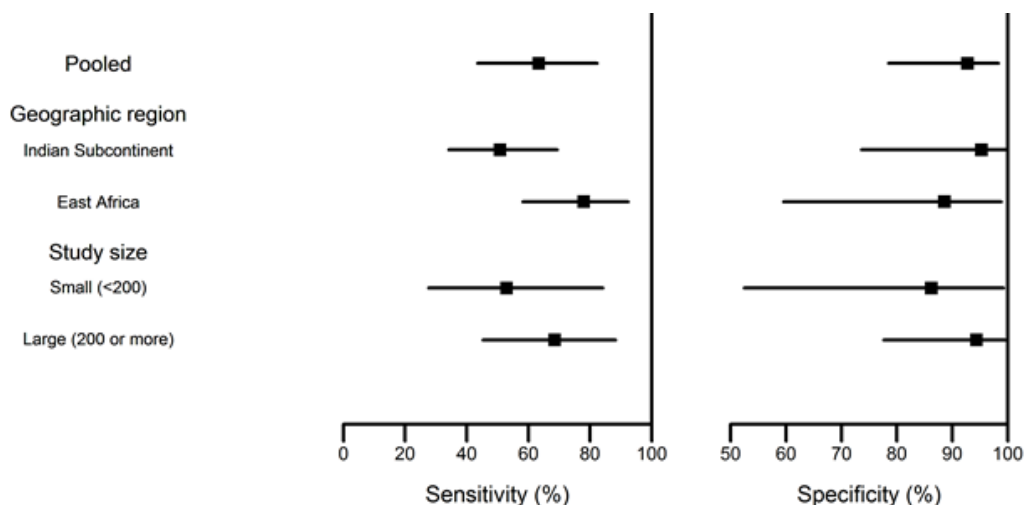
Figure 13. Latex agglutination test: basic summary using the bivariate normal model with a complementary log-log link function. This analysis combines data from all studies without accounting for covariates. The average sensitivity (95% confidence interval) is 63.6% (40.9 to 85.6) and the average specificity is 92.9% (76.7 to 99.2). The confidence region is indicated with a full line, the prediction region with a dotted line. The 95% prediction intervals are very wide: 16.5% to 99.6% for the sensitivity and 41.3% to 100% for the specificity.



Assessment of heterogeneity

Of the predefined sources of heterogeneity, geographic region, disease prevalence, and study size showed sufficient variability across studies to warrant a meta-regression. All studies used the KAtex latex agglutination test manufactured by Kalon biological. Five of the six studies had an unclear risk of bias. All studies in east Africa reported prevalences of VL < 65% and all studies in the Indian subcontinent reported prevalences \geq 65%. Consequently, the effects of region and disease prevalence in the sample could not be separated. There were no significant differences in diagnostic accuracy of the latex agglutination test among regions or study sizes (Figure 14 and Table 5).

Figure 14. Latex agglutination test; summary of the heterogeneity assessment through meta-analysis: sensitivity and specificity estimates by covariates.



Sensitivity analysis

As a sensitivity analysis to the choice of the set of estimates of a given study, we made an alternative selection of the reference standard for those studies presenting results for two reference standards. The results were very similar to the primary analysis: the average (95% CI) Se was 63.4% (40.8 to 85.4) and the average Sp was 92.8% (76.3 to 99.2) (Table 6).

One study assessed the accuracy of the latex agglutination test in a population of 85 mostly HIV-positive participants (93%) in Spain (Vilaplana 2004). According to a reference standard based on culture or direct parasitological examination of bone marrow aspirate, 12 participants were classified as having VL. The estimated Se and Sp of the KAtex test were very high: the Se was 100% (12/12; 95% CI: 74 to 100) and the Sp was 96% (70/73; 95% CI: 88 to 99).

Diagnostic accuracy in HIV-positive participants

FAST

Two studies assessed the accuracy of the Fast agglutination screening test (FAST), one in Ethiopia (Hailu 2006) and one in Turkey (Kilic 2008). Both studies were small (89 participants in Ethiopia and 59 in Turkey). The prevalence of VL in the study population was 51% in Ethiopia and 41% in Turkey. The reference standard was the culture or direct parasitological examination of spleen or lymph node samples (Ethiopia) or of bone marrow samples (Turkey). In Ethiopia, the estimated Se of FAST was 91.1% and the Sp 70.5%. In Turkey, the estimated Se was 95.8% and the Sp 85.7%.

rK26-based rapid diagnostic test

One study assessed the accuracy of an rK26-based rapid diagnostic test manufactured by InBios International (Seattle, Washington,

USA) in a population of 352 patients with suspected VL in India (Sundar 2007). According to a reference standard combining direct parasitological examination of spleen aspirate with DAT serology and response to VL treatment, 282 participants had a diagnosis of VL. The sensitivity of the rK26 ICT was 21.3% and the specificity 100%.

rKE16 immunochromatographic test

One study evaluated an rKE16-based ICT (Signal KA, Span Diagnostics, India) in Kenya (Mbui 2013). This study included 219 patients suspected to have VL. Based on a reference standard consisting of direct smear examination of a splenic aspirate sample, 131 participants were classified as having VL. The Se of the rKE16 ICT was 77.1% and the Sp 95.5%.

Summary of findings

Population: Patients suspected to have visceral leishmaniasis disease ¹

Setting: Health services in endemic areas of the Indian subcontinent ²

New test: rK39 immunochromatographic test ³

Reference standard: (1) direct smear test or culture of splenic aspirate; (2) composite reference standard based on one or more of the following: parasitology, serology, or response to treatment; or (3) latent class analysis ⁴

Pooled sensitivity: 0.97 (95% CI 0.90 to 1.00) | **Pooled specificity:** 0.90 (95% CI 0.76 to 0.98)

Setting ⁵	Positive predictive value ⁶	Negative predictive value ⁶	Number of participants (studies)	Quality of the evidence (QUADAS-2) ⁷
Peripheral health centre with a prior probability of disease of 40%	87%	98%	1468 (6 studies)	Risk of bias: none of the studies had a low risk of bias in all domains. One study had a high risk of bias (domain of flow and timing). Five studies had an unclear risk of bias (domains of index test or reference standard)
Referral centre with a prior probability of disease of 60%	94%	95%		Applicability: low concerns in all studies and in all domains

Interpretation: When the rK39 ICT is used ¹ in the Indian subcontinent, in a setting where the prior probability of VL among clinical suspects is 40%, which is typically seen in a peripheral health centre in an endemic area, the positive predictive value of the test is 87%. This means that out of 100 patients with a positive rK39 result, 87 would have VL (true positive result) and 13 would have another disease (false positive). The negative predictive value is 98%, meaning that out of 100 patients with a negative rK39 ICT result, 98 would have another disease (true negative) and 2 would have VL (false negative)

When the same test is used in a setting with a prior probability of VL of 60%, which is more typical for a referral centre in an endemic area, the positive predictive value is 94% and the negative predictive value is 95%

A likelihood ratio is another way of expressing how informative a diagnostic test is: it indicates to what extent the rK39 ICT result changes the odds that a patient has VL. The likelihood ratio of a positive rK39 ICT result is 9.90, and the likelihood ratio of a negative test result is 0.03. This means that in the Indian subcontinent, a positive rK39 ICT result is a strong argument in favour of VL (ruling in) and that a negative rK39 ICT result is a strong argument against VL (ruling out)

CI: confidence interval

Boelaert M, Verdonck K, Menten J, Sunyoto T, van Griensven J, Chappuis F, Rijal S. Rapid diagnostic tests for visceral leishmaniasis. Cochrane Database of Systematic Reviews 2011, Issue 6. Art. No.: CD009135. DOI: 10.1002/14651858.CD009135

1. The rK39 immunochromatographic test must be used in combination with a clinical case definition (fever and splenomegaly for more than two weeks and no previous history of visceral leishmaniasis). Studies with mainly HIV-positive patients were not included in the pooled analyses.
2. The results of the meta-analysis showed considerable heterogeneity, which was partly explained by the geographic region.
3. This rapid diagnostic test has been developed specifically for field use. It is less invasive, less time-consuming, and easier to perform than the alternative parasitological or serological tests.
4. Latent class analysis is a modelling technique that allows us to estimate the sensitivity and specificity of a set of diagnostic tests in situations in which there is no good reference standard.

5. Two hypothetical situations: a peripheral health centre and a referral centre with a different prior probability of disease
6. A narrative explanation of the predictive values is given in [Appendix 3](#).
7. QUADAS-2 is a tool for the assessment of the quality of diagnostic accuracy studies. The tool comprises four domains: patient selection, index test, reference standard, and flow and timing. Each domain is assessed in terms of risk of bias, and the first three domains are also assessed in terms of concerns regarding applicability.

Population: Patients suspected to have visceral leishmaniasis disease ¹					
Setting: Health services in endemic areas of east Africa ²					
New test: rK39 immunochromatographic test ³					
Reference standard: (1) direct smear test or culture of splenic aspirate; (2) composite reference standard based on one or more of the following: parasitology, serology, or response to treatment; or (3) latent class analysis ⁴					
Pooled sensitivity: 0.85 (95% CI 0.75 to 0.93) ; Pooled specificity : 0.91 (95% CI 0.80 to 0.97)					
Setting ⁵	Positive predictive value ⁶	Negative predictive value ⁶	Number of participants (studies)	Quality of the evidence (QUADAS-2) ⁷	
Peripheral health centre with a prior probability of disease of 40%	86%	90%	1692 (9 studies)	Risk of bias: three studies had a low risk of bias across all domains; four studies had an unclear risk of bias (domains of index test or reference standard); two studies had a high risk of bias (domains of patient selection or flow and timing, or both) Applicability: one study with high concerns about applicability (domain of patient selection); low concerns for all other studies and all other domains	
Referral centre with a prior probability of disease of 60%	93%	81%			
Interpretation: When the rK39 ICT is used ¹ in east Africa, in a setting where the prior probability of VL is 40%, which is typically seen in a peripheral health centre in an endemic area, the positive predictive value of the test is 86%. This means that out of 100 patients with a positive rK39 ICT result, 86 would have VL (true positive result) and 14 would have another disease (false positive). The negative predictive value is 90%, meaning that out of 100 patients with a negative rK39 ICT result, 90 would have another disease (true negative) and 10 would have VL (false negative) When the same test is used in a setting with a prior probability of VL of 60%, which is more typical for a referral centre in an endemic area, the positive predictive value is 93% and the negative predictive value is 81% In east Africa, the likelihood ratio of a positive rK39 ICT result is 9.58, and the likelihood of a negative rK39 ICT result is 0.16. This means that a positive rK39 ICT result is strong argument in favour of VL (ruling in), and that a negative rK39 ICT result is not an absolute argument against VL (does not allow to rule out VL completely)					
CI: confidence interval					

1. The rK39 immunochromatographic test must be used in combination with a clinical case definition (fever and splenomegaly for more than two weeks and no previous history of visceral leishmaniasis). Studies with mainly HIV-positive patients were not included in the pooled analyses.
2. The results of the meta-analysis showed considerable heterogeneity, which was partly explained by the geographic region.
3. This rapid diagnostic test has been developed specifically for field use. It is less invasive, less time-consuming, and easier to perform than the alternative parasitological or serological tests.
4. Latent class analysis is a modelling technique that allows to estimate the sensitivity and specificity of a set of diagnostic tests in situations in which there is no good reference standard.
5. Two hypothetical situations: a peripheral health centre and a referral centre with a different prior probability of disease
6. A narrative explanation of the predictive values is given in [Appendix 3](#).
7. QUADAS-2 is a tool for the assessment of the quality of diagnostic accuracy studies. The tool comprises four domains: patient selection, index test, reference standard, and flow and timing. Each domain is assessed in terms of risk of bias, and the first three domains are also assessed in terms of concerns regarding applicability.

Population: Patients suspected to have visceral leishmaniasis disease ¹

Setting: Health services in endemic areas ²

New test: Latex agglutination test in urine ³

Reference standard: (1) direct smear test or culture of splenic aspirate; (2) composite reference standard based on one or more of the following: parasitology, serology, or response to treatment; or (3) latent class analysis ⁴

Pooled sensitivity : 0.64 (95% CI 0.41 to 0.86) ; **Pooled specificity :** 0.93 (95% CI 0.77 to 0.99) ⁵

Setting ⁶	Positive predictive value	Negative value	predictive	Number of participants (studies)	Quality of the evidence (QUADAS-2) ⁷
Peripheral health centre with a prior probability of disease of 40%	86%	79%		1374 (6 studies)	Risk of bias: none of the studies had a low risk of bias in all domains. One study had a high risk of bias (domain of flow and timing). Five studies had an unclear risk of bias (domain of reference standard)
Referral centre with a prior probability of disease of 60%	93%	63%			Applicability: low concerns in all studies and in all domains

CI: confidence interval

1. Studies with mainly HIV-positive patients were not included in the pooled analyses.

2. The studies included in this review were conducted in Ethiopia, Kenya, Sudan, India and Nepal.
3. This rapid diagnostic test has been developed specifically for field use. It is less invasive, less time-consuming, and easier to perform than the alternative parasitological or serological tests.
4. Latent class analysis is a modelling technique that allows to estimate the sensitivity and specificity of a set of diagnostic tests in situations in which there is no good reference standard.
5. The results of the meta-analysis showed considerable heterogeneity.
6. Two hypothetical situations: a peripheral health centre and a referral centre with a different prior probability of disease.
7. QUADAS-2 is a tool for the assessment of the quality of diagnostic accuracy studies. The tool comprises four domains: patient selection, index test, reference standard, and flow and timing. Each domain is assessed in terms of risk of bias, and the first three domains are also assessed in terms of concerns regarding applicability.

DISCUSSION

Summary of main results

The rK39 ICT, developed in 1998, is the rapid diagnostic test for VL that has been most thoroughly evaluated so far. We retrieved 21 publications that corresponded to our inclusion criteria for the review, and from these, ultimately 18 independent sets of sensitivity (Se) and specificity (Sp) estimates could be included in a formal meta-analysis of diagnostic accuracy of the rK39 ICT. In patients with febrile splenomegaly and no previous history of VL, the overall sensitivity of the rK39 ICT was 91.9% (95% CI 84.8 to 96.5) and the specificity 92.4% (95% CI 85.6 to 96.8). Sensitivity was significantly lower in East Africa (84.3%; 95% CI 74.5 to 93.2) than in the Indian subcontinent (97.0%; 95% CI 90.0 to 99.5), but there was no significant difference in specificity between geographic regions (Summary of findings 1, Summary of findings 2, and Appendix 3). This assessment of heterogeneity did not include the Latin American and Mediterranean region as the number of studies from those regions was too low. There was no significant difference according to prevalence of disease in the sample or manufacturer of the rK39 ICT, but the number of estimates in some categories of the latter was low (11/18 InBios, 4/18 DiaMed, and 3/18 other brands). Generally, care should be taken when interpreting this heterogeneity assessment due to the limited number of studies included in the meta-analysis. Apart from the lack of power, several risk factors may be correlated inducing confounding between different covariates.

For the KAtex test, a latex agglutination test in urine, when used in a clinical setting in patients with febrile splenomegaly, overall sensitivity was 63.6% (95% CI 40.9 to 85.6) and specificity 92.9% (95% CI 76.7 to 99.2) (Summary of findings 3). Seven studies were included in the review and six in the meta-analysis. Because of the limited number of estimates, no complete heterogeneity assessment could be made. There were no significant differences in diagnostic accuracy of the KAtex test among regions or study sizes, but the number of included studies was too small to allow for definite conclusions.

The number of studies addressing three other rapid diagnostic tests (FAST, rK26 ICT, and rKE16 ICT) was too low to allow for a meta-analysis. Our conclusions do not apply to HIV-positive patients. In summary, completeness of evidence was most satisfactory for the rK39 ICT, only partially complete for the latex agglutination test in urine, and too incomplete for the other three tests. The methodological quality of the evidence on the rK39 ICT and the latex agglutination test was acceptable, although we had concerns about the risk of bias in some studies (see below).

Strengths and weaknesses of the review

Our review protocol explicitly set out to select clinical evaluations in a phase III design, based on consecutively recruited patients all suspect for VL, as this type of evidence level is required for making policy recommendations for clinical practice. Therefore, we had to exclude many records using a phase II (case control) design. However, in some records, the study design was not clearly reported, which made it difficult to distinguish between a phase III and a case control design. Because of this poor reporting, we may have missed some true phase III studies.

We believe that the procedures used for study selection were quite exhaustive, as we did not impose any a priori language barriers (excluding only one study in Chinese ex-post), and we searched several databases including regional ones such as LILACS. Therefore, we believe that the information presented here reflects the body of evidence accumulating over the past 15 years rather well. In addition, assessment of possible publication bias did not indicate that smaller studies without favourable results for the RDTs were less likely to be reported.

Most important was the striking heterogeneity in methods across the studies. Although the included studies all complied with the criteria for an acceptable reference standard set out in our protocol, there was a remarkable variety in the reference standards that were used. These reference standards were based on different tests, done in different orders, with different cut-off values, and making different use of clinical information. We formally assessed the risk of bias in each record using the revised QUADAS-2 instrument. In five out of the 21 included records, we had clear concerns related to the patient and sample flow and the timing. For 14 records, the risk of bias was difficult to assess in the domains of the index test and the reference standard mainly because authors failed to report explicitly if and how blinding was applied. When we categorized records into low ($n = 3$), unclear ($n = 11$), and high ($n = 7$) risk of bias, there was no significant difference in Se and Sp of the rK39 ICT, but high-risk studies tended to have a higher sensitivity and a lower specificity. One possible explanation is that the reference standards in studies classified as at high risk of bias were of sub-optimal sensitivity. This could result in a selection of cases being primarily severe cases which may be easy to diagnose with rapid diagnostic tests, while the control group may contain some undiagnosed true VL cases (falsely negative in the reference standard). We tried to assess the influence of this quality in study design on the estimates of accuracy in a sensitivity analysis in the case of the rK39 ICT. Neither the use of a secondary analysis set that was based on a different reference standard, nor an analysis allowing for an imperfect reference standard affected the main conclusions of the meta-analysis on the rK39 ICT, although the secondary analysis using an alternative reference standard in five studies tended to give a lower specificity estimate. This may be related to a lower Se of the reference standards included in this analysis, which would result in an underestimation of the Sp of the index test.

Finally, the findings seemed sufficiently homogeneous to allow for a meaningful summary estimate for the specificity of rK39 ICT as

well as the Se and Sp of the latex agglutination test in urine. However, the sensitivity of rK39 ICT was different according to geographic region, and seems substantially lower in East Africa than in the Indian subcontinent. This finding is in line with an earlier meta-analysis conducted by our team (Chappuis 2006) and with a multi-country study conducted by WHO/TDR (data not included) (Cunningham 2012). Several hypotheses have been raised to explain this lower sensitivity in east Africa: a lower level of circulating antibodies, different age group, or parasitological differences (Bhattacharyya 2013; Harhay 2011).

Applicability of findings to the review question

Our review focused on clinical studies (phase III) of diagnostic accuracy, recruiting participants representative for the spectrum of patients encountered in clinical practice in whom a rapid diagnostic test for VL would be warranted. In summary, those patients would be in line with the WHO case definition and defined as “patients with a clinical suspicion of VL, that is, those who are febrile for more than two weeks and have splenomegaly, presenting at health services in endemic areas.” (Source: protocol). Patients who were previously treated for VL (non-responders or relapsed cases), and those who had signs and symptoms of other forms of leishmaniasis, such as post kala-azar dermal leishmaniasis (PKDL) were excluded. All clinical studies included in the review clearly corresponded with these inclusion criteria. Nevertheless, the settings of these clinical studies varied: some were conducted in tertiary care centres, and some in smaller hospitals of an intermediary level. Few studies recruited patients at the first-contact point, the most peripheral level of the health service, ie health centres or health posts. Although the level of health service leads to variable prior probability of disease in the study sample, and as such is of influence in diagnostic accuracy studies, we believe that our findings are applicable across the several levels of the health system as an analysis of heterogeneity according to disease prevalence did not reveal any difference in the Se or Sp estimates.

Our review does not apply to HIV-positive patients, and caution should be taken in areas with high HIV-prevalence. Because HIV-status is known to influence diagnostic accuracy (ter Horst 2009 - HIV neg; ter Horst 2009 - HIV pos; Alvar 2008), we originally intended to analyse the accuracy of the RDTs in HIV-positive and HIV-negative patients separately. Unfortunately, the number of included studies of HIV-positive people was too low (n = 3) for a separate analysis or a comparison. Furthermore, the information from the Mediterranean region and from Latin America was too limited to draw robust conclusions. Finally, it was not possible to precisely evaluate the importance of the type of sample (serum, plasma or blood; fresh or frozen urine) or the manufacturer or version of a certain test, because these parameters did not vary enough among the included studies. In particular, Cunningham et al. pointed to important variations of sensitivity in a head-to-head comparison of five brands of RDT where two tests based on

rK16 antigen performed substantially less well in east Africa and Brazil (Cunningham 2012).

In summary, we conclude that for current practice of VL control, there is one rapid diagnostic test, the rK39 ICT, which has been sufficiently evaluated, showing high sensitivity and specificity on average, but with a notably lower sensitivity in east Africa compared to the Indian subcontinent. In the Indian subcontinent, the rK39 ICT clearly complies with the normative requirements put forward before, of a sensitivity above 95% and a specificity above 90% (Boelaert 2007). In east Africa, the sensitivity of the rK39 ICT does not comply with these criteria, and can therefore not be used as a stand-alone test to reliably rule out VL in a patient who is suspected to have VL. This does not mean that the use of an rK39 ICT is precluded, but it should be embedded in a test algorithm with a clear instruction on what to do in case of a negative rK39 ICT (second test, referral, come back after two weeks for repeat testing or other).

For the latex agglutination test, although there were only six studies included in the meta-analysis, we can conclude that the low sensitivity makes it unsuitable for use in current clinical practice, though it does not preclude that its performance could be further improved in the future.

No recommendations can be made concerning the FAST test, the rK26 ICT or the rKE16 ICT because of paucity of evidence.

AUTHORS' CONCLUSIONS

Implications for practice

The rK39 ICT can be clearly recommended as a rapid diagnostic test for use in clinical care in the Indian subcontinent in patients with febrile splenomegaly and no previous history of VL.

In east Africa, the rK39 ICT can replace the DAT and other diagnostic tests as the basis for the therapeutic decision in patients suspected to have VL. However, because of the low sensitivity of the rK39 ICT, a negative test result does not rule out VL. Therefore, additional actions are needed in case of a negative result (eg second or different test, referral, coming back after two weeks for repeat testing or other). Too little evidence has accrued so far from other regions to make specific recommendations.

It is important to remember that this antibody-based test has to be used in combination with a clinical case definition (fever of more than two weeks, splenomegaly and no previous history of VL) that needs to be strictly adhered to.

The sensitivity of the KATex latex agglutination test in urine is too low to recommend it for standard practice guidelines for detection of VL in similar settings.

Implications for research

Although this review yielded solid evidence for recommending the rK39 ICT for clinical practice in the Indian subcontinent and East Africa, it would be helpful to conduct in the future more head-to-head comparisons of available tests by region and by test format, as was done by Cunningham et al. (Cunningham 2012), on a regular basis, as this would inform policy makers for their purchasing policies and quality assurance.

Several rapid diagnostic tests such as the latex agglutination test and the FAST should be further developed and evaluated before any recommendation on wide-scale use can be made. Furthermore, better tests are needed, ie tests that are more specific for the acute stage of disease and more sensitive in all geographic regions, especially east Africa.

Last but not least, phase III clinical prospective studies are an essential element in the research and development process of a diagnostic device because they provide the basis for clinical guide-

lines. More studies of this type are needed, although in the case of VL, they are more complex in terms of reference standards. In addition, further standardization of evaluation methodology and a broader awareness of the QUADAS-2 and STARD criteria by investigators would be beneficial for the quality of future evidence in this field.

ACKNOWLEDGEMENTS

We thank Anne-Marie Trooskens from the Institute of Tropical Medicine in Antwerp for expert administrative support and Vittoria Lutje, Information Specialist of the Cochrane Infectious Diseases Group for the literature searches.

The editorial base of the Cochrane Infectious Diseases Group (CIDG) is funded by UKaid from the UK Government for the benefit of low- and middle-income countries.

REFERENCES

References to studies included in this review

Boelaert 2004 - classic {published data only}

Boelaert M, Rijal S, Regmi S, Singh R, Karki B, Jacquet D, et al. A comparative study of the effectiveness of diagnostic tests for visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 2004;**70**(1):72–7.

Boelaert 2004 - LCA {published data only}

Boelaert M, Rijal S, Regmi S, Singh R, Karki B, Jacquet D, et al. A comparative study of the effectiveness of diagnostic tests for visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 2004;**70**(1):72–7.

Boelaert 2008 - Ethiopia {published data only}

Boelaert M, El-Safi S, Hailu A, Mukhtar M, Rijal S, Sundar S, 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. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2008;**102**(1): 32–40.

Boelaert 2008 - India {published data only}

Boelaert M, El-Safi S, Hailu A, Mukhtar M, Rijal S, Sundar S, 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. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2008;**102**(1): 32–40.

Boelaert 2008 - Kenya {published data only}

Boelaert M, El-Safi S, Hailu A, Mukhtar M, Rijal S, Sundar S, 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. *Transactions of the*

Royal Society of Tropical Medicine and Hygiene 2008;**102**(1): 32–40.

Boelaert 2008 - Nepal {published data only}

Boelaert M, El-Safi S, Hailu A, Mukhtar M, Rijal S, Sundar S, 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. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2008;**102**(1): 32–40.

Boelaert 2008 - Sudan {published data only}

Boelaert M, El-Safi S, Hailu A, Mukhtar M, Rijal S, Sundar S, 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. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2008;**102**(1): 32–40.

Chappuis 2003 {published data only}

Chappuis F, Rijal S, Singh R, Acharya P, Karki BM, Das ML, et al. 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. *Tropical Medicine & International Health : TM & IH* 2003; **8**(3):277–85.

Chappuis 2005 - DiaMed {published data only}

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 rapid diagnosis of visceral leishmaniasis in northeastern Uganda. *Journal of Clinical Microbiology* 2005;**43**(12): 5973–7.

- Chappuis 2005 - InBios** *{published data only}*
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 rapid diagnosis of visceral leishmaniasis in northeastern Uganda. *Journal of Clinical Microbiology* 2005;**43**(12): 5973–7.
- Chappuis 2006b** *{published data only}*
Chappuis F, Rijal S, Jha UK, Desjeux P, Karki BM, Koirala S, et al. Field validity, reproducibility and feasibility of diagnostic tests for visceral leishmaniasis in rural Nepal. *Tropical Medicine & International Health : TM & IH* 2006;**11**(1):31–40.
- Cota 2013 - composite 1** *{published data only}*
Cota GF, de Sousa MR, de Freitas Nogueira BM, Gomes LI, Oliveira E, Assis TS, et al. Comparison of parasitological, serological, and molecular tests for visceral leishmaniasis in HIV-infected patients: a cross-sectional delayed-type study. *American Journal of Tropical Medicine and Hygiene* 2013;**89**: 570–7.
- Cota 2013 - composite 2** *{published data only}*
Cota GF, de Sousa MR, de Freitas Nogueira BM, Gomes LI, Oliveira E, Assis TS, et al. Comparison of parasitological, serological, and molecular tests for visceral leishmaniasis in HIV-infected patients: a cross-sectional delayed-type study. *American Journal of Tropical Medicine and Hygiene* 2013;**89**: 570–7.
- Diro 2007** *{published data only}*
Diro E, Techane Y, Tefera T, Assefa Y, Kebede T, Genetu A, et al. Field evaluation of FD-DAT, rK39 dipstick and KATEX (urine latex agglutination) for diagnosis of visceral leishmaniasis in northwest Ethiopia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2007;**101**(9): 908–14.
- Hailu 2006** *{published data only}*
Hailu A, Schoone GJ, Diro E, Tesfaye A, Techane Y, Tefera T, et al. Field evaluation of a fast anti-Leishmania antibody detection assay in Ethiopia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2006;**100**(1):48–52.
- Kilic 2008** *{published data only}*
Kilic S, Taylan Ozkan A, Babur C, Tanir G, Schallig H. Evaluation of serological tests for the diagnosis of visceral leishmaniasis [Visseral Leishmaniasis Tansnda Serolojik Testlerin Değerlendirilmesi]. *Turkish Journal of Medical Sciences* 2008;**38**(1):13–9.
- Machado de Assis 2011** *{published data only}*
de Assis TS, Braga AS, Pedras MJ, Oliveira E, Barral A, de Siqueira IC, et al. Multi-centric prospective evaluation of rK39 rapid test and direct agglutination test for the diagnosis of visceral leishmaniasis in Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2011;**105**(2):81–5.
- Machado de Assis 2012** *{published data only}*
Machado de Assis TS, Rabello A, Werneck GL. Latent class analysis of diagnostic tests for visceral leishmaniasis in Brazil. *Tropical Medicine and International Health* 2012;**17**(10):1202–7.
- Mbui 2013** *{published data only}*
Mbui J, Wasunna M, Balasegaram M, Laussermayer A, Juma R, Njenga SN, et al. Validation of two rapid diagnostic tests for visceral leishmaniasis in Kenya. *PLoS Neglected Tropical Diseases* 2013;**7**(9):e2441.
- Peruhype-Magalhaes 2012** *{published data only}*
Peruhype-Magalhaes V, Machado-de-Assis TS, Rabello A. Use of the Kala-Azar Detect(R) and IT-LEISH(R) rapid tests for the diagnosis of visceral leishmaniasis in Brazil. *Memorias do Instituto Oswaldo Cruz.* 2012;**107**:951–2.
- Rijal 2004** *{published data only}*
Rijal S, Boelaert M, Regmi S, Karki BM, Jacquet D, Singh R, et al. Evaluation of a urinary antigen-based latex agglutination test in the diagnosis of kala-azar in eastern Nepal. *Tropical Medicine & International Health : TM & IH* 2004;**9**(6):724–9.
- Ritmeijer 2006** *{published data only}*
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. *American Journal of Tropical Medicine and Hygiene* 2006;**74**(1):76–80.
- Sundar 1998** *{published data only}*
Sundar S, Reed SG, Singh VP, Kumar PC, Murray HW. Rapid accurate field diagnosis of Indian visceral leishmaniasis. *Lancet* 1998;**351**(9102):563–5.
- Sundar 2007** *{published data only}*
Sundar S, Singh RK, Bimal SK, Gidwani K, Mishra A, Maurya R, et al. Comparative evaluation of parasitology and serological tests in the diagnosis of visceral leishmaniasis in India: a phase III diagnostic accuracy study. *Tropical Medicine & International Health : TM & IH* 2007;**12**(2): 284–9.
- ter Horst 2009 - HIV neg** *{published data only}*
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. *American Journal of Tropical Medicine and Hygiene* 2009;**80**(6):929–34.
- ter Horst 2009 - HIV pos** *{published data only}*
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. *American Journal of Tropical Medicine and Hygiene* 2009;**80**(6):929–34.
- Toz 2004** *{published data only}*
Toz SO, Chang KP, Ozbel Y, Alkan MZ. Diagnostic value of rK39 dipstick in zoonotic visceral leishmaniasis in Turkey. *Journal of Parasitology* 2004;**90**(6):1484–6. [PUBMED: 15715249]
- Veeken 2003 - composite** *{published data only}*
Veeken H, Ritmeijer K, Seaman J, Davidson R. Comparison of an rK39 dipstick rapid test with direct agglutination

test and splenic aspiration for the diagnosis of kala-azar in Sudan. *Tropical Medicine & International Health : TM & IH* 2003;**8**(2):164–7.

Veeken 2003 - spleen {published data only}

Veeken H, Ritmeijer K, Seaman J, Davidson R. Comparison of an rK39 dipstick rapid test with direct agglutination test and splenic aspiration for the diagnosis of kala-azar in Sudan. *Tropical Medicine & International Health : TM & IH* 2003;**8**(2):164–7.

Vilaplana 2004 {published data only}

Vilaplana C, Blanco S, Dominguez J, Gimenez M, Ausina V, TUrul C, et al. Noninvasive method for diagnosis of visceral leishmaniasis by a latex agglutination test for detection of antigens in urine samples. *Journal of Clinical Microbiology* 2004; Vol. 42, issue 4:1853–4.

References to studies excluded from this review

Abeijon 2013 {published data only}

Abeijon C, Campos-Neto A. Potential non-invasive urine-based antigen (protein) detection assay to diagnose active visceral leishmaniasis. *PLoS Neglected Tropical Diseases* 2013;**7**:e2161.

Ahsan 2010 {published data only}

Ahsan MM, Islam MN, Mollah AH, Hoque MA, Hossain MA, Begum Z, et al. Evaluation of latex agglutination test (KAtex) for early diagnosis of kala-azar. *Mymensingh Medical Journal* 2010;**19**(3):335–9.

Akhoundi 2010 {published data only}

Akhoundi B, Mohebbali M, Babakhan L, Edrissian GH, Eslami MB, Keshavarz H, et al. Rapid detection of human *Leishmania infantum* infection: a comparative field study using the fast agglutination screening test and the direct agglutination test. *Travel Medicine and Infectious Disease* 2010;**8**(5):305–10.

Alam 2008 {published data only}

Alam M, Shamsuzzaman A, Musa A, Khan A, Mahmud M, Hossain M, et al. Antigen detection from urine of Kala-azar cases by latex agglutination test. *Mymensingh Medical Journal* 2008;**17**(1):22–7.

Alborzi 2006 {published data only}

Alborzi A, Rasouli M, Nademi Z, Kadivar MR, Pourabbas B. Evaluation of rK39 strip test for the diagnosis of visceral leishmaniasis in infants. *Eastern Mediterranean Health Journal = La revue de sante de la Mediterranee orientale = al-Majallah al-sibhiyah li-sharq al-mutawassit* 2006;**12**(3–4): 294–9.

Al-Nahhas 2003 {published data only}

Al-Nahhas S, Shabaan M, Hammoud L, Al-Taweel A, Al-Jorf S. Visceral leishmaniasis in the Syrian Arab Republic: early detection using rK39. *Eastern Mediterranean Health Journal = La revue de sante de la Mediterranee orientale = al-Majallah al-sibhiyah li-sharq al-mutawassit* 2003;**9**(4): 856–62.

Al-Nahhas 2008 {published data only}

Al-Nahhas SA, Al-Taweel AA, Al-Taweel MA. Assessment of the direct agglutination test, fast agglutination screening

test, and rK39 dipstick test for the sero-diagnosis of visceral leishmaniasis in Syria. *Saudi Medical Journal* 2008;**29**(9): 1250–4.

Amato 2009 {published data only}

Amato Neto V, Amato VS, Tuon FF, Gakiya E, de Marchi CR, de Souza RM, et al. False-positive results of a rapid K39-based strip test and Chagas disease. *International Journal of Infectious Diseases : IJID : official publication of the International Society for Infectious Diseases* 2009;**13**(2): 182–5.

Arya 1997 {published data only}

Arya SC. Latex agglutination test for visceral leishmaniasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1997; Vol. 91, issue 3:366.

Arya 2001 {published data only}

Arya SC. Field performance of the direct agglutination test and the K39 dipstick test to diagnose visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 2001; Vol. 65, issue 5:403–4.

Attar 2001 {published data only}

Attar ZJ, Chance ML, el-Safi S, Carney J, Azazy A, El-Hadi M, et al. Latex agglutination test for the detection of urinary antigens in visceral leishmaniasis. *Acta Tropica* 2001;**78**(1): 11–6.

Azazy 2004 {published data only}

Azazy AA. Detection of circulating antigens in sera from visceral leishmaniasis patients using dot-ELISA. *Journal of the Egyptian Society of Parasitology* 2004;**34**(1):35–43. [PUBMED: 15125515]

Bagchi 1998 {published data only}

Bagchi AK, Tiwari S, Gupta S, Katiyar JC. The latex agglutination test: standardization and comparison with direct agglutination and dot-ELISA in the diagnosis of visceral leishmaniasis in India. *Annals of Tropical Medicine and Parasitology* 1998;**92**(2):159–63.

Bern 2000 {published data only}

Bern C, Jha SN, Joshi AB, Thakur GD, Bista MB. Use of the recombinant K39 dipstick test and the direct agglutination test in a setting endemic for visceral leishmaniasis in Nepal. *American Journal of Tropical Medicine and Hygiene* 2000;**63**(3–4):153–7.

Brandonisio 2002 {published data only}

Brandonisio O, Fumarola L, Maggi P, Cavaliere R, Spinelli R, Pastore G. Evaluation of a rapid immunochromatographic test for serodiagnosis of visceral leishmaniasis. *European Journal of Clinical Microbiology & Infectious Diseases : official publication of the European Society of Clinical Microbiology* 2002;**21**(6):461–4.

Canavate 2011 {published data only}

Canavate C, Herrero M, Nieto J, Cruz I, Chicharro C, Aparicio P, et al. Evaluation of two rK39 dipstick tests, direct agglutination test, and indirect fluorescent antibody test for diagnosis of visceral leishmaniasis in a new epidemic site in highland Ethiopia. *American Journal of Tropical Medicine and Hygiene* 2011;**84**(1):102–6.

- Carvalho 2003** *{published data only}*
Carvalho SF, Lemos EM, Corey R, Dietze R. Performance of recombinant K39 antigen in the diagnosis of Brazilian visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 2003;**68**(3):321–4.
- Cruz 2006** *{published data only}*
Cruz I, Chicharro C, Nieto J, Bailo B, Canavate C, Figueras MC, et al. Comparison of new diagnostic tools for management of pediatric Mediterranean visceral leishmaniasis. *Journal of Clinical Microbiology* 2006;**44**(7):2343–7.
- de Assis 2008** *{published data only}*
Machado de Assis T, da Costa Braga AS, Junqueira Pedras M, Prado Barral A, de Siqueira IC, Nery Costa CH, ET AL. Validation of the rapid immunochromatographic test IT-LEISH® for the diagnosis of human visceral leishmaniasis [Validação do teste imunocromatográfico rápido IT-LEISH® para o diagnóstico da leishmaniose visceral humana]. *Epidemiologia e Serviços de Saúde, Brasília* 2008;**17**(2):107–16.
- Delgado 2001** *{published data only}*
Delgado O, Feliciangeli MD, Coraspe V, Silva S, Perez A, Arias J. Value of a dipstick based on recombinant RK39 antigen for differential diagnosis of American visceral leishmaniasis from other sympatric endemic diseases in Venezuela. *Parasite (Paris, France)* 2001;**8**(4):355–7.
- Edrissian 2003** *{published data only}*
Edrissian GH, Shamsi S, Mohebbi M, Hajjarian H, Mamishi S, Desjeux P. Evaluation of rapid “dipstick rK39” test in diagnosis and serological survey of visceral leishmaniasis in humans and dogs in Iran. *Archives of Iranian Medicine* 2003;**6**(1):29–31.
- El-Safi 2003** *{published data only}*
El-Safi SH, Abdel-Haleem A, Hammad A, El-Basha I, Omer A, Kareem HG, et al. Field evaluation of latex agglutination test for detecting urinary antigens in visceral leishmaniasis in Sudan. *Eastern Mediterranean Health Journal = La revue de sante de la Mediterranee orientale = al-Majallah al-sihbiyah li-sharq al-mutawassit* 2003;**9**(4):844–55.
- Ferreira Dourado 2007** *{published data only}*
Ferreira Dourado Z, Delleon Silva H, Silveira-Lacerda E, García-Zapata MTA. Historical panorama of laboratorial diagnostic of visceral leishmaniasis until the appearance of immunochromatographic tests [Panorama histórico do diagnóstico laboratorial da leishmaniose visceral até o surgimento dos testes imunocromatográficos (rK39)]. *Revista de Patologia Tropical* 2007;**36**(3):205–14.
- Gavvani 2008** *{published data only}*
Gavvani AM, Vatan SK, Ghazanchaei A. KAtex antigen-detection test as a diagnostic tool for latent visceral leishmaniasis cases. *African Journal of Biotechnology* 2008;**7**(7):852–9.
- Goswami 2003** *{published data only}*
Goswami RP, Bairagi B, Kundu PK. K39 strip test—easy, reliable and cost-effective field diagnosis for visceral leishmaniasis in India. *Journal of the Association of Physicians of India* 2003;**51**:759–61.
- Goswami 2007** *{published data only}*
Goswami RP, Rahman M, Guha SK. Utility of K39 strip test in visceral leishmaniasis (VL) and HIV co-infected patients: an early report from Eastern India. *Journal of the Association of Physicians of India* 2007; Vol. 55:154–5.
- Goswami 2012** *{published data only}*
Goswami RP, Das S, Ray Y, Rahman M. Testing urine samples with rK39 strip as the simplest non-invasive field diagnosis for visceral leishmaniasis: an early report from eastern India. *Journal of Postgraduate Medicine* 2012;**58**:180–4.
- Gupta 1994** *{published data only}*
Gupta S, Srivastava JK, Pal A, Katiyar JC, Saxena KC, Dhawan BN, et al. Direct agglutination test and dot-ELISA in the serodiagnosis of visceral leishmaniasis (Kala-azar) - a comparative study. *Serodiagnosis and Immunotherapy in Infectious Disease* 1994;**6**:154–8.
- Hatam 2009** *{published data only}*
Hatam GR, Ghatee MA, Hossini SM, Sarkari B. Improvement of the newly developed latex agglutination test (Katex) for diagnosis of visceral leishmaniasis. *Journal of Clinical Laboratory Analysis* 2009;**23**(4):202–5.
- Hommel 2001** *{published data only}*
Hommel M, Sarkari B, Carney J, Chance ML. Katex for the diagnosis of human visceral leishmaniasis [Le Katex pour le diagnostic de la leishmaniose viscerale humaine]. *Médecine Tropicale: Revue du Corps de Santé Colonial* 2001;**61**(6):503–5. [PUBMED: 11980401]
- Iqbal 2002** *{published data only}*
Iqbal J, Hira PR, Saroj G, Philip R, Al-Ali F, Madda PJ, et al. Imported visceral leishmaniasis: diagnostic dilemmas and comparative analysis of three assays. *Journal of Clinical Microbiology* 2002;**40**(2):475–9.
- Jelinek 1999** *{published data only}*
Jelinek T, Eichenlaub S, Loscher T. Sensitivity and specificity of a rapid immunochromatographic test for diagnosis of visceral leishmaniasis. *European Journal of Clinical Microbiology & Infectious Diseases: official publication of the European Society of Clinical Microbiology* 1999;**18**(9):669–70.
- Khan 2009** *{published data only}*
Khan AM, Pandey K, Kumar V, Dutta P, Das P, Mahanta J. Sample survey for indigenous cases of kala-azar in Assam by rK39 dipstick test. *The Indian Journal of Medical Research* 2009; Vol. 129, issue 3:327–8.
- Khan 2010** *{published data only}*
Khan MG, Alam MS, Podder MP, Itoh M, Jamil KM, Haque R, et al. Evaluation of rK-39 strip test using urine for diagnosis of visceral leishmaniasis in an endemic area in Bangladesh. *Parasites & Vectors* 2010;**3**:114.
- Kumar 2006** *{published data only}*
Kumar R, Pai K, Kumar P, Pandey HP, Sundar S. Sero-epidemiological study of kala-azar in a village of Varanasi

- district, India. *Tropical Medicine & International Health : TM & IH* 2006;**11**(1):41–8.
- Lemos 2003** *{published data only}*
Lemos EM, Carvalho SF, Corey R, Dietze R. Evaluation of a rapid test using recombinant k39 antigen in the diagnosis of visceral leishmaniasis in Brazil [Avaliação do teste rápido utilizando o antígeno recombinante k39 no diagnóstico da leishmaniose visceral no Brasil]. *Revista da Sociedade Brasileira de Medicina Tropical* 2003;**36** Suppl 2:36–8.
- López Corbalán 2012** *{published data only}*
López Corbalán JC, Seguí Ripoll JM, Estebán Fernández JM^a. Urinary antigen as an effective and safe diagnosis of visceral leishmaniasis [Antígeno urinario como diagnóstico eficaz y seguro de leishmaniasis visceral]. *Atención Farmacéutica* 2012;**14**(4):298–9.
- Mandal 2008** *{published data only}*
Mandal J, Khurana S, Dubey ML, Bhatia P, Varma N, Malla N. Evaluation of direct agglutination test, rK39 Test, and ELISA for the diagnosis of visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 2008;**79**(1):76–8.
- Mansour 2009** *{published data only}*
Mansour D, Abass EM, Mahamoud A, el Harith A. Qualitative and semi-quantitative comparison of an rK39 strip test and direct agglutination test for detection of anti-Leishmania donovani antibodies in the Sudan. *Iranian Journal of Immunology : IJI* 2009;**6**(4):208–15.
- Marty 2007** *{published data only}*
Marty P, Delaunay P, Fissore C, Le Fichoux Y. Mediterranean leishmaniasis caused by *Leishmania infantum*. Update on the utility of the IT-Leish and ID-Pagia leishmaniasis tests [La leishmaniose méditerranéenne due a *Leishmania infantum* Mise au point—interets des tests de diagnostic rapide: IT–Leish et ID–Pagia leishmaniasis.]. *Médecine Tropicale: Revue du Corps de Santé Colonial* 2007;**67**(1):79–85.
- Mathur 2005** *{published data only}*
Mathur P, Samantaray J, Chauhan NK. Evaluation of a rapid immunochromatographic test for diagnosis of kala-azar & post kala-azar dermal leishmaniasis at a tertiary care centre of north India. *The Indian Journal of Medical Research* 2005;**122**(6):485–90.
- Matlashewski 2013** *{published data only}*
Matlashewski G, Das V N, Pandey K, Singh D, Das S, Ghosh A K, et al. Diagnosis of visceral leishmaniasis in Bihar India: comparison of the rK39 rapid diagnostic test on whole blood versus serum. *PLoS Neglected Tropical Diseases* 2013;**7**:e2233.
- Mbui 2011** *{published data only}*
Mbui J, Wasunna M, Kirigi G, Kinoti D, Riongoita M, De La Tour R, et al. Validation of two rapid tests for diagnosis of visceral leishmaniasis in Kenya. *Tropical Medicine and International Health* 2011; **Conference: 7th European Congress on Tropical Medicine and International Health Barcelona Spain. Conference Start: 20111003 Conference End: 20111006. Conference Publication: (var.pagings).** 16:283.
- Mohapatra 2010** *{published data only}*
Mohapatra TM, Singh DP, Sen MR, Bharti K, Sundar S. Comparative evaluation of rK9, rK26 and rK39 antigens in the serodiagnosis of Indian visceral leishmaniasis. *Journal of Infection in Developing Countries* 2010;**4**(2):114–7.
- Monno 2009** *{published data only}*
Monno R, Giannelli G, Rizzo C, De Vito D, Fumarola L. Recombinant K39 immunochromatographic test for diagnosis of human leishmaniasis. *Future Microbiology* 2009;**4**(2):159–70.
- Moura 2013** *{published data only}*
Moura AS, Lopes HM, Mourao MV, Morais MH. Performance of a rapid diagnostic test for the detection of visceral leishmaniasis in a large urban setting. *Revista da Sociedade Brasileira de Medicina Tropical* 2013;**46**:589–93.
- Mueller 2013** *{published data only}*
Mueller YK, Kolaczinski JH, Koech T, Lokwang P, Riongoita M, Velilla E, et al. Clinical epidemiology, diagnosis and treatment of visceral leishmaniasis in the Pokot endemic area of Uganda and Kenya. *American Society of Tropical Medicine and Hygiene* 2013;**90**(1):33–9.
- Ozderdem 2009** *{published data only}*
Ozderdem D, Eroglu F, Genç A, Demirkazık M, Koltas IS. Comparison of microscopic examination, rK39, and PCR for visceral leishmaniasis diagnosis in Turkey. *Parasitology Research* 2009;**106**(1):197–200.
- Ozkan 2008** *{published data only}*
Ozkan AT, Yalcinkaya T, Kilic S, Babur C, Schallig HD. Investigation of *Leishmania infantum* seropositivity in HIV/AIDS patients [HIV/AIDS hastalarinda *Leishmania infantum* seropozitifliginin arastirilmesi]. *Mikrobiyoloji Bulteni* 2008;**42**(1):113–7.
- Pappas 1983** *{published data only}*
Pappas MG, Hajkowsky R, Hockmeyer WT. Dot enzyme-linked immunosorbent assay (Dot-ELISA): a micro technique for the rapid diagnosis of visceral leishmaniasis. *Journal of Immunological Methods* 1983;**64**(1-2):205–14.
- Pappas 1984** *{published data only}*
Pappas MG, Hajkowsky R, Cannon LT Sr, Hockmeyer WT. Dot enzyme-linked immunosorbent assay (Dot-ELISA): comparison with standard ELISA and complement fixation assays for the diagnosis of human visceral leishmaniasis. *Veterinary Parasitology* 1984;**14**(3-4):239–49.
- Pappas 1984a** *{published data only}*
Pappas MG, Hajkowsky R, Hockmeyer WT. Standardization of the dot enzyme-linked immunosorbent assay (Dot-ELISA) for human visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 1984;**33**(6):1105–11.
- Pattabhi 2010** *{published data only}*
Pattabhi S, Whittle J, Mohamath R, El-Safi S, Moulton GG, Guderian JA, et al. Design, development and evaluation of rK28-based point-of-care tests for improving rapid diagnosis of visceral leishmaniasis. *PLoS Neglected Tropical Diseases* 2010;**4**(9):e822.

- Qu 1987** *{published data only}*
Qu JQ, Bao YF. Dot-ELISA using monoclonal antibodies for identification of Leishmania donovani. *Chinese Medical Journal* 1987;**100**(10):823–6.
- Qu 2000** *{published data only}*
Qu JQ, Guan LR, Shulidan I, Zuo XP, Chai JJ, Chen SB, et al. Rapid screening with a recombinant antigen (rK39) for diagnosis of visceral leishmaniasis using dipstick. *Zhongguo ji sheng chong xue yu ji sheng chong bing za zhi = Chinese Journal of Parasitology & Parasitic Diseases* 2000;**18**(3): 155–8.
- Riera 2004** *{published data only}*
Riera C, Fisa R, Lopez P, Ribera E, Carrio J, Falco V, et al. 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. *European Journal of Clinical Microbiology & Infectious Diseases : official publication of the European Society of Clinical Microbiology* 2004;**23**(12): 899–904.
- Rouf 2009** *{published data only}*
Rouf MA, Rahman ME, Islam MN, Islam MN, Ferdous NN, Hossain MA. Sensitivity, specificity and predictive values of immunochromatographic strip test in diagnosis of childhood kala-azar. *Mymensingh Medical Journal* 2009;**18** (1 Suppl):S1–5.
- Saghrouni 2009** *{published data only}*
Saghrouni F, Gaied-Mekis S, Fathallah A, Amri F, Ach H, Guizani I, et al. Immunochromatographic rK39 strip test in the serodiagnosis of visceral leishmaniasis in Tunisia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2009;**103**(12):1273–8.
- Saha 2011** *{published data only}*
Saha S, Goswami R, Pramanik N, Guha S K, Saha B, Rahman M, et al. Easy test for visceral Leishmaniasis and post-Kala-azar Dermal Leishmaniasis. *Emerging Infectious Diseases* 2011;**17**:1304–6.
- Salam 2008** *{published data only}*
Salam MA. Immunochromatographic strip test detection of anti-rK39 antibody for the diagnosis of kala-azar in an endemic zone of Bangladesh. *Pakistan Journal of Medical Sciences* 2008;**24**(4):497–501.
- Salam 2011** *{published data only}*
Salam MA, Khan MG, Mondal D. Urine antigen detection by latex agglutination test for diagnosis and assessment of initial cure of visceral leishmaniasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2011;**105**(5): 269–72.
- Salotra 2005** *{published data only}*
Salotra P, Singh R. Rapid & reliable diagnostic tests for visceral leishmaniasis. *The Indian Journal of Medical Research* 2005;**122**(6):464–7.
- Sarkari 2005** *{published data only}*
Sarkari B, Chance M, Hommel M. A capture ELISA for the diagnosis of visceral leishmaniasis using a monoclonal antibody against a leishmanial urinary antigen. *Iranian Biomedical Journal* 2005;**9**(3):117–22.
- Sarker 2003** *{published data only}*
Sarker CB, Momen A, Jamal MF, Siddiqui NI, Siddiqui FM, Chowdhury KS, et al. Immunochromatographic (rK39) strip test in the diagnosis of visceral leishmaniasis in Bangladesh. *Mymensingh Medical Journal* 2003;**12**(2):93–7.
- Schallig 2002** *{published data only}*
Schallig HD, Canto-Cavalheiro M, da Silva ES. Evaluation of the direct agglutination test and the rK39 dipstick test for the sero-diagnosis of visceral leishmaniasis. *Memorias do Instituto Oswaldo Cruz* 2002;**97**(7):1015–8.
- Schoone 2001** *{published data only}*
Schoone GJ, Hailu A, Kroon CC, Nieuwenhuys JL, Schallig HD, Oskam L. A fast agglutination screening test (FAST) for the detection of anti-Leishmania antibodies. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2001; **95**(4):400–1.
- Scott 1991** *{published data only}*
Scott JM, Shreffler WG, Ghalib HW, el Asad A, Siddig M, Badaro R, et al. A rapid and simple diagnostic test for active visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 1991;**44**(3):272–7.
- Senaldi 1996** *{published data only}*
Senaldi G, Xiao-su H, Hoessli DC, Bordier C. Serological diagnosis of visceral leishmaniasis by a dot-enzyme immunoassay for the detection of a Leishmania donovani-related circulating antigen. *Journal of Immunological Methods* 1996;**193**(1):9–15.
- Shamsuzzaman 2003** *{published data only}*
Shamsuzzaman AK, Hossain MA, Musa AK, Hasan MU, Dhar DK. A preliminary report on culture of Leishmania donovani in Mymensingh Medical College and evaluation of new immuno-chromatography test (ICT). *Mymensingh Medical Journal* 2003;**12**(1):51–4.
- Sharma 2009** *{published data only}*
Sharma NL, Mahajan VK, Negi AK, Verma GK. The rK39 immunochromatographic dipstick testing: a study for K39 seroprevalence in dogs and human leishmaniasis patients for possible animal reservoir of cutaneous and visceral leishmaniasis in endemic focus of Satluj river valley of Himachal Pradesh (India). *Indian Journal of Dermatology, Venereology, and Leprology* 2009;**75**(1):52–5.
- Silva 2005** *{published data only}*
Silva ES, Schoone GJ, Gontijo CM, Brazil RP, Pacheco RS, Schallig HD. Application of direct agglutination test (DAT) and fast agglutination screening test (FAST) for sero-diagnosis of visceral leishmaniasis in endemic area of Minas Gerais, Brazil. *Kinetoplastid Biology and Disease* 2005;**4**:4.
- Singh 2009** *{published data only}*
Singh DP, Sundar S, Mohapatra TM. The rK39 strip test is non-predictor of clinical status for kala-azar. *BMC Research Notes* 2009;**2**:187.
- Singh 2010** *{published data only}*
Singh DP, Goyal RK, Singh RK, Sundar S, Mohapatra TM. In search of an ideal test for diagnosis and prognosis of kala-

- azar. *Journal of Health, Population, and Nutrition* 2010;**28**(3):281–5.
- Sinha 2008** *{published data only}*
Sinha PK, Bimal S, Pandey K, Singh SK, Ranjan A, Kumar N, et al. A community-based, comparative evaluation of direct agglutination and rK39 strip tests in the early detection of subclinical *Leishmania donovani* infection. *Annals of Tropical Medicine and Parasitology* 2008;**102**(2): 119–25.
- Srivastava 1988** *{published data only}*
Srivastava L, Singh VK. Diagnosis of Indian kala-azar by dot enzyme-linked immunosorbent assay (dot-ELISA). *Annals of Tropical Medicine and Parasitology* 1988;**82**(4):331–4.
- Sundar 2002** *{published data only}*
Sundar S, Sahu M, Mehta H, Gupta A, Kohli U, Rai M, et al. Noninvasive management of Indian visceral leishmaniasis: clinical application of diagnosis by K39 antigen strip testing at a kala-azar referral unit. *Clinical Infectious Diseases : an official publication of the Infectious Diseases Society of America* 2002;**35**(5):581–6.
- Sundar 2002a** *{published data only}*
Sundar S, Pai K, Sahu M, Kumar V, Murray HW. Immunochromatographic strip-test detection of anti-K39 antibody in Indian visceral leishmaniasis. *Annals of Tropical Medicine and Parasitology* 2002;**96**(1):19–23.
- Sundar 2002b** *{published data only}*
Sundar S, Rai M. Laboratory diagnosis of visceral leishmaniasis. *Clinical and Diagnostic Laboratory Immunology* 2002;**9**(5):951–8.
- Sundar 2005** *{published data only}*
Sundar S, Agrawal S, Pai K, Chance M, Hommel M. Detection of leishmanial antigen in the urine of patients with visceral leishmaniasis by a latex agglutination test. *American Journal of Tropical Medicine and Hygiene* 2005;**73**(2):269–71.
- Sundar 2005a** *{published data only}*
Sundar S. Serological diagnosis of Indian visceral leishmaniasis - direct agglutination test versus rK39 strip test and comparison of two strip formats. 2005; Vol. 73, issue 6:302.
- Sundar 2006** *{published data only}*
Sundar S, Singh RK, Maurya R, Kumar B, Chhabra A, Singh V, et al. Serological diagnosis of Indian visceral leishmaniasis: direct agglutination test versus rK39 strip test. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2006;**100**(6):533–7.
- Sundar 2006a** *{published data only}*
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. *Journal of Clinical Microbiology* 2006;**44**(1): 251–3.
- Teran-Angel 2010** *{published data only}*
Teran-Angel G, Rodriguez V, Silva R, Zerpa O, Schallig H, Ulrich M, et al. Non invasive diagnostic tools for visceral leishmaniasis: a comparison of the immunoserological tests DAT, rK26 and rK39 [Herramientas no invasivas en Venezuela: comparacion entre las pruebas inmunoserologicas DAT, rK26 y rK39 en el diagnostico de leishmaniasis visceral.]. *Biomedica : Revista del Instituto Nacional de Salud* 2010;**30**(1):39–45.
- Walton 1986** *{published data only}*
Walton BC, Pappas MG, Sierra M Jr, Hajkowsky R, Jackson P, Custodio R. Field use of the Dot-ELISA test for visceral leishmaniasis in Honduras. *Bulletin of the Pan American Health Organization* 1986;**20**(2):147–56.
- Walton 1987** *{published data only}*
Walton BC, Pappas MG, Sierra M, Hajkowsky R, Jackson P, Custodio R. Field use of the DOT-ELISA test for visceral leishmaniasis in Honduras [La prueba DOT-ELISA para la detección de campo de la leishmaniasis visceral en Honduras]. *Boletín de la Oficina Sanitaria Panamericana* 1987;**102**(1):19–28.
- Welch 2008** *{published data only}*
Welch RJ, Anderson BL, Litwin CM. Rapid immunochromatographic strip test for detection of anti-K39 immunoglobulin G antibodies for diagnosis of visceral leishmaniasis. *Clinical and Vaccine Immunology : CVI* 2008;**15**(9):1483–4.
- Zijlstra 2001** *{published data only}*
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. *Tropical Medicine & International Health : TM & IH* 2001;**6**(2):108–13.

Additional references

- Alvar 2008**
Alvar J, Aparicio P, Aseffa A, Den Boer M, Canavate C, Dedet JP, et al. The relationship between leishmaniasis and AIDS: the second 10 years. *Clinical Microbiology Reviews* 2008;**21**(2):334–59.
- Alvar 2012**
Alvar J, Velez ID, Bern C, Herrero M, Desjeux P, Cano J, et al. Leishmaniasis worldwide and global estimates of its incidence. *PLoS One* 2012;**7**(5):e35671.
- Babiker 2007**
Babiker ZO, Davidson R, Mazinda C, Kipnetich S, Ritmeijer K. Utility of lymph node aspiration in the diagnosis of visceral leishmaniasis in Sudan. *American Journal of Tropical Medicine and Hygiene* 2007;**76**(4): 689–93.
- Bachmann 2006**
Bachmann LM, Puhon MA, ter Riet G, Bossuyt PM. Sample sizes of studies on diagnostic accuracy: literature survey. *BMJ (Clinical research ed.)* 2006;**332**(7550):1127–9.
- Baughman 2008**
Baughman AL, Bisgard KM, Cortese MM, Thompson WW, Sanden GN, Strelbel PM. Utility of composite reference standards and latent class analysis in evaluating the

- clinical accuracy of diagnostic tests for pertussis. *Clinical and Vaccine Immunology* : CVI 2008;**15**(1):106–14.
- Bhattacharyya 2013**
Bhattacharyya T, Boelaert M, Miles MA. Comparison of visceral leishmaniasis diagnostic antigens in African and Asian *Leishmania donovani* reveals extensive diversity and region-specific polymorphisms. *PLoS Neglected Tropical Diseases* 2013;**7**(2):e2057.
- Bhattarai 2009**
Bhattarai NR, Van der Auwera G, Khanal B, De Doncker S, Rijal S, Das ML, et al. PCR and direct agglutination as *Leishmania* infection markers among healthy Nepalese subjects living in areas endemic for Kala-Azar. *Tropical Medicine & International Health : TM & IH* 2009;**14**(4): 404–11.
- Black 2002**
Black MA, Craig BA. Estimating disease prevalence in the absence of a gold standard. *Statistics in Medicine* 2002;**21** (18):2653–69.
- Boelaert 1999**
Boelaert M, El Safi S, Jacquet D, de Muynck A, van der Stuyft P, Le Ray D. Operational validation of the direct agglutination test for diagnosis of visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 1999;**60** (1):129–34.
- Boelaert 2000**
Boelaert M, Criel B, Leeuwenburg J, Van Damme W, Le Ray D, Van der Stuyft P. Visceral leishmaniasis control: a public health perspective. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2000;**94**(5):465–71.
- Boelaert 2004**
Boelaert M, Rijal S, Regmi S, Singh R, Karki B, Jacquet D, et al. A comparative study of the effectiveness of diagnostic tests for visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 2004;**70**(1):72–7.
- Boelaert 2007**
Boelaert M, Bhattacharya S, Chappuis F, El Safi SH, Hailu A, Mondal D, et al. Evaluation of rapid diagnostic tests: visceral leishmaniasis. *Nature Reviews Microbiology* November 2007;**5**(11):S30–S39.
- Branscum 2005**
Branscum AJ, Gardner IA, Johnson WO. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Preventive Veterinary Medicine* 2005;**68**(2-4): 145–63.
- Chappuis 2006**
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 (Clinical Research Ed.)* 2006;**333** (7571):723.
- Chappuis 2007**
Chappuis F, Sundar S, Hailu A, Ghalib H, Rijal S, Peeling RW, et al. Visceral leishmaniasis: what are the needs for diagnosis, treatment and control?. *Nature Reviews Microbiology* 2007;**5**(11):873–82.
- Cunningham 2012**
Cunningham J, Hasker E, Das P, El Safi S, Goto H, Mondal D, et al. A global comparative evaluation of commercial immunochromatographic rapid diagnostic tests for visceral leishmaniasis. *Clinical Infectious Diseases : an official publication of the Infectious Diseases Society of America* 2012;**55**(10):1312–9.
- De Almeida 2006**
De Almeida Silva L, Romero HD, Prata A, Costa RT, Nascimento E, Carvalho SF, et al. Immunologic tests in patients after clinical cure of visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 2006;**75** (4):739–43.
- Deeks 2005**
Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of Clinical Epidemiology* 2005;**58**(7):882–93.
- den Boer 2006**
den Boer M, Davidson RN. Treatment options for visceral leishmaniasis. *Expert Review of Anti-infective Therapy* 2006;**4**(2):187–97.
- Dendukuri 2001**
Dendukuri N, Joseph L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics* 2001;**57**(1):158–67.
- Enoe 2000**
Enoe C, Georgiadis MP, Johnson WO. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Preventive Veterinary Medicine* 2000;**45**(1-2):61–81.
- Evans 1992**
Evans TG, Teixeira MJ, McAuliffe IT, Vasconcelos I, Vasconcelos AW, Sousa Ade A, et al. Epidemiology of visceral leishmaniasis in northeast Brazil. *Journal of Infectious Diseases* 1992;**166**(5):1124–32.
- Gelman 1995**
Gelman A, Carlin JC, Stern H, Rubin DB. *Bayesian Data Analysis*. New York: Chapman and Hall, 1995.
- Greenland 2006**
Greenland S. Bayesian perspectives for epidemiological research: I. Foundations and basic methods. *International Journal of Epidemiology* 2006;**35**:765–75.
- Hailu 1990**
Hailu A. Pre- and post-treatment antibody levels in visceral leishmaniasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1990;**84**(5):673–5.
- Hailu 2009**
Hailu A, Gramiccia M, Kager PA. Visceral leishmaniasis in Aba-Roba, south-western Ethiopia: prevalence and incidence of active and subclinical infections. *Annals of Tropical Medicine and Parasitology* 2009;**103**(8):659–70.
- Harhay 2011**
Harhay MO, Olliaro PL, Vaillant M, Chappuis F, Lima MA, Ritmeijer K, et al. Who is a typical patient with

- visceral leishmaniasis? Characterizing the demographic and nutritional profile of patients in Brazil, East Africa, and South Asia. *American Journal of Tropical Medicine and Hygiene* 2011;**84**(4):543–50.
- Harith 1986**
Harith AE, Kolk AH, Kager PA, Leeuwenburg J, Muigai R, Kiugu S, et al. A simple and economical direct agglutination test for serodiagnosis and sero-epidemiological studies of visceral leishmaniasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1986;**80**(4):583–6.
- Herwaldt 1999**
Herwaldt BL. Leishmaniasis. *Lancet* 1999;**354**(9185): 1191–9.
- Ho 1982**
Ho M, Siongok TK, Lyerly WH, Smith DH. Prevalence and disease spectrum in a new focus of visceral leishmaniasis in Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1982;**76**(6):741–6.
- Hui 1980**
Hui SL, Walter SD. Estimating the error rates of diagnostic tests. *Biometrics* 1980;**36**(1):167–71.
- Macaskill 2010**
Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C editor(s). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 1.0*. The Cochrane Collaboration, 2010.
- Menten 2013**
Menten J, Boelaert M, Lesaffre E. Bayesian meta-analysis of diagnostic tests allowing for imperfect reference standards. *Statistics in Medicine* 2013;**32**(30):5398–413.
- Murray 2005**
Murray HW, Berman JD, Davies CR, Saravia NG. Advances in leishmaniasis. *Lancet* 2005;**366**(9496):1561–77.
- Ostyn 2011**
Ostyn B, Gidwani K, Khanal B, Picado A, Chappuis F, Singh SP, et al. Incidence of symptomatic and asymptomatic *Leishmania donovani* infections in high-endemic foci in India and Nepal: a prospective study. *PLoS Neglected Tropical Diseases* 2011;**5**(10):e1284.
- Pasquau 2005**
Pasquau F, Ena J, Sanchez R, Cuadrado JM, Amador C, Flores J, et al. Leishmaniasis as an opportunistic infection in HIV-infected patients: determinants of relapse and mortality in a collaborative study of 228 episodes in a Mediterranean region. *European Journal of Clinical Microbiology* 2005;**24**(6):411–8.
- PATH 2008**
PATH. RDT Info. Current information on rapid diagnostic tests. www.rapid-diagnostics.org/index.htm (accessed 16 August 2010).
- Peeling 2007**
Peeling RW, Smith PG, Bossuyt PMM. A guide for diagnostic evaluations. *Nature Reviews Microbiology* 2007;**4** (9 Suppl):S2–S6.
- R Development Team 2010**
R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2010.
- Raguenaud 2007**
Raguenaud ME, Jansson A, Vanlerberghe V, Deborggraeve S, Dujardin JC, Orfanos G, et al. Epidemiology and clinical features of patients with visceral leishmaniasis treated by an MSF clinic in Bakool region, Somalia, 2004–2006. *PLoS Neglected Tropical Diseases* 2007;**1**(1):e85.
- Reitsma 2005**
Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982–90.
- Rittig 2000**
Rittig MG, Bogdan C. Leishmania-host-cell interaction: complexities and alternative views. *Parasitology Today* 2000;**16**(7):292–7.
- Sarker 2004**
Sarker CB, Alam KS, Jamal MF, Rahman S, Huq MH, Musa AK, et al. Sensitivity of splenic and bone marrow aspirate study for diagnosis of kala-azar. *Mymensingh Medical Journal* 2004;**13**(2):130–3.
- Seaman 1996**
Seaman J, Mercer AJ, Sondorp E. The epidemic of visceral leishmaniasis in western Upper Nile, southern Sudan: course and impact from 1984 to 1994. *International Journal of Epidemiology* 1996;**25**(4):862–71.
- Spiegelhalter 2003**
Spiegelhalter D, Thomas A, Best N, Lunn D. WinBUGS User Manual Version 1.4. Cambridge: MRC Biostatistics Unit, 2003.
- Sullivan 2003**
Sullivan Pepe M. *The Statistical Evaluation of Medical tests for Classification and Prediction*. Oxford: Oxford University Press, 2003.
- van Griensven 2010**
van Griensven J, Balasegaram M, Meheus F, Alvar J, Lynen L, Boelaert M. Combination therapy for visceral leishmaniasis. *Lancet Infectious Diseases* 2010;**10**(3):184–94.
- Veeken 2003**
Veeken H, Ritmeijer K, Seaman J, Davidson R. Comparison of an rK39 dipstick rapid test with direct agglutination test and splenic aspiration for the diagnosis of kala-azar in Sudan. *Tropical Medicine and International Health* 2003;**8** (2):164–7.
- Verde 2010**
Verde PE. Meta-analysis of diagnostic test data: a bivariate Bayesian modeling approach. *Statistics in Medicine* 2010;**29** (30):3088–102. [PUBMED: 21170904]
- Whitehead 2002**
Whitehead, A. *Meta-Analysis of Controlled Clinical Trials*. Chichester: John Wiley & Sons, 2002.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529–36. [PUBMED: 22007046]

WHO 2005

World Health Organization. Regional strategic framework for elimination of kala-azar from the South-east Asia region (2005-2015). New Delhi: WHO Regional Office for South-East Asia 2005.

WHO 2010

World Health Organization. Control of the leishmaniasis: report of a meeting of the WHO Expert Committee on the Control of Leishmaniasis, Geneva, 22-26 March 2010. WHO technical report series 2010, issue 949.

Zhou 2002

Zhou XH, McClish DK, Obuchowski NA. *Statistical Methods in Diagnostic Medicine*. New York: John Wiley, 2002.

Zijlstra 1992

Zijlstra EE, Ali MS, el-Hassan AM, el-Toum IA, Satti M, Ghalib HW, et al. Kala-azar: a comparative study of parasitological methods and the direct agglutination test in diagnosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1992;**86**(5):505–7.

Zijlstra 1994

Zijlstra EE, el-Hassan AM, Ismael A, Ghalib HW. Endemic kala-azar in eastern Sudan: a longitudinal study on the incidence of clinical and subclinical infection and post-kala-azar dermal leishmaniasis. *American Journal of Tropical Medicine and Hygiene* 1994;**51**(6):826–36.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Boelaert 2004 - classic

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 310</p> <p>Age (reported for 181 new VL cases): median 25, interquartile range 13-36</p> <p>Sex (reported for 181 new VL cases): male:female ratio 1.7:1</p> <p>Presenting signs and symptoms: fever of 14 days or more and splenomegaly</p> <p>Frequency of VL: 59%</p> <p>HIV: not reported</p> <p>Clinical setting: tertiary care centre (B.P. Koirala Institute of Health Sciences in Dharan)</p> <p>Country: Nepal</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test</p> <p>Brand: InSure Rapid test for Visceral Leishmaniasis, InBios International, Washington, USA</p> <p>Sample: serum</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL (3/310 participants were relapse cases)</p> <p>Sample: bone marrow; if negative, possible and no malaria: spleen</p> <p>Technique: direct smear (Giemsa stain)</p> <p>Definition of VL: bone marrow or spleen parasitology positive</p> <p>Definition of non-VL: bone marrow negative, and spleen negative or spleen aspirate not done</p> <p>Reference standard category: parasitology including spleen aspirate - no serology</p>		
Flow and timing	one participant was excluded because of a missing serum sample		
Comparative			
Notes	Same study as Boelaert 2004 - LCA but different analysis		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		

Boelaert 2004 - classic (Continued)

Did the study avoid inappropriate exclusions?	Yes			
				Low
DOMAIN 2: Index Test All tests				
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear			
				Low
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Boelaert 2004 - LCA

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 310</p> <p>Age (reported for 181 new VL cases): median 25, interquartile range 13-36</p> <p>Sex (reported for 181 new VL cases): male:female ratio 1.7:1</p> <p>Presenting signs and symptoms: fever of 14 days or more and splenomegaly</p> <p>Frequency of VL: 59%</p> <p>HIV: not reported</p> <p>Clinical setting: tertiary care centre (B.P. Koirala Institute of Health Sciences in Dharan)</p> <p>Country: Nepal</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test</p> <p>Brand: InSure Rapid test for Visceral Leishmaniasis, InBios International, Washington, USA</p> <p>Sample: serum</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL (3/310 participants were relapse cases)</p> <p>Approach: latent class analysis</p> <p>Tests included in latent class analysis: rK39 immunochromatographic test, pancytopenia, IFAT, DAT, formol-gel test, and parasitology (direct smear; Giemsa stain) of bone marrow or spleen samples</p>		
Flow and timing	One participant was excluded because of a missing serum sample		
Comparative			
Notes	Same study as Boelaert 2004 - classic but different analysis		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			

Boelaert 2004 - LCA (Continued)

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear			
Low				
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear			
Low				
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Boelaert 2008 - Ethiopia

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 38</p> <p>Age: median 25 years; range 16-49. No children \leq 12 years</p> <p>Sex: 95% men</p> <p>Presenting signs and symptoms: fever of 2 weeks or more, and splenomegaly or lymphadenopathy or both</p> <p>Pregnant women were excluded. Patients with a thick-film-positive malaria episode were excluded</p>

	<p>Estimated frequency of VL: 57% HIV: not tested Clinical setting: university hospital (Kahsay Abera Hospital in Humera and Gondar College of Medical Sciences) Country: Ethiopia Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; Sample: serum Type: latex agglutination test in urine; Brand: KAtex, Kalon Biological Ltd, Guildford, UK; Sample: fresh urine</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL Approach: latent class analysis Tests included in latent class analysis: rK39 immunochromatographic test, latex agglutination test, DAT, and parasitology (culture and direct smear (Giemsa stain)) of lymph node, bone marrow or spleen sample</p>		
Flow and timing	Three participants were excluded because of missing data.		
Comparative			
Notes	Same study as Diro 2007 but different analysis.		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
			Low

Boelaert 2008 - Ethiopia (Continued)

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard?	Yes		

Boelaert 2008 - India

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 352</p> <p>Age (computed on 260/352 participants): median 15 years; range 2-65; 44% of the study population was \leq 12 years old</p> <p>Sex (computed on 260/352 participants): 55% men</p> <p>Presenting signs and symptoms: fever (more or less than 2 weeks), and splenomegaly or lymphadenopathy or both</p> <p>No children < 2 years old; no pregnant women</p> <p>Estimated frequency of VL: 80%</p> <p>HIV: not reported</p> <p>Clinical setting: research centre (Kala-Azar Medical Research and Training Centre in Muzaffarpur)</p> <p>Country: India</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>

Index tests	Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; Sample: serum Type: latex agglutination test in urine; Brand: KAtex, Kalon Biological Ltd, Guildford, UK; Sample: fresh urine		
Target condition and reference standard(s)	Target condition: clinical VL Approach: latent class analysis Tests included in latent class analysis: rK39 immunochromatographic test, latex agglutination test, DAT, and parasitology (direct smear (Giemsa stain)) of lymph node, bone marrow or spleen sample		
Flow and timing	No withdrawals		
Comparative			
Notes	Same study as Sundar 2007 but different analysis		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Boelaert 2008 - India (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear			
Low				
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Boelaert 2008 - Kenya

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 308 Age: median 13 years; range 1-75; 45% of the study population was ≤12 years old Sex: 65% men Presenting signs and symptoms: fever for 2 weeks or more, and splenomegaly or lymphadenopathy or both; no children < 2 years old; no pregnant women; no patients with a thick-film-positive malaria episode Estimated frequency of VL: 61% HIV: not reported Clinical setting: hospital (Kabarnet hospital in Baringo district) Country: Kenya Endemic <i>Leishmania species:</i> <i>L. donovani</i></p>
Index tests	<p>Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; Sample: serum Type: latex agglutination test in urine; Brand: KAtex, Kalon Biological Ltd, Guildford, UK; Sample: fresh urine</p>

Target condition and reference standard(s)	Target condition: clinical VL Approach: latent class analysis Tests included in latent class analysis: rK39 immunochromatographic test, latex agglutination test, DAT, and parasitology (direct smear (Giemsa stain)) of lymph node, bone marrow or spleen sample		
Flow and timing	One participant was excluded because of missing data.		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Boelaert 2008 - Kenya (Continued)

				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Boelaert 2008 - Nepal

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 158</p> <p>Age: median 23 years; range 2-68; 25% of the study population was ≤ 12 years old</p> <p>Sex: 59% men</p> <p>Presenting signs and symptoms: fever for 2 weeks or more, and splenomegaly or lymphadenopathy or both;</p> <p>No children < 2 years old; no pregnant women</p> <p>Estimated frequency of VL: 71%</p> <p>HIV: not reported</p> <p>Clinical setting: tertiary care centre (B.P. Koirala Institute of Health Sciences in Dharan)</p> <p>Country: Nepal</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>
Index tests	<p>Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; Sample: serum</p> <p>Type: latex agglutination test in urine; Brand: KAtext, Kalon Biological Ltd, Guildford, UK; Sample: fresh urine</p>
Target condition and reference standard(s)	<p>Target condition: clinical VL</p> <p>Approach: latent class analysis</p> <p>Tests included in latent class analysis: rK39 immunochromatographic test, latex agglutination test, DAT, and parasitology (direct smear (Giemsa stain)) of lymph node, bone marrow or spleen sample</p>

Boelaert 2008 - Nepal (Continued)

Flow and timing	No withdrawals		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and ref-	Yes		

Boelaert 2008 - Nepal (Continued)

reference standard?			
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard?	Yes		

Boelaert 2008 - Sudan

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 294 Age: median 20 years; range 2-72; 36% of the study population was ≤12 years old Sex: 64% men Presenting signs and symptoms: fever for 2 weeks or more, and splenomegaly or lymphadenopathy or both No children < 2 years old; no pregnant women; no patients with a thick-film-positive malaria episode Estimated frequency of VL: 37% HIV: not reported Clinical setting: two health centres (Umalkhair and Tabarakelleh) Country: Sudan Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>
Index tests	<p>Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; Sample: serum Type: latex agglutination test in urine; Brand: KAtex, Kalon Biological Ltd, Guildford, UK; Sample: fresh urine</p>
Target condition and reference standard(s)	<p>Target condition: clinical VL Approach: latent class analysis Tests included in latent class analysis: rK39 immunochromatographic test, latex agglutination, DAT, and parasitology (direct smear (Giemsa stain)) of lymph node, bone marrow or spleen sample</p>
Flow and timing	Three participants were excluded because of missing data.
Comparative	
Notes	
Methodological quality	

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Boelaert 2008 - Sudan (Continued)

Did all patients receive a reference standard?	Yes		

Chappuis 2003

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 195 Age (data available for 184 included participants): mean 23 years Sex (data available for 184 included participants): 59% men Presenting signs and symptoms: fever for 2 weeks or more and clinical splenomegaly Frequency of VL: 76% HIV: 0% Clinical setting: tertiary care centre (B.P. Koirala Institute of Health Sciences in Dharan) Country: Nepal Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: InSure Rapid Test for Visceral Leishmaniasis, InBios Interational, Washington, USA; Sample: serum</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL Sample: bone marrow; if bone marrow negative and spleen aspiration possible: spleen Technique: direct smear examination Reference standard category: parasitology including spleen aspirate - no serology</p>		
Flow and timing	Eleven participants with an uncertain diagnosis were excluded		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		

Chappuis 2003 (Continued)

Did the study avoid inappropriate exclusions?	Yes		
Low			
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Low			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Low			
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard?	Yes		

Chappuis 2005 - DiaMed

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 255</p> <p>Age (among 131 VL patients): mean age 14 years (standard deviation 11)</p> <p>Age (among 112 non-VL patients): mean age 16 years (standard deviation 13)</p> <p>Sex (among 131 VL patients): 73% men;</p> <p>Sex (among 112 non-VL patients): 53% men</p> <p>Presenting signs and symptoms: fever for 2 weeks or more and either clinical splenomegaly or wasting syndrome</p> <p>Frequency of VL: 54%</p> <p>6/112 non-VL patients had previous history of treatment for VL</p> <p>HIV: not reported</p> <p>Clinical setting: a 120-bed district hospital (Amudat Hospital)</p> <p>Country: Uganda</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test (dual test for visceral leishmaniasis and malaria); Brand: DUAL-IT L/M dipstick, Diamed AG, Switzerland; Sample: blood</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL</p> <p>Combination of parasitology of spleen sample (direct smear, Giemsa stain), PCR of spleen sample, serology (DAT), and response to VL therapy</p> <p>Definition of VL: (1) positive spleen aspirate in two laboratories (by microscopic reading in local or regional laboratory and by PCR in reference laboratory) or (2) DAT >1:12800 and response to VL therapy</p> <p>Definition of non-VL: (1) DAT titres of <1:1600; or (2) borderline DAT titres with negative spleen aspirate in two laboratories with no diagnosis of VL made during the following 6 months</p> <p>Reference standard category: combination of parasitology and serology</p>		
Flow and timing	Twelve participants were excluded: 2 persons died and 10 had discrepant results for the reference standard		
Comparative			
Notes	Same patient population as Chappuis 2005 - InBios but different brand of rK39 immunochromatographic test. Considered as two studies in meta-analysis		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Chappuis 2005 - DiaMed (Continued)

Was a case-control design avoided?	Yes			
Did the study avoid inappropriate exclusions?	Yes			
				Low
DOMAIN 2: Index Test All tests				
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes			
				Low
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Chappuis 2005 - InBios

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 255</p> <p>Age (among 131 VL patients): mean age 14 years (standard deviation 11)</p> <p>Age (among 112 non-VL patients): mean age 16 years (standard deviation 13)</p> <p>Sex (among 131 VL patients): 73% men;</p> <p>Sex (among 112 non-VL patients): 53% men</p> <p>Presenting signs and symptoms: fever for 2 weeks or more and either clinical splenomegaly or wasting syndrome</p> <p>Frequency of VL: 54%</p> <p>6/112 non-VL patients had previous history of treatment for VL</p> <p>HIV: not reported</p> <p>Clinical setting: a 120-bed district hospital (Amudat Hospital)</p> <p>Country: Uganda</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: InBios International, Washington, USA; Sample: plasma</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL</p> <p>Combination of parasitology of spleen sample (direct smear, Giemsa stain), PCR of spleen sample, serology (DAT), and response to VL therapy</p> <p>Definition of VL: (1) positive spleen aspirate in two laboratories (by microscopic reading in local or regional laboratory and by PCR in reference laboratory) or (2) DAT >1:12800 and response to VL therapy</p> <p>Definition of non-VL: (1) DAT titres of <1:1600; or (2) borderline DAT titres with negative spleen aspirate in two laboratories with no diagnosis of VL made during the following 6 months</p> <p>Reference standard category: combination of parasitology and serology</p>		
Flow and timing	Twelve participants were excluded: 2 persons died and 10 had discrepant results for the reference standard		
Comparative			
Notes	Same patient population as Chappuis 2005 - DiaMed but different brand of rK39 immunochromatographic test. Considered as two studies in meta-analysis		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Chappuis 2005 - InBios (Continued)

Was a case-control design avoided?	Yes			
Did the study avoid inappropriate exclusions?	Yes			
				Low
DOMAIN 2: Index Test All tests				
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes			
				Low
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Chappuis 2006b

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 154 Age (among 85 VL patients): mean age 22 years (standard deviation 13) Age (among 57 non-VL patients): mean age 26 (standard deviation 15) Sex (among 85 VL patients): 57% men Age (among 57 non-VL patients): 65% men Presenting signs and symptoms: fever for 2 weeks or more and clinical splenomegaly All patients received malaria treatment 12/142 participants had previous history of VL Frequency of VL: 60% HIV: not tested Clinical setting: 15-bed first-referral government hospital (Rangeli District Hospital) Country: Nepal Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: InSure One-Step Rapid Test for Visceral Leishmaniasis, InBios International, Washington, USA; Sample: serum</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL Combination of parasitology of bone marrow sample (direct smear) and serology (DAT) and response to therapy Definition of VL: (1) positive bone marrow aspirate during initial evaluation or follow-up; or (2) DAT titre \geq 1:3200 and absence of response to antimalarial treatment but successful response to anti-leishmanial therapy Definition of non-VL: negative bone marrow aspirate with (1) DAT titre < 1:3200 or with (2) positive DAT but a definite cure with non-VL specific therapy Reference standard category: combination of parasitology and serology</p>		
Flow and timing	Twelve patients were excluded: 4 with incomplete diagnostic work-up and 8 with an uncertain diagnosis		
Comparative			
Notes	Exclusion of patients with previous history of VL had no significant effect on Se and Sp estimates (data not reported). Data on accuracy of latex agglutination test were incomplete and not included		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Chappuis 2006b (Continued)

Was a case-control design avoided?	Yes			
Did the study avoid inappropriate exclusions?	Yes			
				Low
DOMAIN 2: Index Test All tests				
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear			
				Low
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Cota 2013 - composite 1

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 113 Age (VL cases): mean 41.0 years (standard deviation 10.8) Age (non-VL cases): mean 40.0 years (standard deviation 9.8) Sex (VL cases): 24% men (11 men out of 46) Sex (non-VL cases): 42% men (28 men out of 67) Presenting signs and symptoms: clinical suspicion of VL, ie more than 14 days of fever or splenomegaly or cytopenia Previous VL diagnosis: 24 out of 113 participants had a previous history of VL Frequency of VL: 41% HIV: 100% Clinical setting: HIV cohort in reference hospital in Belo Horizonte (Eduardo de Menezes Hospital, Fundação Hospitalar do Estado de Minas Gerais) Country: Brazil Endemic <i>Leishmania</i> species: <i>L. infantum</i></p>		
Index tests	Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; sample: serum		
Target condition and reference standard(s)	<p>Target condition: clinical VL Samples: bone marrow aspirate and venous blood Techniques: direct smear and culture of bone marrow aspirate; PCR on peripheral blood; and three serological tests: index test, indirect fluorescent antibody test (IFI, Bio-Manguinhos) and direct agglutination test (prototype kit) Definition of VL and non-VL: decided after clinical follow-up by an adjudication committee with four members. They evaluated results of all tests available, including tests performed for other diagnostic possibilities Reference standard category: parasitology and serology</p>		
Flow and timing			
Comparative			
Notes	Another estimate of Se and Sp based on latent class analysis in the same study population was not included because there was no information about the assessment of the conditional independence assumption (explained in the Criteria for considering studies for this review and in Appendix 1).		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Cota 2013 - composite 1 (Continued)

Was a case-control design avoided?	Yes			
Did the study avoid inappropriate exclusions?	Yes			
				High
DOMAIN 2: Index Test All tests				
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes			
				Low
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
Were the reference standard results interpreted without knowledge of the results of the index tests?	No			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Cota 2013 - composite 2

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 113</p> <p>Age (VL cases): mean 41.0 years (standard deviation 10.8)</p> <p>Age (non-VL cases): mean 40.0 years (standard deviation 9.8)</p> <p>Sex (VL cases): 24% men (11 men out of 46)</p> <p>Sex (non-VL cases): 42% men (28 men out of 67)</p> <p>Presenting signs and symptoms: clinical suspicion of VL, ie more than 14 days of fever or splenomegaly or cytopenia</p> <p>Previous VL diagnosis: 24 out of 113 participants had a previous history of VL</p> <p>Frequency of VL: 37%</p> <p>HIV: 100%</p> <p>Clinical setting: HIV cohort in reference hospital in Belo Horizonte (Eduardo de Menezes Hospital, Fundação Hospitalar do Estado de Minas Gerais)</p> <p>Country: Brazil</p> <p>Endemic <i>Leishmania</i> species: <i>L. infantum</i></p>
Index tests	Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; sample: serum
Target condition and reference standard(s)	<p>Target condition: clinical VL</p> <p>Samples: bone marrow aspirate</p> <p>Techniques: direct smear and culture</p> <p>Reference standard category: parasitology not including spleen aspirate - no serology</p>
Flow and timing	Two patients were not included in the analysis because bone marrow aspiration was not done
Comparative	
Notes	Another estimate of Se and Sp based on latent class analysis in the same study population was not included because there was no information about the assessment of the conditional independence assumption (explained in the Criteria for considering studies for this review and in Appendix 1).

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Cota 2013 - composite 2 (Continued)

		High
DOMAIN 2: Index Test All tests		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
		Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	No	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
		Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard?	Yes	

Diro 2007

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL

Patient characteristics and setting	<p>Sample size: 38 Age: median 25 years; range 16-49. No children \leq 12 years Sex: 95% men Presenting signs and symptoms: fever of 2 weeks or more, and splenomegaly or lymphadenopathy or both Pregnant women and people with a previous history of VL treatment were not included. Patients with a thick-film-positive malaria episode were excluded Frequency of VL: 61% HIV: not tested Clinical setting: university hospital (Kahsay Abera Hospital in Humera and Gondar College of Medical Sciences) Country: Ethiopia Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; Sample: serum Type: Latex agglutination test in urine; Brand: KAtex, Kalon Biologicals Ltd, Guildford, UK; Sample: fresh urine</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL Sample: spleen or lymph node aspirate Technique: parasitology: culture or direct smear (Giemsa stain) Reference standard category: parasitology including spleen aspirate - no serology</p>		
Flow and timing			
Comparative			
Notes	Findings in context of active case detection were not included in this review. Same study as Boelaert 2008 - Ethiopia but different analysis.		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low

Diro 2007 (Continued)

DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard?	Yes		

Hailu 2006

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 103</p> <p>Age: not reported</p> <p>Sex: not reported</p> <p>Presenting signs and symptoms: fever for 2 weeks or more and splenomegaly</p>

	Not clear whether VL relapse suspects were excluded Frequency of VL: 51% (45/89) HIV: not tested Clinical setting: health centres in rural northern Ethiopia (Humera and Gondar) Country: Ethiopia Endemic <i>Leishmania</i> species: <i>L. donovani</i>		
Index tests	Type: FAST; Brand: not applicable; Sample: not reported		
Target condition and reference standard(s)	Target condition: clinical VL Sample: lymph node or spleen Technique: direct smear (Giemsa stain) or culture Reference standard category: parasitology including spleen aspirate - no serology		
Flow and timing	Fourteen patients were excluded, probably because the FAST results could not be read under field conditions		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			

Hailu 2006 (Continued)

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Low			
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard?	Yes		

Kilic 2008

Study characteristics	
Patient sampling	Consecutive enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 59</p> <p>Age (data available for 24 VL patients): median age 7 years; 22/24 cases <14 years</p> <p>Sex: 71% men</p> <p>Presenting signs and symptoms: prolonged fever + splenomegaly or hepatomegaly + anaemia/pancytopenia</p> <p>Frequency of VL: 41%</p> <p>HIV: not reported</p> <p>Clinical setting: tertiary care centre</p> <p>Country: Turkey</p> <p>Endemic <i>Leishmania</i> species: <i>L. infantum</i></p>
Index tests	Type: FAST; Brand: not applicable; Sample: serum

Kilic 2008 (Continued)

Target condition and reference standard(s)	Target condition: clinical VL Sample: bone marrow Technique: direct smear (Giemsa stain) and culture Reference standard category: parasitology not including spleen aspirate - no serology		
Flow and timing	No information about exclusion		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Unclear
Did all patients receive a reference standard?	Yes

Machado de Assis 2011

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 332</p> <p>Age: median 13 years; range 1 month - 76 years</p> <p>Sex: 58% men</p> <p>Presenting signs and symptoms: fever + coming from endemic area + one of the following: splenomegaly, hepatomegaly, anaemia, leucopenia, and thrombocytopenia</p> <p>Frequency of VL: 64%</p> <p>HIV: people with known immunodeficiency were excluded</p> <p>Clinical setting: two university hospitals and two research centres in four different states</p> <p>Country: Brazil</p> <p>Endemic <i>Leishmania</i> species: <i>L. infantum</i></p>
Index tests	Type: rK39 immunochromatographic test; Brand: IT-LEISH, DiaMed Latino-America S.A, Cressier sur Morat, Switzerland; Sample: blood
Target condition and reference standard(s)	<p>Target condition: clinical VL</p> <p>Sample: bone marrow</p> <p>Technique: culture or direct smear (Giemsa stain); at least 2 smears per patient</p> <p>Definition of VL: positive parasitological test</p> <p>Definition of non-VL: negative parasitological test and firm diagnosis of another disease</p> <p>Reference standard category: parasitology not including spleen aspirate - no serology</p>
Flow and timing	People with missing data or uncertain diagnosis were not mentioned
Comparative	

Notes	Same study as Machado de Assis 2012 but different analysis. Same study as Peruhype-Magalhaes 2012 but different brand of index test.		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Machado de Assis 2011 (Continued)

Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard?	Yes		

Machado de Assis 2012

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 404 Age: median 13 years; range 1 month - 77 years; standard deviation 17 years Sex: 58% men Presenting signs and symptoms: fever + coming from endemic area + one of the following: splenomegaly, hepatomegaly, anaemia, leucopenia, and thrombocytopenia Frequency of VL: 67% HIV: people with known immunodeficiency were excluded Clinical setting: two university hospitals and two research centres in four different states Country: Brazil Endemic <i>Leishmania</i> species: <i>L. infantum</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: IT-LEISH, DiaMed Latino-America S.A, Cressier sur Morat, Switzerland; Sample: finger prick sample of blood</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL Sample: bone marrow Approach: latent class analysis Tests included in latent class analysis: rK39 immunochromatographic test, rK39-ELISA, IFAT, DAT, microscopy of bone marrow sample</p>		
Flow and timing	People with missing data or inconclusive test results were not mentioned		
Comparative			
Notes	Same study as Machado de Assis 2011 but different analysis. Same study as Peruhype-Magalhaes 2012 but different brand of index test.		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Machado de Assis 2012 (Continued)

Was a consecutive or random sample of patients enrolled?	Yes			
Was a case-control design avoided?	Unclear			
Did the study avoid inappropriate exclusions?	Unclear			
				Low
DOMAIN 2: Index Test All tests				
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear			
				Low
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Unclear			
Did all patients receive a reference standard?	Yes			

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL \geq 5 years with suspected VL and presenting at outpatient department		
Patient characteristics and setting	<p>Sample size: 249 (total number of participants)</p> <p>Age (VL cases): median 16 years (interquartile range 10-25 years)</p> <p>Age (non-VL cases): median 16.5 years (interquartile range 11-25 years)</p> <p>Sex (VL cases): 66.4% men</p> <p>Sex (non-VL cases): 79.6% men</p> <p>Presenting signs and symptoms: fever for 2 weeks or more + clinical splenomegaly + malaria ruled out by a negative rapid test (Paracheck)</p> <p>Frequency of VL: 59.8%</p> <p>HIV: 0.5% (195 participants were tested; 1 non-VL case was HIV positive)</p> <p>Clinical setting: Kimalel Health Centre, Baringo district and Kacheliba Kala-azar Treatment Centre, Pokot North district</p> <p>Country: Kenya</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: DiaMed-IT LEISH, DiaMed AG, Switzerland; Sample: serum</p> <p>Type: rKE16 immunochromatographic test; Brand: Signal KA, Span Diagnostics Ltd, India; Sample: serum</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL</p> <p>Sample: spleen aspirate</p> <p>Technique: direct smear examination (Giemsa stain)</p> <p>Definition of VL: positive splenic aspirate</p> <p>Definition of non-VL: negative splenic aspirate</p> <p>Reference standard category: parasitology including spleen aspirate - no serology</p>		
Flow and timing	Thirty participants were excluded: 27 had contra-indications for spleen aspiration and 3 had an inconclusive smear reading		
Comparative			
Notes	Unclear if study population includes people with previous history of VL		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		

Did the study avoid inappropriate exclusions?	Yes			
				Low
DOMAIN 2: Index Test All tests				
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes			
				Low
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Peruhype-Magalhaes 2012

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 278 Age: mean 12 years; range 1 month - 77 years Sex: 58% (161/278) men Presenting signs and symptoms: fever + coming from endemic area + one of the following: splenomegaly, hepatomegaly, anaemia, leucopenia, and thrombocytopenia Frequency of VL: 69% (193/278) HIV: people with known immunodeficiency were excluded Clinical setting: two university hospitals and two research centres in four different states Country: Brazil Endemic <i>Leishmania</i> species: <i>L. infantum</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: Kalazar Detect, InBios International, Washington, USA; Sample: stored serum</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL Sample: bone marrow Technique: culture or direct smear (Giemsa stain); at least 2 smears per patient Definition of VL: positive parasitological test Definition of non-VL: negative parasitological test and firm diagnosis of another disease Reference standard category: parasitology not including spleen aspirate - no serology</p>		
Flow and timing	Unclear why there are 54 participants less than in Machado de Assis 2011		
Comparative			
Notes	Same study population as Machado de Assis 2011 and Machado de Assis 2012 . Evaluation of a new index test on stored serum samples.		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
			Low

DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard?	Yes		

Rijal 2004

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	Sample size: 269 Age (155 VL patients): median age 23 (interquartile range 13-26) Age (77 non-VL patients): median age 20 (interquartile range 10-30) Sex: not reported

	<p>Presenting signs and symptoms: fever for 2 weeks or more and clinical splenomegaly Frequency of VL: 67% HIV: done but not reported; at least 3 HIV-positive participants among non-VL patients Clinical setting: tertiary care centre (B.P. Koirala Institute of Health Sciences in Dharan) Country: Nepal Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	Type: latex agglutination test in urine; Brand: KAtex, Kalon Biological Ltd, Aldershot, UK; Sample: frozen urine		
Target condition and reference standard(s)	<p>Target condition: clinical VL Combination of parasitology (direct smear, Giemsa stain) of bone marrow or spleen aspirate and serology (DAT) Definition of VL: parasitology positive Definition of non-VL: parasitology negative and DAT \leq 1:3200 Reference standard category: combination of parasitology and serology</p>		
Flow and timing	Thirty-seven patients were not included: the urine samples of 8 patients were lost; 1 person left against medical advice; and 28 had an uncertain diagnosis (parasitology negative and DAT positive)		
Comparative			
Notes	Not clear whether VL relapse suspects were excluded		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
			Low

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard?	Yes		

Ritmeijer 2006

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 356</p> <p>Age: not reported</p> <p>Sex: not reported</p> <p>Presenting signs and symptoms: fever for more than 2 weeks + malaria ruled out + wasting + splenomegaly or lymphadenopathy</p> <p>Frequency of VL: 67% (228/341)</p> <p>HIV: not reported</p> <p>Clinical setting: three Kala-azar treatment centres in eastern and southern Sudan, supported by Médecins sans frontières</p> <p>Country: Sudan</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>

Index tests	Type: rK39 immunochromatographic test; Brand: DiaMed-IT Leish, DiaMed AG, Cressier sur Morat, Switzerland; Sample: blood		
Target condition and reference standard(s)	Target condition: clinical VL Combination of parasitology (direct smear, Giemsa stain) of lymph node or spleen aspirate sample and serology (DAT) Definition of VL: parasitology positive or DAT positive $\geq 1:6400$ Definition of non-VL: DAT $\leq 1/400$ or borderline DAT with negative aspirate Reference standard category: combination of parasitology and serology		
Flow and timing	Fifteen patients were excluded because of an uncertain diagnosis		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Ritmeijer 2006 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear			
Low				
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			
Did all patients receive a reference standard?	Yes			

Sundar 1998

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 323</p> <p>Age: mean age of VL patients: 25 years (standard error 1); among non-VL patients, age is reported in subgroups: 73 patients with documented other infections: mean age 26 (standard error 2); 83 patients with presumed other infections: mean age 23 (standard error 2); and 40 patients without final diagnosis: mean age 25 years (standard error 2)</p> <p>Sex: 69% men</p> <p>Frequency of VL: 39%</p> <p>Presenting signs and symptoms: persistent fever with or without splenomegaly and suspected VL. Some patients had been empirically treated with antimalarial or antibacterial agents before referral, without response</p> <p>HIV: not reported</p> <p>Clinical setting: kala-azar units in Varanasi and Muzaffarpur, linked to research centre</p> <p>Country: India</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>
Index tests	Type: rK39 immunochromatographic test; Brand: not applicable; Sample: blood

Sundar 1998 (Continued)

Target condition and reference standard(s)	Target condition: clinical VL Combination of parasitology (direct smear Giemsa stain) of spleen aspirate sample and diagnosis of other diseases Reference standard category: parasitology including spleen aspirate - no serology		
Flow and timing	Reference standard was conditional on result of index test.		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
			Low

Sundar 1998 (Continued)

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Did all patients receive a reference standard?	No

Sundar 2007

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 352 Age: median 15 years; interquartile range 2-65 Sex: 55% men Presenting signs and symptoms: fever for 2 weeks or more and splenomegaly No children < 1 year old; no pregnant women; no people with past kala-azar history Frequency of VL: 80% HIV: known HIV infection exclusion criterion Clinical setting: two research centres, in Muzaffarpur and Patna Country: India Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>
Index tests	<p>Type: rK39 immunochromatographic test; Brand: InBios International, Washington, USA; Sample: serum or blood Type: rK26 immunochromatographic test; Brand: InBios International, Washington, USA; Sample: serum or blood Type: latex agglutination test in urine; Brand: KAtex, Kalon Biologicals, Aldershot, UK; Sample: fresh urine</p>
Target condition and reference standard(s)	<p>Target condition: clinical VL Combination of parasitology (direct smear, Giemsa stain) of spleen aspirate sample and serology (DAT) and clinical diagnosis / follow-up / response to therapy Definition of VL: (1) spleen parasitology positive; or (2) clinical diagnosis and DAT positive \geq 1:3200 and response to anti-leishmanial therapy Definition of non-VL: spleen parasitology negative and alternative diagnosis and no VL during follow-up of six months</p>

	Reference standard category: combination of parasitology and serology		
Flow and timing	No withdrawals explained.		
Comparative			
Notes	Same study as Boelaert 2008 - India but different analysis and additional test (rK26 immunochromatographic test). Group of 100 healthy endemic controls not included in this review		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
			Low
DOMAIN 4: Flow and Timing			

Sundar 2007 (Continued)

Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard?	Yes		

ter Horst 2009 - HIV neg

Study characteristics	
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 179</p> <p>Age: reported for larger population (n = 699): mean age is 25,4 (standard deviation 10,1)</p> <p>Sex: reported for larger population (n = 699): ratio men:women is 15:1</p> <p>Presenting signs and symptoms: clinical VL case definition (not specified)</p> <p>No patients with previous history of VL treatment.</p> <p>Frequency of VL: 94%</p> <p>HIV: all HIV negative</p> <p>Clinical setting: kala-azar treatment centre supported by Médecins sans frontières (Kahsay Abera Hospital in Humera)</p> <p>Country: Ethiopia</p> <p>Endemic <i>Leishmania species:</i> <i>L. donovani</i></p>
Index tests	Type: rK39 immunochromatographic test; Brand: DiaMed AG, Switzerland; Sample: blood
Target condition and reference standard(s)	<p>Target condition: clinical VL</p> <p>Combination of parasitology (direct smear, Giemsa stain) of spleen aspirate sample and serology (DAT)</p> <p>Definition of VL: (1) spleen parasitology positive; or (2) DAT positive \geq 1:3200</p> <p>Definition of non-VL: (1) DAT \leq 1:400; or (2) DAT \leq 1:3200 and spleen parasitology negative</p> <p>Reference standard category: combination of parasitology and serology</p>
Flow and timing	Many patients with unknown HIV status were not included.
Comparative	
Notes	

Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Unclear		

ter Horst 2009 - HIV neg (Continued)

Were all patients included in the analysis?	No		
Did all patients receive a reference standard?	Yes		

ter Horst 2009 - HIV pos

Study characteristics			
Patient sampling	Consecutive and prospective enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 71 Age: reported for larger population (n = 699): mean age is 25.4 (standard deviation 10.1) Sex: reported for larger population (n = 699): ratio men:women is 15:1 Presenting signs and symptoms: clinical VL case definition (not specified) No patients with previous history of VL treatment. Frequency of VL: 92% HIV: all HIV positive Clinical setting: kala-azar treatment centre supported by Médecins sans frontières (Kahsay Abera Hospital in Humera) Country: Ethiopia Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	Type: rK39 immunochromatographic test; Brand: DiaMed AG, Switzerland; Sample: blood		
Target condition and reference standard(s)	<p>Target condition: clinical VL Combination of parasitology (direct smear, Giemsa stain) of spleen aspirate sample and serology (DAT) Definition of VL: (1) spleen parasitology positive; or (2) DAT positive \geq 1:3200 Definition of non-VL: (1) DAT \leq 1:400; or (2) DAT \leq 1:3200 and spleen parasitology negative Reference standard category: combination of parasitology and serology</p>		
Flow and timing	Many patients with unknown HIV status were not included.		
Comparative			
Notes	Not included in meta-analysis because this is the only evaluation of the rK39 immunochromatographic test in a population of HIV-positive patients		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

ter Horst 2009 - HIV pos (Continued)

Was a consecutive or random sample of patients enrolled?	Yes			
Was a case-control design avoided?	Yes			
Did the study avoid inappropriate exclusions?	Yes			
				Low
DOMAIN 2: Index Test All tests				
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes			
				Low
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Unclear			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Unclear			
Were all patients included in the analysis?	No			
Did all patients receive a reference standard?	Yes			

Study characteristics			
Patient sampling	Consecutive enrolment of patients with suspected VL		
Patient characteristics and setting	<p>Sample size: 58</p> <p>Age: only reported for 13 cases with all tests positive: 12 children between 1 and 15 years and 1 adult of 48 years</p> <p>Sex: not reported</p> <p>Frequency of VL: 28%</p> <p>Presenting signs and symptoms: ≥ 1 of the following signs: fever, anaemia, thrombocytopenia, leucopenia, splenomegaly, hepatomegaly, and weight loss</p> <p>HIV: not reported</p> <p>Clinical setting: various regional hospitals</p> <p>Country: Turkey</p> <p>Endemic <i>Leishmania</i> species: <i>L. infantum</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: InBios International, Washington, USA; Sample: serum or plasma</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL</p> <p>Combination of parasitology (direct smear, Giemsa stain) of lymph node or bone marrow aspirate sample and serology (IFAT)</p> <p>Definition of VL: (1) lymph node or bone marrow parasitology positive; or (2) IFAT positive ≥ 1: 128</p> <p>Definition of non-VL: parasitology and IFAT negative</p> <p>Reference standard category: combination of parasitology and serology</p>		
Flow and timing	No information about excluded patients.		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Low

DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard?	Yes		

Veeken 2003 - composite

Study characteristics	
Patient sampling	Retrospective selection of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 77</p> <p>Age: median 11; range 4-66 years</p> <p>Sex: 53% men</p> <p>Presenting signs and symptoms: fever, and splenomegaly or wasting</p>

Veeken 2003 - composite (Continued)

	<p>Frequency of VL: 70% HIV: not reported Clinical setting: kala-azar treatment centre at Um el Kher supported by Médecins sans frontières Country: Sudan Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>		
Index tests	<p>Type: rK39 immunochromatographic test; Brand: Amrad ICT, Brookdale, Australia; Sample: serum</p>		
Target condition and reference standard(s)	<p>Target condition: clinical VL Combination of parasitology (direct smear, Giemsa stain) of spleen aspirate sample and serology (DAT) Definition of VL: spleen parasitology positive or DAT positive \geq 1:6400 Definition of non-VL: all other patients Reference standard category: combination of parasitology and serology</p>		
Flow and timing	<p>Only people with stored serum and available results of DAT and parasitological test of spleen aspirate were included. According to the protocol by Médecins sans frontières, splenic aspiration was done in patients with suspected VL who were critically ill and in those with borderline DAT titres</p>		
Comparative			
Notes	<p>Same study as Veeken 2003 - spleen but different analysis.</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low

Veeken 2003 - composite (Continued)

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard?	Unclear		

Veeken 2003 - spleen

Study characteristics	
Patient sampling	Retrospective selection of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 77</p> <p>Age: median 11; range 4-66 years</p> <p>Sex: 53% men</p> <p>Presenting signs and symptoms: fever, and splenomegaly or wasting</p> <p>Frequency of VL: 65%</p> <p>HIV: not reported</p> <p>Clinical setting: kala-azar treatment centre at Um el Kher supported by Médecins sans frontières</p> <p>Country: Sudan</p> <p>Endemic <i>Leishmania</i> species: <i>L. donovani</i></p>
Index tests	Type: rK39 immunochromatographic test; Brand: Amrad ICT, Brookdale, Australia; Sample: serum

Veeken 2003 - spleen (Continued)

Target condition and reference standard(s)	Target condition: clinical VL Sample: spleen Technique: parasitology (direct smear, Giemsa stain) Reference standard category: parasitology including spleen aspirate - no serology		
Flow and timing	Only people with stored serum and available results of DAT and parasitological test of spleen aspirate were included. According to the protocol by Médecins sans frontières, splenic aspiration was done in patients with suspected VL who were critically ill and in those with borderline DAT titres		
Comparative			
Notes	Same study as Veeken 2003 - composite but different analysis.		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Veeken 2003 - spleen (Continued)

			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard?	Unclear		

Vilaplana 2004

Study characteristics	
Patient sampling	Consecutive enrolment of patients with suspected VL
Patient characteristics and setting	<p>Sample size: 89 Age: not reported Sex: not reported Presenting signs and symptoms: not specified (suggestive clinical signs and symptoms of VL) Frequency of VL: 14% (12/85) HIV: 100% of cases and 91% of non-cases Clinical setting: not reported Country: Spain Endemic <i>Leishmania species:</i> <i>L. infantum</i></p>
Index tests	Type: latex agglutination test in urine; Brand: KAtex, Kalon Biological Ltd, Aldershot, UK
Target condition and reference standard(s)	<p>Target condition: clinical VL Sample: bone marrow Technique: parasitology: culture or direct smear (Giemsa stain) Reference standard category: parasitology not including spleen aspirate - no serology</p>
Flow and timing	Three patients are not included in the analysis
Comparative	

Vilaplana 2004 (Continued)

Notes	Not included in meta-analysis because only evaluation of latex agglutination test in population with mainly HIV-positive patients		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Vilaplana 2004 (Continued)

Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard?	Yes		

DAT: direct agglutination test

ELISA: enzyme-linked immuno-sorbent assay

FAST: fast agglutination screening test

IFAT: indirect fluorescence antibody test

LCA: latent class analysis

PCR: polymerase chain reaction

Se: sensitivity

Sp: specificity

VL: visceral leishmaniasis

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Abeijon 2013	index test is not a rapid test
Ahsan 2010	not phase III diagnostic accuracy study
Akhoundi 2010	not phase III diagnostic accuracy study
Al-Nahhas 2003	target condition is not clinical visceral leishmaniasis
Al-Nahhas 2008	not phase III diagnostic accuracy study
Alam 2008	not phase III diagnostic accuracy study
Alborzi 2006	not phase III diagnostic accuracy study
Amato 2009	not phase III diagnostic accuracy study
Arya 1997	not original research article
Arya 2001	not original research article

(Continued)

Attar 2001	not phase III diagnostic accuracy study
Azazy 2004	not phase III diagnostic accuracy study
Bagchi 1998	not phase III diagnostic accuracy study
Bern 2000	not phase III diagnostic accuracy study
Brandonisio 2002	not phase III diagnostic accuracy study
Canavate 2011	not phase III diagnostic accuracy study
Carvalho 2003	not phase III diagnostic accuracy study
Cruz 2006	not phase III diagnostic accuracy study
de Assis 2008	publication of the same study in more than one record
Delgado 2001	not phase III diagnostic accuracy study
Edrissian 2003	not phase III diagnostic accuracy study
El-Safi 2003	not phase III diagnostic accuracy study
Ferreira Dourado 2007	not original research article
Gavгани 2008	not phase III diagnostic accuracy study
Goswami 2003	not phase III diagnostic accuracy study
Goswami 2007	not phase III diagnostic accuracy study
Goswami 2012	not phase III diagnostic accuracy study
Gupta 1994	not phase III diagnostic accuracy study
Hatam 2009	not phase III diagnostic accuracy study
Hommel 2001	not original research article
Iqbal 2002	not phase III diagnostic accuracy study
Jelinek 1999	target condition is not clinical visceral leishmaniasis
Khan 2009	not phase III diagnostic accuracy study

(Continued)

Khan 2010	not phase III diagnostic accuracy study
Kumar 2006	not phase III diagnostic accuracy study
Lemos 2003	not phase III diagnostic accuracy study
López Corbalán 2012	not phase III diagnostic accuracy study
Mandal 2008	reference standard is not according to criteria
Mansour 2009	not phase III diagnostic accuracy study
Marty 2007	not phase III diagnostic accuracy study
Mathur 2005	not phase III diagnostic accuracy study
Matlashewski 2013	reference standard is not according to criteria
Mbui 2011	publication of the same study in more than one record
Mohapatra 2010	not phase III diagnostic accuracy study
Monno 2009	not phase III diagnostic accuracy study
Moura 2013	reference standard is not according to criteria
Mueller 2013	not phase III diagnostic accuracy study
Ozerdem 2009	reference standard is not according to criteria
Ozkan 2008	not phase III diagnostic accuracy study
Pappas 1983	not phase III diagnostic accuracy study
Pappas 1984	not phase III diagnostic accuracy study
Pappas 1984a	not phase III diagnostic accuracy study
Pattabhi 2010	not phase III diagnostic accuracy study
Qu 1987	target condition is not clinical visceral leishmaniasis
Qu 2000	not phase III diagnostic accuracy study
Riera 2004	not phase III diagnostic accuracy study

(Continued)

Rouf 2009	not phase III diagnostic accuracy study
Saghrouni 2009	not phase III diagnostic accuracy study
Saha 2011	not phase III diagnostic accuracy study
Salam 2008	not phase III diagnostic accuracy study
Salam 2011	target condition is not clinical visceral leishmaniasis
Salotra 2005	not original research article
Sarkari 2005	not phase III diagnostic accuracy study
Sarker 2003	not phase III diagnostic accuracy study
Schallig 2002	not phase III diagnostic accuracy study
Schoone 2001	not phase III diagnostic accuracy study
Scott 1991	not phase III diagnostic accuracy study
Senaldi 1996	not phase III diagnostic accuracy study
Shamsuzzaman 2003	not phase III diagnostic accuracy study
Sharma 2009	target condition is not clinical visceral leishmaniasis
Silva 2005	not phase III diagnostic accuracy study
Singh 2009	not phase III diagnostic accuracy study
Singh 2010	not phase III diagnostic accuracy study
Sinha 2008	not phase III diagnostic accuracy study
Srivastava 1988	not phase III diagnostic accuracy study
Sundar 2002	reference standard is not according to criteria
Sundar 2002a	not phase III diagnostic accuracy study
Sundar 2002b	not original research article
Sundar 2005	not phase III diagnostic accuracy study

(Continued)

Sundar 2005a	publication of the same study in more than one record
Sundar 2006	not phase III diagnostic accuracy study
Sundar 2006a	not phase III diagnostic accuracy study
Teran-Angel 2010	not phase III diagnostic accuracy study
Walton 1986	not phase III diagnostic accuracy study
Walton 1987	publication of the same study in more than one record
Welch 2008	not phase III diagnostic accuracy study
Zijlstra 2001	not phase III diagnostic accuracy study

DATA

Presented below are all the data for all of the tests entered into the review.

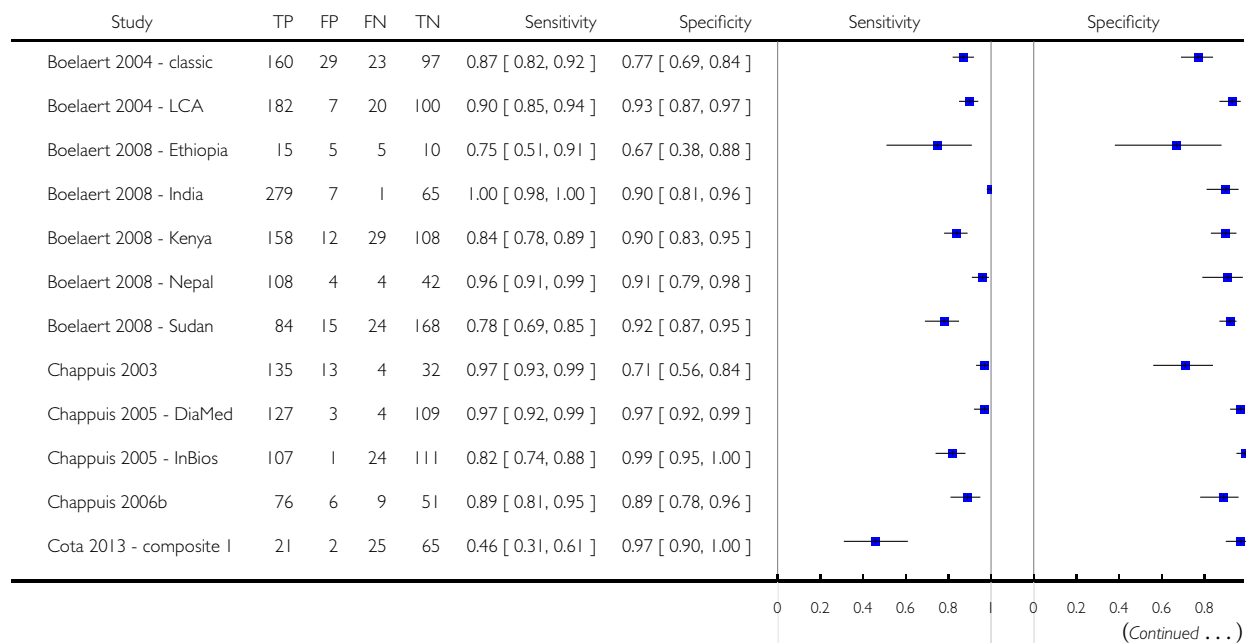
Tests. Data tables by test

Test	No. of studies	No. of participants
1 rK39 immunochromatographic test	26	5544
2 KAtex	9	1848
3 FAST	2	148
4 rK26 immunochromatographic test	1	352
5 rK39 Primary Analysis	16	3574
6 rKE16 immunochromatographic test	1	219

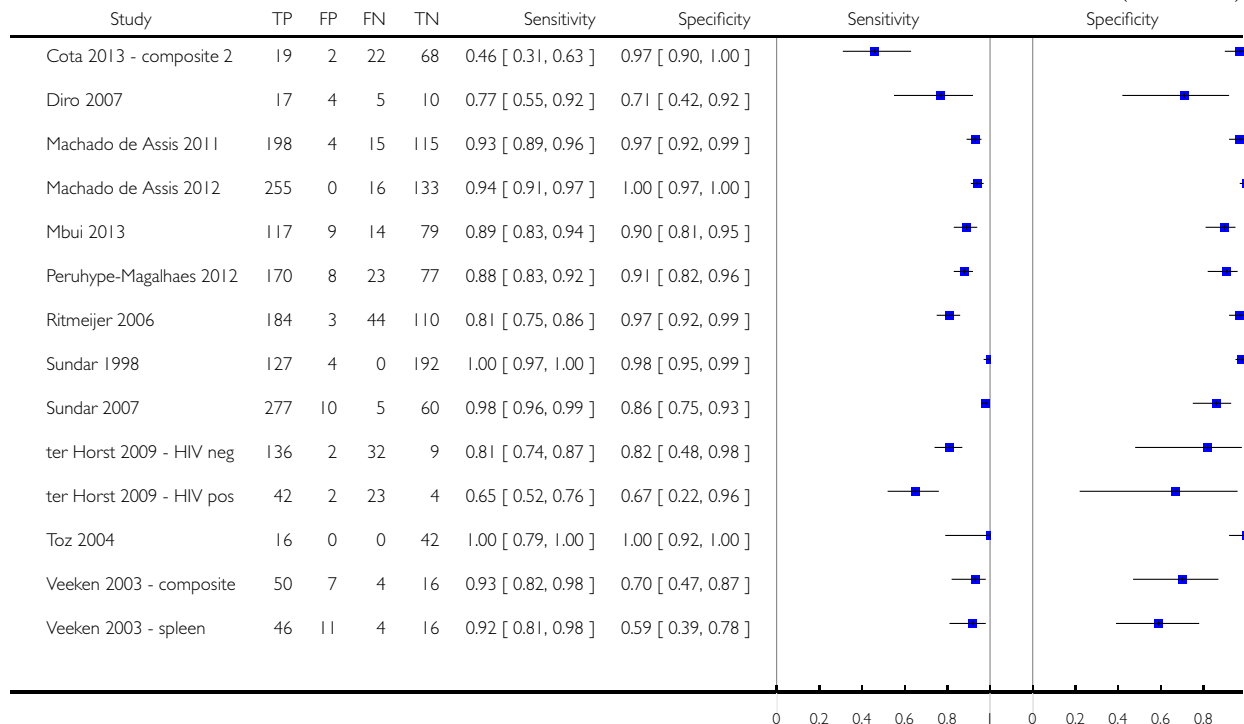
Test 1. rK39 immunochromatographic test.

Review: Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease

Test: 1 rK39 immunochromatographic test



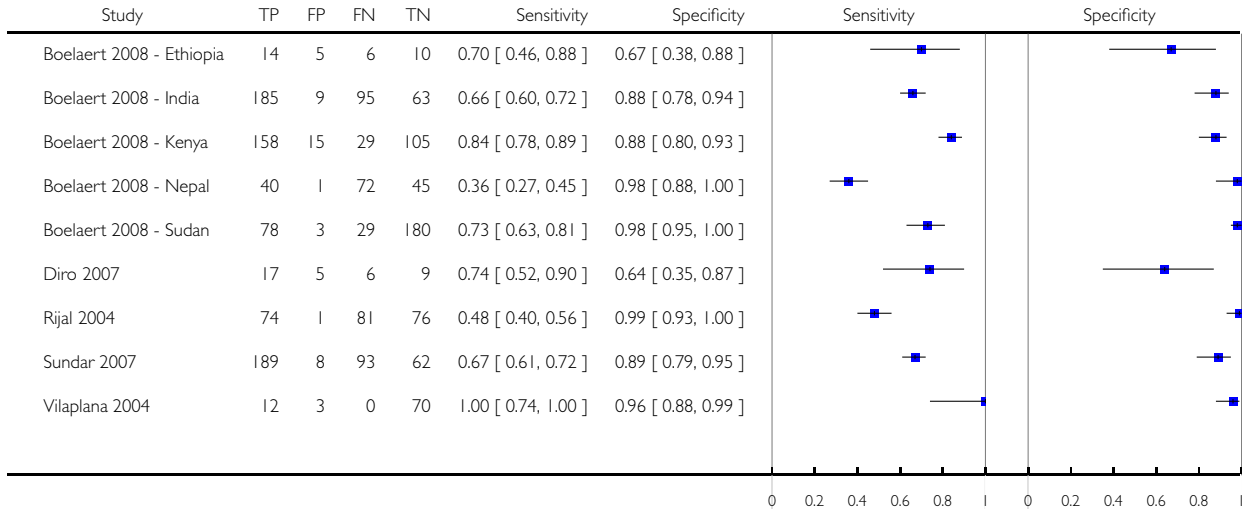
(... Continued)



Test 2. KAtex.

Review: Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease

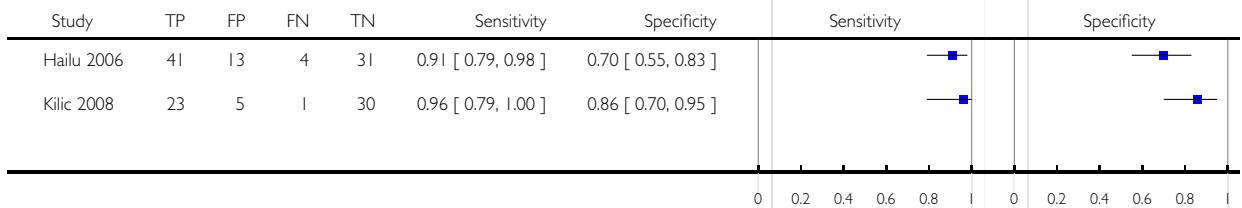
Test: 2 KAtex



Test 3. FAST.

Review: Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease

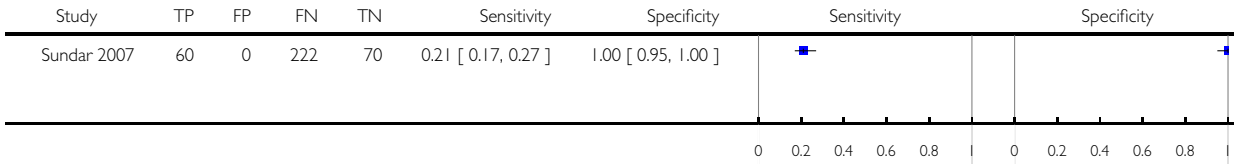
Test: 3 FAST



Test 4. rK26 immunochromatographic test.

Review: Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease

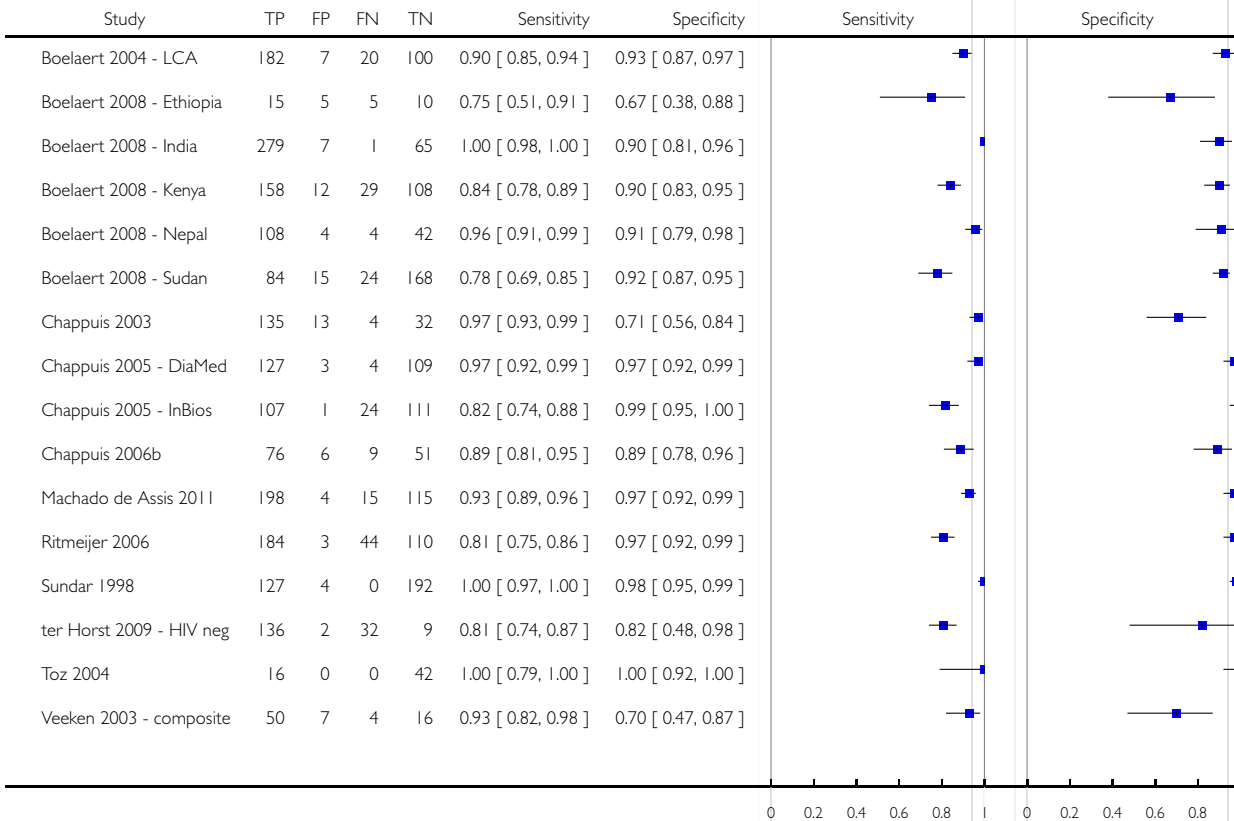
Test: 4 rK26 immunochromatographic test



Test 5. rK39 Primary Analysis.

Review: Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease

Test: 5 rK39 Primary Analysis



Test 6. rKE16 immunochromatographic test.

Review: Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease

Test: 6 rKE16 immunochromatographic test

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Mbui 2013	101	4	30	84	0.77 [0.69, 0.84]	0.95 [0.89, 0.99]		

ADDITIONAL TABLES

Table 1. Studies on rK39 ICT presenting two sets of Se and Sp estimates based on different reference standards: estimates included in primary analysis and sensitivity analysis

Study	Selected for primary analysis	Selected for sensitivity analysis
Boelaert 2004 - LCA & Boelaert 2004 - classic	LCA	Parasitology including spleen aspirate - no serology
Boelaert 2008 - Ethiopia & Diro 2007	LCA	Parasitology including spleen aspirate - no serology
Boelaert 2008 - India & Sundar 2007	LCA	Parasitology and serology
Cota 2013 - composite 1 & Cota 2013 - composite 2 *	Parasitology and serology	Parasitology not including spleen aspirate - no serology
Machado de Assis 2012 & Machado de Assis 2011	LCA	Parasitology not including spleen aspirate - no serology
Veeken 2003 - composite & Veeken 2003 - spleen	Parasitology and serology	Parasitology including spleen aspirate - no serology

* Cota 2013 describes HIV-infected patients only and is not included in the formal meta-analysis

Table 2. Overall Analysis Summary

	Number of studies	Se (95% CI)	Sp (95% CI)
rK39 immunochromatographic test			
Overall	18	91.9 (84.8 to 96.5)	92.4 (85.6 to 96.8)
Indian subcontinent	6	97.0 (90.0 to 99.5)	90.2 (76.1 to 97.7)
East Africa	9	85.3 (74.5 to 93.2)	91.1 (80.3 to 97.3)
Latex agglutination test			
Overall	6	63.6 (40.9 to 85.6)	92.9 (76.7 to 99.2)

Table 3. Heterogeneity assessment of the rK39 ICT meta-analysis

	Sensitivity		Specificity	
	Estimate	(95% CI)	Estimate	(95% CI)
Geographic region				
Indian subcontinent	97.0	(90.0 to 99.5)	90.2	(76.1 to 97.7)
East Africa	85.3	(74.5 to 93.2)	91.1	(80.3 to 97.3)
Commercial brand				
DiaMed	86.4	(70.7 to 95.9)	94.4	(80.5 to 99.2)
InBios	91.3	(83.8 to 96.3)	91.2	(83.1 to 96.3)
Other	97.8	(88.7 to 99.9)	92.5	(75.5 to 99.0)
Disease prevalence in sample				
Low (< 65%)	91.0	(82.2 to 96.9)	94.4	(87.2 to 98.3)
High (≥ 65%)	92.4	(84.5 to 97.4)	89.8	(79.7 to 95.8)
Study size				
Small (< 250)	91.5	(82.7 to 96.7)	89.8	(80.6 to 95.6)
Large (≥ 250)	92.3	(83.4 to 97.4)	94.5	(87.8 to 98.3)

Table 3. Heterogeneity assessment of the rK39 ICT meta-analysis (Continued)

QUADAS-2: risk of bias				
Low	89.5	(69.5 to 98.3)	96.0	(82.8 to 99.7)
Unclear	91.9	(83.7 to 97.1)	92.0	(84.0 to 96.7)
High	93.0	(77.9 to 99.1)	87.8	(70.7 to 97.2)
Reference standard				
Parasitology - no serology *	95.6	(84.3 to 99.6)	89.1	(72.4 to 97.4)
Parasitology and serology	89.6	(78.6 to 96.7)	94.6	(85.2 to 98.7)
Latent class analysis	91.0	(79.8 to 97.2)	91.5	(80.7 to 97.3)

* The categories “parasitology including spleen aspirate - no serology” and “parasitology not including spleen aspirate - no serology” were combined because the latter category contained only one study.

Table 4. Sensitivity analysis results of the rK39 ICT meta-analysis

	Sensitivity		Specificity	
	Estimate	(95% CI)	Estimate	(95% CI)
Main Analysis				
Indian subcontinent	97.0	(90.0 to 99.5)	90.2	(76.1 to 97.7)
East Africa	85.3	(74.5 to 93.2)	91.1	(80.3 to 97.3)
Alternative Analysis Set				
Indian subcontinent	96.1	(88.9 to 99.2)	86.7	(69.2 to 99.2)
East Africa	85.2	(75.0 to 92.7)	90.1	(77.0 to 97.4)
Bayesian analysis allowing for imperfect reference standards: expert priors - using main analysis set				
Indian subcontinent	97.3	(91.9 to 99.5)	93.7	(74.9 to 99.7)
East Africa	86.7	(77.5 to 94.0)	95.7	(84.3 to 99.8)
Bayesian analysis allowing for imperfect reference standards: vague priors - using main analysis set				

Table 4. Sensitivity analysis results of the rK39 ICT meta-analysis (Continued)

Indian subcontinent	97.3	(91.9 to 99.5)	93.0	(77.8 to 99.3)
East Africa	86.1	(77.2 to 93.3)	94.3	(83.8 to 99.6)

Table 5. Heterogeneity assessment of the Latex agglutination test meta-analysis

	Sensitivity		Specificity	
	Estimate	(95% CI)	Estimate	(95% CI)
Geographic region				
Indian subcontinent	50.8	(34.1 to 69.3)	95.3	(73.6 to 99.8)
East Africa	77.9	(58.2 to 92.3)	88.6	(59.5 to 92.3)
Study size				
Small (< 200)	52.9	(27.6 to 84.1)	86.2	(52.5 to 99.2)
Large (≥ 200)	68.5	(45.2 to 88.2)	94.5	(49.0 to 100.0)

Table 6. Sensitivity analysis results of the Latex agglutination test meta-analysis

	Sensitivity		Specificity	
	Estimate	(95% CI)	Estimate	(95% CI)
Main Analysis	63.6	(40.9 to 85.6)	92.9	(76.7 to 99.2)
Alternative Analysis Set	63.4	(40.8 to 85.4)	92.8	(76.3 to 99.2)

APPENDICES

Appendix I. Statistical methods

A short summary of some less common statistical methods used in this review is given below. A more extensive description of the model used to correct for imperfect reference standards can be found in [Menten 2013](#).

1) Latent class analysis (LCA)

LCA is a modelling technique that can be used in situations in which there is no good reference standard. It assumes that the true disease status in a study population is unknown (or latent). The LCA model estimates the sensitivity and specificity of a set of diagnostic tests (A, B, C, ...) on the basis of observed frequencies in test patterns (ABC+++, ABC++-, ABC+--, ...). As such, the LCA model provides a model-based estimate of the gold standard classification; ie the best way to group study participants in diseased or non-diseased.

The basic latent class model assumes that the observed variables are conditionally independent. This means that there should be no associations between the results of the diagnostic tests within each category of the latent variable (disease status). If this assumption does not hold, more advanced techniques (eg based on Bayesian statistical methodology) can be used. To be selected for this review, studies using LCA had to assess the conditional independence assumption between the diagnostic tests, and if conditional dependence was expected, they had to use appropriate statistical methods to take this into account. If a study was selected, the sensitivity and specificity estimates derived from the final LCA model were included in this review.

More information can be found in [Hui 1980](#), [Black 2002](#), [Branscum 2005](#), and [Baughman 2008](#) among other references.

2) The complementary Log-Log function

In the bivariate model a “link” function $g(y)$ is used to allow the use of the Normal (Gaussian) distribution to model the underlying study-specific sensitivity (Se) and specificity (Sp) of each study included in the meta-analysis. The standard link function used in the bivariate model is the logit link:

$$g(y) = \log(y/(1-y))$$

An alternative link function is the complementary log log (cloglog) link:

$$g(y) = \log(-\log(1-y))$$

Both link functions transform Se and Sp, which are in the interval [0,1], to any real number between minus infinity and plus infinity. The advantage of the cloglog link with our data is that it approaches infinity less quickly when y approaches 1 and consequently it mitigates the influence of studies that report 100% Se or Sp. This also reduces the inflation of the random-effects standard deviations as is apparent from comparison of Figure 6 and Figure 10. With the logit link, the prediction region extends to below the line of no diagnostic value ($Se + Sp < 1$), while the study which reports the lowest diagnostic value has $Se = 0.75$ and $Sp = 0.70$ (Figure 10). On the other hand, some observed data with high Se and Sp are not contained within the prediction region. The prediction region of the model with the cloglog link contains all observed data points, while not extending far beyond the studies with lowest observed Se and Sp (Figure 6). This is reflected in a lower deviance information criterion (DIC), a measure of model fit, for the cloglog model formulation compared to the logit formulation. The model with the lowest DIC shows the best fit to the data.

3) WinBUGS code for the primary model

WinBUGS is a statistical software for Bayesian analysis using Markov chain Monte Carlo methods. WinBUGS (or its recent open-source version OpenBUGS) provides a flexible Bayesian framework for model fitting. It can be used to fit both the bivariate and HSROC models.

Below is the code to fit the basic bivariate model, allowing for data from studies that use a reference standard and for studies that use latent class analysis.

Assuming there are $N=N1+N2$ studies:

- for the $N1$ studies that use a reference standard, the data extracted is:
 - $NDiseased[i]$ and $NNotDiseased[i]$: the (true) number of diseased and non-diseased subjects in study i
 - $TP[i]$, $TN[i]$: the number of true positives and true negatives in study i
- for the $N2$ studies that use latent class analysis, the data extracted is:
 - $Y1[i]$ and $Y2[i]$: the estimates of $g(Se)$ and $g(Sp)$ obtained from the results of the LCA reported in the publication for study i
 - $W1[i]$ and $W2[i]$: the estimates of the standard error of $Y1[i]$ and $Y2[i]$ for study i
 - with $g()$ the link function (logit or cloglog)

The code is as follows:

```
model{
```

```

# Binomials for Studies that Use a Reference Standard
# (where the reference standard is presumed to be perfect)
for(i in 1:N1){
TP[i] ~ dbin(SE[i],NDiseased[i])
TN[i] ~ dbin(SP[i],NNotDiseased[i])
}
# Normals for Studies that Use LCA
for(i in 1:N2){
Y1[i] ~ dnorm(alpha[i+N1,1],W1[i])
Y2[i] ~ dnorm(alpha[i+N1,2],W2[i])
}
# Bivariate normals for g(Se) and g(Sp)
for(i in 1:N){
# Implement the link function g()
# If logit is used
logit(SE[i]) <- alpha[i,1]
logit(SP[i]) <- alpha[i,2]
OR
# If cloglog is used
SE[i] <- 1-exp(-exp(alpha[i,1]))
SP[i] <- 1-exp(-exp(alpha[i,2]))
# Specify the bivariate normal
alpha[i,1:2] ~ dnorm(mu[,R[,])
}
# Specify Non-Informative Priors
# for means:
mu[1] ~ dnorm(0.0,.37)
mu[2] ~ dnorm(0.0,.37)
# for variance-covariance matrix
R[1:2,1:2] ~ dwish(Omega[,], 2) }
Note: Prior for Omega are provided as data:
Omega = matrix(c(.001,0,0,.001),nrow=2,byrow=T)

```

Appendix 2. Search strategy

Detailed search strategy

Search set	MEDLINE	EMBASE
1	Exp Leishmaniasis, visceral [MeSH]	Exp Visceral leishmaniasis [Emtree]
2	Exp Leishmania donovani [MeSH]	Exp Leishmania donovani [Emtree]
3	Exp Leishmania infantum [MeSH]	Exp Leishmania infantum [Emtree]
4	Kala azar OR kala-azar ti, ab	Exp Leishmania chagasi [Emtree]
5	Leishmania chagasi ti, ab	Kala azar OR kala-azar ti, ab
6	Visceral leishmania* ti, ab	Visceral leishmania* ti, ab

(Continued)

7	1-6/OR	1-6/OR
8	Rapid diagnostic test* ti, ab	Rapid diagnostic test* ti, ab
9	RDT ti, ab	RDT ti, ab
10	Antigen* detect* ti, ab	Antigen detection [Emtree]
11	Antibod* detect* ti, ab	Antibody detection [Emtree]
12	Latex Fixation Tests [MeSH]	Latex agglutination test [Emtree]
13	Lateral flow test ti, ab	Lateral flow test ti, ab
14	Enzyme-Linked Immunosorbent Assay [MeSH]	Latex fixation test* ti, ab
15	Serodiagnostic test* ti, ab	Enzyme-linked immunosorbent assay [Emtree]
16	ELISA ti, ab	ELISA ti, ab
17	Direct agglutination test* ti, ab	Agglutination test [Emtree]
18	Dipstick* ti, ab	Dipstick* ti, ab
19	K39 antigen, Leishmania [Substance Name]	Test strip [Emtree]
20	K26 antigen, Leishmania [Substance Name]	K39 Or rK39 ti, ab
21	K39 Or rK39 ti, ab	K26 or rK26 ti, ab
22	Strip test* ti, ab	Serodiagnosis [Emtree]
23	Reagent kits, diagnostic [MeSH]	Immunoblotting [Emtree]
24	Immunoblotting [MeSH]	8-12/OR
25	Serological tests [MeSH]	7 AND 24
26	8-25/OR	Limit 25 to humans
27	7 AND 26	
28	Limit 27 to humans	
27		
28		

Appendix 3. Interpretation of the clinical relevance of the findings of this review - predictive values and likelihood ratios

What follows is an interpretation of the clinical relevance of the findings of this review regarding the rK39 immunochromatographic test (ICT) when it is used to diagnose visceral leishmaniasis (VL) among patients with febrile splenomegaly and no previous history of VL.

Indian subcontinent

When the rK39 ICT is used in the Indian subcontinent, in a setting where the prior probability of VL among clinical suspects is 40%, which is typically seen in a peripheral health centre in an endemic area, the positive predictive value of the test is 87%. This means that out of 100 patients with a positive rK39 result, 87 would have VL (true positive result) and 13 would have another disease (false positive). The negative predictive value is 98%, meaning that out of 100 patients with a negative rK39 ICT result, 98 would have another disease (true negative) and 2 would have VL (false negative).

When the same test is used in a setting with a prior probability of VL of 60%, which is more typical for a referral centre in an endemic area, the positive predictive value is 94% and the negative predictive value is 95%.

A likelihood ratio is another way of expressing how informative a diagnostic test is: it indicates to what extent the rK39 ICT result changes the odds that a patient has VL. The likelihood ratio of a positive rK39 ICT result is 9.90, and the likelihood ratio of a negative test result is 0.03. This means that in the Indian subcontinent, a positive rK39 ICT result is a strong argument in favour of VL (ruling in) and that a negative rK39 ICT result is a strong argument against VL (ruling out).

East Africa

When the rK39 ICT is used in east Africa, in a setting where the prior probability of VL is 40%, which is typically seen in a peripheral health centre in an endemic area, the positive predictive value of the test is 86%. This means that out of 100 patients with a positive rK39 ICT result, 86 would have VL (true positive result) and 14 would have another disease (false positive). The negative predictive value is 90%, meaning that out of 100 patients with a negative rK39 ICT result, 90 would have another disease (true negative) and 10 would have VL (false negative).

When the same test is used in a setting with a prior probability of VL of 60%, which is more typical for a referral centre in an endemic area, the positive predictive value is 93% and the negative predictive value is 81%.

In east Africa, the likelihood ratio of a positive rK39 ICT result is 9.58, and the likelihood of a negative rK39 ICT result is 0.16. This means that a positive rK39 ICT result is strong argument in favour of VL (ruling in), and that a negative rK39 ICT result is not an absolute argument against VL (does not allow to rule out VL completely).

CONTRIBUTIONS OF AUTHORS

Marleen Boelaert, Joris Menten, François Chappuis and Suman Rijal performed previous work that was the foundation of the current study. Marleen Boelaert, Joris Menten and Temmy Sunyoto conceived and designed the review. Marleen Boelaert, Temmy Sunyoto, Kristien Verdonck and Johan van Griensven screened the retrieved papers against the inclusion criteria. Marleen Boelaert, Johan van Griensven and Kristien Verdonck extracted data and appraised the quality of the papers. Data management for the review was carried out by Joris Menten, Kristien Verdonck and Temmy Sunyoto. Joris Menten did the statistical analysis. Marleen Boelaert, Joris Menten, Temmy Sunyoto and Kristien Verdonck wrote the first draft of the review. All authors contributed to the final manuscript. Marleen Boelaert is the guarantor of the review.

DECLARATIONS OF INTEREST

Marleen Boelaert, Joris Menten, François Chappuis and Suman Rijal are authors of studies included in this review. There are no other known conflicts of interest.

SOURCES OF SUPPORT

Internal sources

- Institute of Tropical Medicine, Antwerp, Belgium.

External sources

- Belgian Development Cooperation, Belgium.
Grant to Temmy Sunyoto
- Department of Economy, Science and Innovation of the Flemish Government, Belgium.
Funding for the Clinical Trials Unit at the Institute of Tropical Medicine, Antwerp

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

For the assessment of the quality of the included records, we planned and started to work with QUADAS. When QUADAS-2 became available, we switched to this new tool ([Whiting 2011](#)). In the review, we only report the information obtained using QUADAS-2.

In order to avoid loss of information, we used a somewhat wider interpretation of the composite reference standard. For the sake of clarity, we slightly reformulated the accepted reference standards in the section [Criteria for considering studies for this review](#).

In the statistical modelling, we used a complementary log-log, rather than a logit link function in the bivariate model. The choice was based on an improved fit of the model to the data as described in the findings section.