VOLUME 17 NO 10 PP 1202–1207 OCTOBER 2012

Latent class analysis of diagnostic tests for visceral leishmaniasis in Brazil

Tália Santana Machado de Assis[1], Ana Rabello[1] and Guilherme Loureiro Werneck[2]

1 *Laboratório de Pesquisas Clínicas, Centro de Pesquisas René Rachou, Fundação Oswaldo Cruz (FIOCRUZ), Belo Horizonte, Minas Gerais, Brazil*
2 *Departamento de Epidemiologia, Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*

Abstract

OBJECTIVE To estimate the sensitivities and specificities of different diagnostic tests for visceral leishmaniasis (VL) using latent class analysis (LCA).

METHODS This study was performed using data from a prospective study conducted in four Brazilian states from May 2004 to May 2007. Five diagnostic tests for VL were evaluated in 285 VL cases and 119 non-cases: microscopy, indirect fluorescence antibody test (IFAT), enzyme-linked immunosorbent assay using recombinant K39 antigen (rK39-ELISA), direct agglutination test (DAT) and the rK39 rapid test.

RESULTS Microscopy showed sensitivity of 77.0% (CI: 71.5–81.5) and specificity of 99.0% (CI: 94.0–99.7). The IFAT and the DAT showed similar sensitivities, 88.3% (CI: 84.0–92.0) and 88.5% (CI: 84.1–92.0), respectively, but the DAT had a higher specificity (95.4%, CI: 89.2–98.1) than did the IFAT (83.0%, CI: 75.0–88.2). The rK39-ELISA and the rK39 rapid test showed sensitivities of 99.0% (CI: 96.3–99.6) and 94.0% (CI: 90.1–96.3), and specificities of 82.5% (CI: 75.0–88.3) and 100% (CI: 97.0–100.0%), respectively.

CONCLUSIONS Considering the lack of an adequate reference standard, LCA proved to be a useful tool in validating diagnostic methods for VL. The DAT and the rK39 rapid test showed better performance. Thus, clinically suspected cases of VL in a Brazilian endemic area could be treated based on the positivity of one of these tests.

keywords visceral leishmaniasis, diagnosis, latent class analysis

Introduction

Diagnostic methods for visceral leishmaniasis (VL) should be carefully validated, because a naïve evaluation may generate biased conclusions, particularly because of the lack of an appropriate reference standard. New tests are usually compared to existing imperfect ones, and their accuracy might seriously be underestimated or overestimated using such approach (Thibodeau 1981; Valenstein 1990).Current recommendations for a definitive diagnosis of VL rely on parasitological confirmation by means of invasive procedures, requiring infrastructure and professional expertise. Unfortunately, the sensitivity of bone marrow and lymph node aspirates is suboptimal, ranging from 53% to 86% (World Health Organization 2010).

Flawed estimates of test accuracy properties have a serious potential impact from the clinical point of view. False-positive results may lead to overtreatment, augmented financial cost, unnecessary exposure of individuals to the side effects of drugs and delay of treatment for other serious conditions. On the other hand, a false-negative result may extend suffering, delay appropriate treatment and aggravate prognosis. An alternative to the classical validation approach using parasitological diagnostic methods as the reference standard is latent class analysis (LCA) (Hui & Walter 1980; Rindskopf & Rindskopf 1986).

Latent class analysis is based on the concept that observed results of different imperfect tests for the same disease are influenced by a latent common variable, the true disease status, which cannot be directly measured. In basic LCA models, the observed variables are assumed to be conditionally independent. In a group of patients with unknown disease status, for whom results from several diagnostic tests are available, LCA will model the probability of each combination of test results on the latent class and will provide an estimate of sensitivity and specificity for each of the diagnostic tests evaluated (Hui & Walter 1980; Rindskopf & Rindskopf 1986).

Several studies have used LCA for the evaluation of diagnostic tests, such as Langhi Junior *et al.* (2002) and Andrade and Gontijo (2008) for Chagas' disease, Girardi *et al.* (2009) for tuberculosis, and Koukounari *et al.* (2009) for schistosomiasis. Boelaert *et al.* (1999, 2004, 2008), using LCA for the diagnosis of human VL, concluded that the model is a useful tool and provides more realistic estimates of the performance of diagnostic tests compared with the classical validation approach. However, these studies were developed in east Africa and in the Indian subcontinent where VL is caused by a different parasite species and presents different epidemiological features. Therefore, the purpose of this study was to apply LCA to estimate the sensitivity and specificity of five diagnostic tests for VL caused by *Leishmania (Leishmania) chagasi* (syn. *Leishmania infantum*) in Brazil.

Methods

The analysis was performed using data from a prospective multicentric study conducted in four Brazilian states (Maranhão, Piauí, Bahia and Minas Gerais) between May 2004 and May 2007 (Machado de Assis *et al.* 2008, 2011).

The following diagnostic tests were evaluated: (i) microscopy (bone marrow smears were stained with Giemsa and evaluated under a 1000× oil immersion lens on an optic microscope); (ii) indirect fluorescence antibody test (IFAT), performed with a commercial kit (Bio-Manguinhos, Rio de Janeiro, Brazil); (iii) enzyme-linked immunosorbent assay using recombinant K39 antigen (rK39-ELISA), performed according to Machado de Assis *et al.* (2008); (iv) direct agglutination test (DAT), performed as in Pedras *et al.* (2008); and (v) the rapid test (IT-LEISH® Diamed Latino-America S. A. - Cressier sur Morat, Switzerland) performed according to Machado de Assis *et al.* (2011). The Research Ethics Committee of the Centro de Pesquisas René Rachou, Fundação Oswaldo Cruz (CPqRR-FIOCRUZ) approved the study (CEPSH/CPqRRnº: 13/2003).

Data analysis

A database containing the epidemiological and clinical characteristics of all patients and the results of the laboratory tests was constructed using SPSS 12.0 software (SPSS Inc., Chicago, IL, USA). Five variables were included in the LCA: the results of microscopy, IFAT, rK39-ELISA, DAT and rapid test. LCA was performed using TAGS software implemented in R version 2.2 (R Development Core Team and R Foundation for Statistical Computing, 2005). In this study, we implemented the basic latent class model, using the assumption of conditional independence given the latent class. In basic LCA, there are no associations between the observed variables within each category of the latent variable. The latent variable is the true status on the disease, and the hypothesis is that there are two latent classes (presence or absence of VL). The fit of LCA model for the assumption of conditional independence was performed through the goodness-of-fit test followed by the evaluation of residual correlations between tests.

The serial reading was determined using the following formulas: Sensitivity OR rule = se A + (1 − se A) × se B and Specificity OR rule = sp A × sp B. The serial reading using OR rule considers that if the first test is positive, the diagnosis is positive; otherwise, the second test is performed. If the second test is positive after a negative first test, then the diagnosis also is positive; otherwise, the diagnosis is negative.

Results

A total of 404 patients with clinical suspicion for VL as defined by fever, accompanied by splenomegaly, hepatomegaly, anaemia, leukopenia or thrombocytopenia, were enrolled in the study. Of these patients, 285 had a firm diagnosis of VL; the diagnosis was reached by parasitological methods in 213 patients and a positive serological test and adequate response to treatment in 72 patients. The other 119 patients had a negative parasitological examination and confirmation of disease from another etiology. The non-cases were diagnosed with various diseases, such as leukemia, liver disease, schistosomiasis, ascariasis, liver fibrosis, lymphoma, rheumatoid arthritis, malaria, mononucleosis, typhoid fever, marrow aplasia, liver cirrhosis, meningitis, lupus erythematosus, encephalitis, tuberculosis, among others. The median age of the patients was 13 years (range: 1 month–76.8 years, standard deviation: 17 years) and 58% were male. The median time for symptoms of the patients was 56 days (range: 3–720 days, standard deviation: 86 days).

The test for evaluating the fit of the model with conditional independence (goodness-of-fit test) proved to be adjusted (P value = 0.06). The residuals correlations between tests were randomly distributed around 0 (rapid test and IFAT = 0.02, rapid test and microscopy = 0.05, rapid test and rK39-ELISA = 0.01, rapid test and DAT = −0.00, IFAT and microscopy = −0.01, IFAT and rK39-ELISA = −0.03, IFAT and DAT = 0.00, microscopy and rK39-ELISA = 0.00, microscopy and DAT = 0.00 e rK39-ELISA and DAT = −0.00).

The disease prevalence estimated by LCA was 67%. The parasitological test showed sensitivity of 77.0% (CI: 71.5–81.5) and specificity of 99.0% (CI: 94.0–99.7). The IFAT and the DAT showed sensitivities of 88.3% (CI: 84.0–92.0) and 88.5% (CI: 84.1–92.0), respectively, but

Table 1 Values of sensitivity and specificity of diagnostic methods for visceral leishmaniasis as estimated by basic latent class analysis

	Microscopy	IFAT	rK39-ELISA	DAT	rK39 Rapid test
Sensitivity (%) (95% CI)	77.0 (71.5–81.5)	88.3 (84.0–92.0)	99.0 (96.3–99.6)	88.5 (84.1–92.0)	94.0 (90.1–96.3)
Specificity (%) (95% CI)	99.0 (94.0–99.7)	83.0 (75.0–88.2)	82.5 (75.0–88.3)	95.4 (89.2–98.1)	100 (97.0–100.0)

Significant differences ($P \leq 0.05$): sensitivity of rapid test *vs.* all others tests evaluated, DAT *vs.* rK39-ELISA, DAT *vs.* microscopy, rK39-ELISA *vs.* IFAT, rK39-ELISA *vs.* microscopy, microscopy *vs.* IFAT and specificity of rapid test *vs.* rK39-ELISA, rapid test *vs.* IFAT, DAT *vs.* rK39-ELISA, DAT *vs.* IFAT, rK39-ELISA *vs.* microscopy, microscopy *vs.* IFAT.
No significant differences ($P > 0.05$): sensitivity of DAT *vs.* IFAT and specificity of rapid test *vs.* DAT, rapid test *vs.* microscopy, DAT *vs.* microscopy and rK39-ELISA *vs.* IFAT.

the specificity of the DAT was higher than the observed for IFAT (95.4%, CI: 89.2–98.1 *vs.* 83.0%, CI: 75.0–88.2).

The rK39-ELISA and the rK39 rapid test showed sensitivities of 99.0% (CI: 96.3–99.6) and 94.0% (CI: 90.1–96.3) and specificities of 82.5% (CI: 75.0–88.3) and 100% (CI: 97.0–100.0%), respectively (Table 1). Table 2 shows the frequencies of diagnostic test patterns. The difference of sensitivity of rapid test and all others tests evaluated, DAT *vs.* rK39-ELISA, DAT *vs.* microscopy, rK39-ELISA *vs.* IFAT, rK39-ELISA *vs.* microscopy, microscopy *vs.* IFAT and the difference of specificity of rapid test *vs.* rK39-ELISA, rapid test *vs.* IFAT, DAT *vs.* rK39-ELISA, DAT *vs.* IFAT, rK39-ELISA *vs.* microscopy, microscopy *vs.* IFAT, were significant ($P \leq 0.05$). DAT *vs.* IFAT showed similar sensitivity ($P > 0.05$) and rapid test *vs.* DAT, rapid test *vs.* microscopy, DAT *vs.* microscopy and rK39-ELISA *vs.* IFAT showed similar specificity ($P > 0.05$).

In the serial reading of diagnostic tests evaluated sensitivities equal or above 99.0% were reached. However, specificities equal or above 95% were obtained only by rapid test *vs.* DAT and rapid test *vs.* microscopy (Table 3).

Table 2 Observed frequencies of tests patterns as estimated by latent class analysis model

Rapid test	IFAT	Microscopy	rK39-ELISA	DAT	Observed frequency
0	0	0	0	0	87
0	0	0	0	1	1
0	0	0	1	0	15
0	0	0	1	1	3
0	0	1	0	0	1
0	0	1	0	1	1
0	1	0	0	0	15
0	1	0	0	1	2
0	1	0	1	0	6
0	1	0	1	1	3
0	1	1	1	1	11
0	1	1	1	0	1
1	0	0	1	1	7
1	0	1	0	0	1
1	0	1	1	0	3
1	0	1	1	1	20
1	1	0	1	0	11
1	1	0	1	1	41
1	1	1	0	1	2
1	1	1	1	0	15
1	1	1	1	1	158

Discussion

The diagnosis of VL is not a simple task, as it shares clinical features with other diseases; therefore, accurate laboratory diagnostic tests are essential. The current reference test for disease diagnosis is the microscopic demonstration of *Leishmania* spp. in spleen, bone marrow, lymph nodes or liver aspirates, but both the aspiration procedure and the reading of slides require a high level of expertise that makes them unsuitable for generalised field use. Diagnostic research in VL has been damaged by the lack of a perfect reference standard. The parasitological test is highly specific, but its sensitivity is influenced by the tissue sample, time and quality of the reading.

Because of the limitations of direct methods, several immunological tests have been evaluated. IFAT is the test utilised by the Brazilian Leishmaniasis Control Programme, with sensitivity and specificity values of 88–92% and 81–92%, respectively (Ministério da Saúde, 2006).

Table 3 Values of sensitivity and specificity of diagnostic methods using serial reading

	Serial reading	
Test combination	Sensitivity (%)	Specificity (%)
Rapid test/IFAT (95% CI)	99.3 (97.5–99.9)	83.0 (74.3–88.7)
Rapid test/DAT (95% CI)	99.3 (97.5–99.9)	95.4 (89.3–98.1)
Rapid test/rK39 ELISA (95% CI)	99.9 (98.1–100.0)	82.5 (74.3–88.7)
Rapid test/Microscopy (95% CI)	99.0 (96.9–99.8)	99.0 (95.4–99.9)

ELISA using rK39 antigen is considered a valuable tool and has estimates of sensitivity of 95–97% and specificity of 84–97% (Machado de Assis *et al.* 2008; Pedras *et al.* 2008). DAT is simple to perform, with sensitivity estimates of 95–99%, and specificity of 88–98% (Sundar *et al.* 2007; Pedras *et al.* 2008; Oliveira *et al.* 2009; Machado de Assis *et al.* 2011). Rapid tests are also simple to perform, do not require laboratory structure and have estimates of sensitivity and specificity varying from 67–100% and from 59–100%, respectively (Sundar *et al.* 1998; Zijlstra *et al.* 2001; Carvalho *et al.* 2003; Veeken *et al.* 2003; Machado de Assis *et al.* 2008).

Sheps and Schechter (1984) report that, in practice, very few real reference standards are available, and one-third of medical articles dealing with diagnostic test evaluation used no well-defined reference standard, and Guyatt *et al.* (1986) report that most new diagnostic technologies have not been assessed adequately to determine whether their application improves public health. Therefore, research on this issue needs a better and more standardised validation methodology, and LCA has been suggested as a potential solution to the problem of imperfect reference standards (Hadgu & Qu 1998), although softwares for this purpose are not widely available (Pouillot *et al.* 2002).

The design of validation studies based on LCA is not necessarily much more expensive than the classical alternative, as a minimum of three tests and roughly 100 observations are required for a model of conditional independence (Boelaert *et al.* 1999). One nice feature is that LCA based on serological tests might provide good estimates of the sensitivity and specificity of tests, avoiding the discomfort of the bone marrow aspiration required to perform the parasitological test. Reviews of publications on diagnostics have shown that although the quality of diagnostic trials is improving, many are still lacking in rigour. Some common design problems are the evaluation in an inappropriate study group or in an inappropriate setting, small sample size and lack of an adequate standard test (Ransohoff & Feinstein 1978; Reid *et al.* 1995; Peeling *et al.* 2006).

In this study, LCA estimated sensitivity of 77% and specificity of 99% for the bone marrow aspirate. These results corroborate the data reported by Boelaert *et al.* (2004), where LCA estimated a sensitivity of 78.1% and a specificity of 94.8%. This strengthens the view that bone marrow aspirate cannot be considered a reference standard for the validation of diagnostic tests for VL and that complementary approaches such as LCA might be useful for studies of validation. Boelaert *et al.* (2007) recommends that in cases where spleen aspiration cannot be used, researchers can opt to use either a composite reference standard or LCA. Spleen aspirate is not recommended by the Brazilian Leishmaniasis Control Programme because of the high risk of severe accidents related to this procedure.

Latent class analysis estimated a sensitivity of 88.3% and a specificity of 83.0% for the IFAT. These findings contrast with those reported by Boelaert *et al.* (2004), analysing patients from Nepal where LCA estimated a sensitivity of 30.0% and a specificity of 98.3%; however, the findings corroborate the data presented by Machado de Assis *et al.* (2008), and Pedras *et al.* (2008), which reported sensitivities ranging from 88% to 92% and specificities ranging from 81% to 88% using classical validation approaches.

Latent class analysis estimated sensitivity of 99.0% and specificity of 82.5% for the rK39-ELISA. This is the first time that the performance of ELISA for VL has been assessed using LCA. The data presented here support those by Machado de Assis *et al.* (2008) and Pedras *et al.* (2008), which reported sensitivities ranging from 95% to 97% and specificities ranging from 84% to 97% for rK39 antigen, using classical validation.

In this study, the DAT showed sensitivity of 88.5% and specificity of 95.4%. The results of the sensitivity of DAT using LCA agree with those presented by Boelaert *et al.* (2008) for the Sudan (85.7%), however, disagree with those presented by Boelaert *et al.* (2004, 2008 for the Ethiopia, Kenya, India and Nepal) (range: 94–98.8%). The results of the specificity of DAT agree with those observed by Boelaert *et al.* (2004, 2008 for the Ethiopia, Sudan, India and Nepal) (range: 91–98.2%), but disagree with those reported by Boelaert *et al.* (2008 for the Kenya) (81.9).

The results of the sensitivity of the rK39 rapid test in this study (94.0%) are in agreement with those reported by Boelaert *et al.* (2004, 2008 for the India and Nepal), (range: 90.1–99.6%); however, they disagree with those presented by Boelaert *et al.* (2008 for the Ethiopia, Kenya and Sudan) (range: 75.4–84.7%). The results of the specificity of the rK39 rapid test in this study (100%) disagree with those presented by Boelaert *et al.* (2004, 2008) (range: 70–93%). Discrepancies between our findings and those of other investigators might be explained by possible differences in the test accuracy between subspecies of the *L. donovani* complex, by genetic differences in patients, by methodological differences between studies and the use of different brands of rapid tests and standardisation of DAT.

One way to improve the performance of diagnostic tests is to use serial reading. Usually, in the serial approach, a simpler and cheaper test is carried out first. Taking into account the performance of the tests evaluated, we recommend that the first test to be performed is a rapid test, which provides results within 20 min, followed, if

necessary, by DAT, which is a non-invasive test and requires minimal structure. In Brazil, the Ministry of Health has recently purchased rapid tests, and hopefully these will be increasingly available for the diagnosis of patients in health services. Studies on the cost effectiveness of such approaches should be conducted to analyse the feasibility of associations between diagnostics tests studied.

In conclusion, as described in other studies in east Africa and in the Indian subcontinent, LCA proved to be a useful tool for the validation of diagnostic methods for human VL caused by *L. infantum*. In the absence of an adequate reference standard, LCA gave consistent estimates of test characteristics. The DAT and the rK39 rapid test showed better performance and should be considered as strong tools to be used under supervised conditions by the Public Health System in Brazil.

Acknowledgement

This work was financially supported by the Secretary of Health Surveillance, Brazilian Ministry of Health, CNPq (National Counsel of Technological and Scientific Development) and the Oswaldo Cruz Foundation – FIOCRUZ.

References

Andrade AQ & Gontijo ED (2008) Triagem neonatal para infecção chagásica congênita: Aplicação de análise de classe latente para avaliação dos testes diagnósticos. *Revista da Sociedade Brasileira de Medicina Tropical* **41**, 615–620.

Boelaert M, Sayda ES, Goetghebeur E, Gomes-Pereira S, Le Ray D & Van der Stuyft P (1999) Latent class analysis permits unbiased estimates of the validity of DAT for the diagnosis of visceral leishmaniasis. *Tropical Medicine and International Health* **4**, 395–401.

Boelaert M, Rijal S, Regmi S *et al.* (2004) A comparative study of the effectiveness of diagnostic tests for visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* **70**, 72–77.

Boelaert M, Bhattacharya S, Chappuis F *et al.* (2007) Evaluation of rapid diagnostic tests: visceral leishmaniasis. *Nature Reviews* **5**, 30–39.

Boelaert M, El-Safi S, Hailu A *et al.* (2008) 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* **102**, 32–40.

Carvalho SFG, Lemos EM, Corey R & Dietze R (2003) Performance of recombinant K39 antigen in diagnosis of brazilian visceral leishmaniasis. *American Journal of Tropical Medicine and Hygiene* **68**, 321–324.

Girardi E, Angeletti C, Puro V *et al.* (2009) Estimating diagnostic accuracy of tests for latent tuberculosis infection without a gold standard among healthcare workers. *European Surveillance* **29**, 14.

Guyatt GH, Tugwell PX, Feeny DH, Haynes RB & Drummond M (1986) A framework for clinical evaluation of diagnostic technologies. *Canadian Medical Association Journal* **134**, 587–594.

Hadgu A & Qu Y (1998) A biomedical application of latent class models with random effects. *Applied Statistics* **47**, 603–616.

Hui SL & Walter SD (1980) Estimating error rates of diagnostic tests. *Biometrics* **36**, 167–171.

Koukounari A, Webster JP, Donnelly CA *et al.* (2009) Sensitivities and specificities of diagnostic tests and infection prevalence of Schistosoma haematobium estimated from data on adults in villages northwest of Accra, Ghana. *American Journal of Tropical Medicine and Hygiene* **80**, 435–441.

Langhi Junior DM, Bordin JO, Castelo A, Walter FD, Moraes-Souza H & Stumpf RJ (2002) The application of latent class analysis for diagnostic test validation of chronic Trypanosoma cruzi infection in blood donors. *The Brazilian Journal of Infectious Diseases* **6**, 181–187.

Machado de Assis TS, Braga ASC, Pedras MJ, Walter FD, Moraes-Souza H & Stumpf RJ (2008) Validação do teste imunocromatográfico rápido IT-LEISH para o diagnóstico da leishmaniose visceral humana. *Revista Epidemiologia e Serviços de Saúde* **17**, 105–116.

Machado de Assis TS, Braga ASC, Pedras MJ *et al.* (2011) 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* **105**, 81–85.

Ministério da Saúde (2006) *Secretaria de Vigilância em Saúde, Brasil*. http://portal.saude.gov.br (accessed 23 September 2009).

Oliveira A, Pedras MJ, De Assis IE & Rabello A (2009) Improvement of direct agglutination test (DAT) for laboratory diagnosis of visceral leishmaniasis in Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **12**, 1279–1281.

Pedras MJ, Viana LG, de Oliveira EJ & Rabello A (2008) Comparative evaluation of direct agglutination test, rK39 and soluble antigen-ELISA and RIFI for the diagnosis of visceral leishmaniasis in Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **102**, 172–178.

Peeling RW, Smith PG & Bossuyt PMM (2006) A guide for diagnostic evaluations. *Nature Reviews Microbiology* **4**, 2–6.

Pouillot R, Gerbier G & Gardner IA (2002) TAGS, a program for the evaluation of test accuracy in the absence of a gold standard. *Preventive Veterinary Medicine* **53**, 67–81.

Ransohoff DF & Feinstein AR (1978) Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *New England Journal of Medicine* **299**, 926–930.

Reid MC, Lachs MS & Feinstein A (1995) Use of methodological standards in diagnostic test research. Getting better but still not good. *Journal American Medical Association* **274**, 645–651.

Rindskopf D & Rindskopf W (1986) The value of latent class analysis in medical diagnosis. *Statistics in Medicine* **5**, 21–27.

Sheps SB & Schechter MT (1984) The assessment of diagnostic tests. *Journal American Medical Association* **252**, 2418–2422.

Sundar S, Reed SG, Singh VP, Kumar PCK & Murray HW. (1998) Rapid accurate field diagnosis of Índian visceral leishmaniasis. *Lancet* **351**, 563–565.

Sundar S, Singh RK, Bimal SK *et al.* (2007) 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* **12**, 284–289.

Thibodeau LA (1981) Evaluating diagnostic tests. *Biometrics* **37**, 801–804.

Valenstein P (1990) Evaluation diagnostic tests with imperfect standards. *American Journal of Clinical Pathology* **93**, 252–258.

Veeken H, Ritmeijer K, Searman J & Davidson R (2003) Comparison of an rK39 dipstick rapid test with direct agglutination test and splenic aspiration for the diagnosis of kala-azar in Sudão. *Tropical Medicine & International Health* **8**, 164–167.

World Health Organization (2010). Technical report series. Control of the Leishmaniasis.

Zijlstra EE, Nur Y, Desjeux P *et al.* (2001) Diagnosing visceral leishmaniasis with the recombinant K39 strip test: experience from the Sudan. *Transactions of the Royal Society of Tropical Medicine and Hygien* **6**, 108–113.

Corresponding Author **Ana Rabello**, Avenida Augusto de Lima, 1715, Barro Preto, Belo Horizonte, Minas Gerais, Brasil, CEP 30190-002. Tel.: +31-33497708; E-mail: ana@cpqrr.fiocruz.br