Latent class analysis of diagnostic tests for visceral leishmaniasis in Brazil

Tânia Santana Machado de Assis¹, Ana Rabello¹ and Guilherme Loureiro Wernick³

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 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).

Hawed 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 Koukouri et al. (2009) for schistosomiasis. Bocken 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 1000x oil immersion lens on an optic microscope); (ii) indirect fluorescence antibody test (IFAT), performed with a commercial kit (BioManguinhos, 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® Diaomed 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 (Cephr-Fiocruz) approved the study (CEPSH/Cephr-Rn; 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

<table>
<thead>
<tr>
<th>Method</th>
<th>Microscopy</th>
<th>IFAT</th>
<th>rK39-ELISA</th>
<th>DAT</th>
<th>rK39 Rapid test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>(95% CI)</td>
<td></td>
<td>(95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>77.0 (71.5-81.5)</td>
<td>88.3 (84.0-92.0)</td>
<td>99.0 (96.3-99.6)</td>
<td>88.5 (84.1-92.0)</td>
<td>94.0 (90.1-96.3)</td>
</tr>
<tr>
<td>Specificity</td>
<td>(95% CI)</td>
<td></td>
<td>(95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>99.0 (94.0-99.7)</td>
<td>83.0 (75.0-88.2)</td>
<td>82.5 (75.0-88.3)</td>
<td>95.4 (89.2-98.1)</td>
<td>100 (97.0-100.0)</td>
</tr>
</tbody>
</table>

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 other 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).

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

<table>
<thead>
<tr>
<th>Rapid test</th>
<th>IFAT</th>
<th>Microscopy</th>
<th>rK39-ELISA</th>
<th>DAT</th>
<th>Observed frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>87</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>41</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>138</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Test combination</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid test/IFAT (95% CI)</td>
<td>99.3 (97.5–99.9)</td>
<td>83.0 (74.3–88.7)</td>
</tr>
<tr>
<td>Rapid test/DAT (95% CI)</td>
<td>99.3 (97.5–99.9)</td>
<td>95.4 (89.3–98.1)</td>
</tr>
<tr>
<td>Rapid test/rK39 ELISA (95% CI)</td>
<td>99.9 (98.1–100.0)</td>
<td>82.5 (74.3–88.7)</td>
</tr>
<tr>
<td>Rapid test/Microscopy (95% CI)</td>
<td>99.0 (96.9–99.8)</td>
<td>99.0 (95.4–99.9)</td>
</tr>
</tbody>
</table>
ELISA using rK39 antigen is considered a valuable tool and has estimates of sensitivity of 93–97% and specificity of 84–97% (Machado de Assis et al. 2008; Pedras et al. 2008). DAT is simple to perform, with sensitivity estimates of 93–99%, 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; Veekon 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 (Hadj & Qu 1998), although software 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 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 Pro-
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 Council of Technological and Scientific Development) and the Oswaldo Cruz Foundation – FIOCRUZ.

References


**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