Epidemiol. Infect. (1999), **123**, 499–506. Printed in the United Kingdom

The potential of Latent Class Analysis in diagnostic test validation for canine *Leishmania infantum* infection

M. BOELAERT[1]*, K. AOUN[2], J. LIINEV[3], E. GOETGHEBEUR[3] AND P. VAN DER STUYFT[1]

[1] *Epidemiology Unit, Department of Public Health, Prince Leopold Institute of Tropical Medicine, Nationalestraat 155, B-2000 Antwerpen, Belgium*
[2] *Laboratoire d'Epidémiologie et d'Ecologie Parasitaire, Institut Pasteur de Tunis, Tunis, Tunisia*
[3] *Department of Applied Mathematics and Informatics, Universiteit Gent, Ghent, Belgium*

(*Accepted 22 July 1999*)

SUMMARY

Accuracy assessment of diagnostic tests may be seriously biased if an imperfect reference test is used such as parasitology in the diagnosis of visceral leishmaniasis. We compared classical validity analysis of serological tests for *Leishmania infantum* with Latent Class Analysis (LCA), to assess whether it circumvented the gold standard problem. Clinical status, three serological tests (IFAT, ELISA and DAT) and parasitological data were recorded for 151 dogs captured in an endemic area. Sensitivity and specificity estimates from the 2×2 contingency tables were broadly corroborated by LCA, but the latter method provided more precise estimates that were robust for the different fitted models. It furthermore yielded a higher prevalence of infection and indicated that parasitology was only 55% sensitive. LCA seems a promising technique for test validation, but caution is required when applying it to sparse data sets. The feasibility and applicability of LCA in infectious disease epidemiology is discussed.

INTRODUCTION

Epidemiologists have shown an increasing interest in Latent Class Analysis (LCA), specifically in the field of diagnostic test validation when the disease under study cannot be accurately diagnosed [1–4]. LCA is a mathematical modelling technique developed in the social sciences and is based on the idea that observed variables are jointly determined by underlying, unobserved constructs [5, 6]. It can be thought of as the analogue of factor analysis for categorical data. LCA attempts to model associations between observed categorical variables by assuming that a non-observed (latent) variable is determining these associations.

In diagnostic test validation, the true disease status of an individual can be considered as a dichotomous latent variable with two categories, 'infected' and 'not infected'. Within a group of individuals with unknown disease status, for whom at least three independent diagnostic test results are available, LCA will model the probability of each combination of test results (or response pattern) conditional on the latent class. An estimate of both disease prevalence and sensitivity and specificity of all tests can be derived from the pattern of diagnostic test results as expected under the LCA model.

In this study, our objective was to assess whether LCA corroborated findings taken from a classical validation of serological tests for *Leishmania infantum* infection. In the Mediterranean basin, a reliable diagnostic test will remain a key element in the transmission control strategy of visceral leishmaniasis for as long as there is no vaccine available [7, 8]. The direct agglutination test (DAT) has been used [9–11], alongside other serological tests [12–14], to detect *L.*

* Author for correspondence.

infantum infection in the dog, which is the main animal reservoir of the disease. These tests used to be validated against parasitology, known to be highly specific but poorly sensitive, as a reference test. The problems posed by the lack of a true gold standard for *L. infantum* infection have been discussed: the difficult interpretation of serological survey data and the fact that test characteristics were previously often determined with respect to disease and not with respect to infection [15, 16].

METHODS

A sample of 152 street dogs at risk of *L. infantum* infection was collected in the regions of Medjez el Bab and Kasserine, Tunisia, by the veterinary department of the Institut Pasteur de Tunis. Clinical data, tissue and serum samples were collected from each dog. Direct examination of a parasitological smear, a parasitological culture, and three serological tests, (indirect immunofluorescence antibody test (IFAT) [17], enzyme-linked immunosorbent assay (ELISA) [18] and direct agglutination test (DAT) [19]), were carried out. All tests were performed at the Institut Pasteur de Tunis, except for the DAT, which was completed in the Laboratory of Protozoology of the Prince Leopold Institute for Tropical Medicine in Antwerp. One serum sample out of 152 was missing in Antwerp, and DAT could not be processed. This case record was deleted from the analysis.

In the classical validation analysis, the clinical and serological test results were compared in 2×2 contingency tables to the parasitological results. A parasitologically positive result was defined as the identification of *L. infantum* on smears in direct examination and/or parasitological culture. Sensitivity and specificity of the other tests were estimated with a 95% exact binomial confidence interval (CI).

Five variables were included in the LCA analysis: clinical status, parasitological result, IFAT, ELISA and DAT. The proportion of dogs with *L. infantum*-infection and sensitivity and specificity of each test were estimated, based on the combined information of clinical, parasitological and serological tests. For more details on latent structure model specification and parameter estimation, see Appendix. Approximate CI for the parameter estimates were calculated, based on their estimated standard error (S.E.).

LCA was performed using LEM, Version 1.0. (Vermunt, 1997, unpublished), and a software program written by Qu [20]. A series of models were fitted starting with a saturated loglinear model for the observed table as well as an independence model that states there is no association between the observed test results. Subsequently a Two Latent Class (2LC)-model (I) was fitted, which states that the dogs can be classified in two mutually exclusive and exhaustive groups: the infected and the non-infected, although this information is not-observed. Furthermore this model assumes that the observed diagnostic test results are independent conditional on latent class, i.e. within the subgroup of the infected as well as the uninfected dogs, there is no association between test results. For reasons explained in the results section, we fitted a less restricted model which allowed for the direct effect of an additional latent characteristic between a pair of tests (II) [21]. Finally we explored a model with three latent classes (III), specifying dogs were a mixture of three mutually exclusive and exhaustive latent classes, which could be biologically interpreted as healthy non-infected, asymptomatically infected and symptomatically infected (ill) dogs. Models were formally compared based on the difference in goodness of fit likelihood ratio statistic ($G^2 = 2\Sigma\ O_{abcde}\ (\log\ O_{abcde}/E_{abcde})$) and Akaike's information criterion (AIC) based on the log-likelihood function (AIC $= -2$ log-likelihood $+ 2 \times$ number of parameters) [22]. Under regularity conditions, the number of degrees of freedom (D.F.) for the χ^2 distribution associated with the G^2 of each model, is given by the number of response patterns (2^5) minus the number of parameters to be estimated minus 1. Goodman's proposal [23] was followed, adding one D.F. for every parameter located on the boundary of the parameter space (either 0 or 1). To check the assumption of conditional independence between the diagnostic tests, the correlation matrix of the residuals was graphically examined following Qu [20].

RESULTS

Eleven out of the 151 dogs (7·3 %) tested parasitologically positive. Table 1 gives the performance of the clinical case definition and the serological test results with parasitological status as a reference test.

Table 2 shows the characteristics of the fitted models. The saturated model for the observed table has by definition a perfect fit and 0 D.F.. In comparison, the independence model does clearly not fit: the conditional likelihood ratio test ($G^2_{ind} - G^2_{sat}$

Table 1. *Sensitivity and specificity of clinical and serological tests for* Leishmania infantum *infection in dogs compared to parasitology as a reference test* (n = *151*)

	Sensitivity in 11 parasitologically positive dogs (n = number of true positives)	Specificity in 140 parasitologically negative dogs (n = number of false positives)
Clinically ill	0·64 (7)	0·92 (11)
DAT	0·91 (10)	0·87 (18)
IFAT	1 (11)	0·94 (9)
ELISA	0·82 (9)	0·87 (18)

= 217·5 for 26 D.F.) is significant at an α-level of 0·05. However, the 2LC model (all observed test results are independent within the latent classes), showed a very good fit: $G^2_{2LC} - G^2_{sat}$ = 14·5 for 23 D.F., P = 0·91. Three parameters in this 2LC model had a boundary estimate: sensitivity and specificity of IFAT, as well as specificity of parasitology were estimated at 100%.

Although this 2LC model (model I) fitted the data well, inspection of correlation between residuals in this model suggested residual correlation between certain test results after accounting for disease status (Fig. 1) [20].

To account for the apparently correlated error between clinical status and parasitology, a more complex latent class model (II) was fitted. It specifies that for a dog belonging to the latent class 'infected', the probability of having simultaneously a false negative result for both the clinical case definition and parasitology was higher than expected under independence. It is, e.g., biologically plausible that infected dogs without clinical symptoms (false-negative on clinical appreciation), might have lower parasite loads than clinically symptomatic dogs and, therefore, the probability of a false–negative parasitological result in the former group might be higher. Hagenaars has used the term 'direct effect' to indicate such correlated test errors [21]. Including a direct effect between clinical definition and parasitology in the group of infected dogs resulted in one additional parameter to estimate compared to the 2LC – model. The number of D.F. for model II was thus $32-12-1+3 = 22$.

The difference between the G^2 statistic for model I and model II was significant at the 0·10 level (G^2 3·6 for 1 D.F. with P = 0·06). A lower AIC indicated a preference for this less parsimonious model.

We explored whether one should distinguish three latent classes in this group of dogs, assuming, (a) a group of non-infected dogs, (b) a group of asymptomatically- infected dogs and (c) a group of symptomatically-infected dogs. A three latent class model (III) provided no better fit to the data as judged from AIC. Table 2 gives an overview of the fitted models.

Table 3 shows both the observed and estimated frequencies by model II.

The parameters of interest, sensitivity, specificity, and prevalence, as estimated by models I and II, were very similar. Table 4 compares the parameter estimates as produced by the classical contingency analysis (2×2) with those produced by LCA (model II).

The overall proportion of *L. infantum* infection in the group of 151 dogs was estimated at 13·2% by LCA, whereas parasitology was positive only in 7·3% of the dogs. This striking difference was caused by the fact that LCA estimated sensitivity of parasitology only at 55%.

DISCUSSION

Diagnostic test evaluation is frequently rendered ineffective by the lack of a 'gold standard', i.e. an exact knowledge of the true disease status of the tested individuals. New tests have to be compared with imperfect existing ones, and the performance of a new test can seriously be under- or overestimated in this way [24–26]. In this study, the high specificity values of serological tests in a classical validity study, with parasitology as a reference test, were corroborated by LCA. A 2LC-model, including a direct effect between two tests in the group of infected dogs, provided the best fit to the data. This model estimated the sensitivity of parasitology, at 55%, as remarkably low. Even though most parameter estimates produced by LCA seem close to those yielded by a 2×2 contingency table, the LCA method provided more precise estimates of sensitivity, and gave some perspective to a comparison of all the tests studied. The LCA-models identified serological tests, and in particular IFAT, as better tests for infection than parasitology. IFAT was found to be 100% sensitive and specific but, as any previously published estimate of test characteristics, these figures are subject to sampling variation. To estimate however this sampling variation, exact procedures should be used in the estimation algorithms and to date the software is not yet available. IFAT was also found to be 100%

Table 2. *Models fitted by LCA in data set (n = 151) with five observed variables: clinical definition, ELISA, IFAT, DAT, parasitology*

Model		Model specification	D.F.	Likelihood-ratio G^2	*P*-value	AIC
	Saturated model	{CEDIP}	0	0	1	438·3
	Independence model	{C, E, D, I, P}	26	217·5	0·0000	603·8
I	2LC	{X, C|X, E|X, D|X, I|X, P|X}	23	14·5	0·91	412·8
II	2LC with direct effect between clinical definition & parasitology in group of infected dogs	{X, CP|X, E|X, D|X, I|X}	22	10·9	0·98	411·2
III	3LC	{X, C|X, E|X, D|X, I|X, P|X}	14	11·6	0·63	421·9

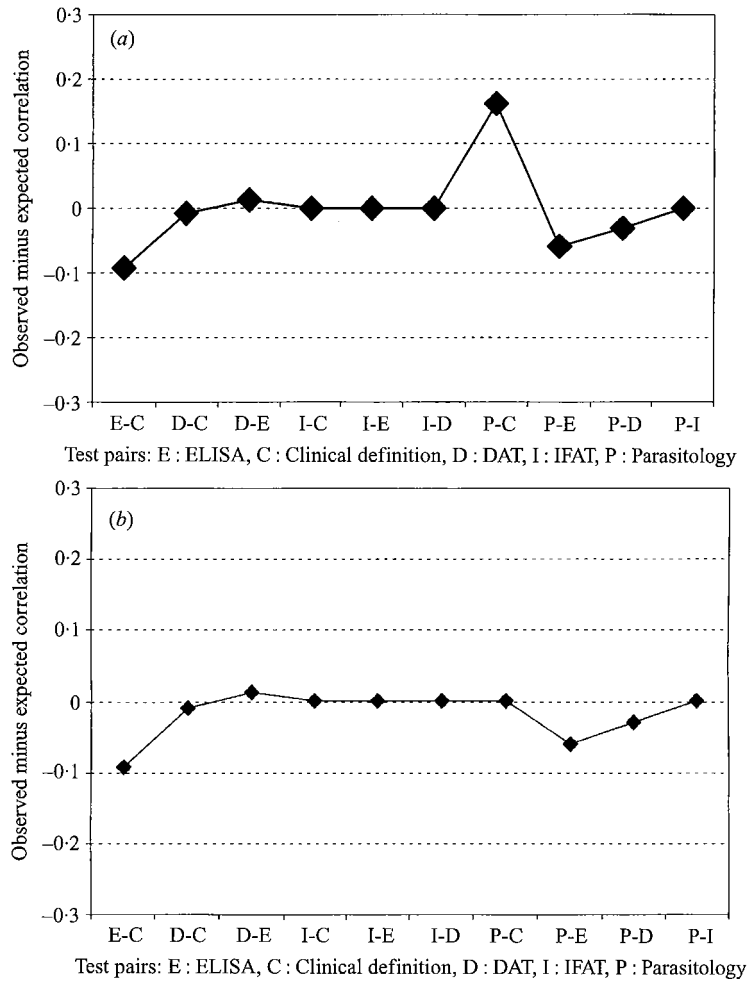

Fig. 1. (*a*): Pair-wise correlation between residuals in the two latent class model (model I). (*b*) Pair-wise correlation between residuals in the two latent class model with a direct effect between clinical status and parasitology result in the group of infected dogs (model II)

sensitive and specific by Mancianti and Meciani [12]. IFAT sensitivity was suboptimal in experimentally infected dogs, where 16 out of 25 (64%) dogs developed IFAT titres > 80 and 23 out of 25 (92%) > 40 [27]. Moreover, in a carefully observed cohort of naturally infected dogs [16], 8–9 months between infection and maximal IFAT sensitivity, 86%, was observed. Berrahal and colleagues found IFAT to be 100% sensitive in a group of symptomatic dogs, but only 1 out of 17 asymptomatic dogs positive on PCR and immunoblot was also IFAT positive [13]. They explained the latter finding hypothesizing that most

Table 3. *Observed and estimated frequencies and standardized residual for 32 response patterns as estimated by model II (n = 151)*

C E D I P*	Observed	Estimated	Standardized residual ((0−E)/√E)
0 0 0 0 0	105·000	105·812	−0·079
0 0 0 0 1	0·000	0·000	0·000
0 0 0 1 0	0·000	0·035	−0·187
0 0 0 1 1	0·000	0·020	−0·141
0 0 1 0 0	8·000	7·806	0·069
0 0 1 0 1	0·000	0·000	0·000
0 0 1 1 0	0·000	0·665	−0·815
0 0 1 1 1	0·000	0·380	−0·616
0 1 0 0 0	8·000	7·806	0·069
0 1 0 0 1	0·000	0·000	0·000
0 1 0 1 0	0·000	0·315	−0·561
0 1 0 1 1	1·000	0·180	1·933
0 1 1 0 0	1·000	0·576	0·559
0 1 1 0 1	0·000	0·000	0·000
0 1 1 1 0	7·000	5·985	0·415
0 1 1 1 1	3·000	3·420	−0·227
1 0 0 0 0	9·000	7·806	0·427
1 0 0 0 1	0·000	0·000	0·000
1 0 0 1 0	0·000	0·010	−0·100
1 0 0 1 1	0·000	0·035	−0·187
1 0 1 0 0	0·000	0·576	−0·759
1 0 1 0 1	0·000	0·000	0·000
1 0 1 1 0	0·000	0·190	−0·436
1 0 1 1 1	2·000	0·665	1·637
1 1 0 0 0	0·000	0·576	−0·759
1 1 0 0 1	0·000	0·000	0·000
1 1 0 1 0	0·000	0·090	−0·300
1 1 0 1 1	0·000	0·315	−0·561
1 1 1 0 0	0·000	0·042	−0·206
1 1 1 0 1	0·000	0·000	0·000
1 1 1 1 0	2·000	1·710	0·222
1 1 1 1 1	5·000	5·985	0·403

* C, clinical definition; E, ELISA; D, DAT; I, IFAT; P, Parasitology. 0, negative; 1, positive.

subjects in the asymptomatic group pertained to the 'type-B' dogs, that self-cure after infection [28]. In our data set, the best model distinguished between two groups of dogs, 'infected' and 'non-infected' ones. The classification did not take into account time elapsed since infection given that only cross-sectional clinical, parasitological and serological information was available. The LCA estimate of 100 % IFAT sensitivity and specificity in this data set should thus be interpreted accordingly.

LCA has become relatively easy to use, since the necessary software is now becoming accessible via internet and add-ons in commercial packages. However, LCA requires an understanding of the basic laws of probability theory and training (or expert assistance) in loglinear modelling. The design of validation studies based on LCA is not necessarily much more expensive than the classical alternative, since a minimum of three tests and roughly 100 observations are required to allow for fitting a 2LC-model of conditional independence. Adding additional tests is most useful if they are non-dependent.

As the gold standard problem is recurrent in infectious disease epidemiology, it might seem reasonable to conclude that LCA has a clear application field [3] and so we might be tempted to endorse the recommendations of previous authors [1, 4]. However, the literature gives little emphasis to the conditions of applicability of the G^2 derived χ^2 test for model comparison [2, 29]. The conditions of applicability of a χ^2 test in a 2×2 contingency table are well known; the expected frequency in each cell of the contingency table should be minimum 5. In our LCA study, this condition was not met in 26 out of 32 cells (81·2 %) of the 2^5 table. In addition, 3 of the 11 parameters in the 2LC model were estimated on the boundary, suggesting a violation of the regularity conditions [18]. One of the three, specificity of parasitology, could be considered *a priori* as 100 %, and, as such, deleted from the list of parameters to estimate. However, this was less clear in the case of sensitivity and specificity of IFAT. In the present study, the distributions of the differences in χ^2 and G^2 statistics were simulated for comparison between models I and II (detailed results not presented). Referring to Goodman's proposal to adjust the D.F. for the number of boundary parameters, both statistics appeared to behave relatively well, although in the higher tail of the distribution the criteria for rejection were slightly more liberal than intended by the theoretical asymptotic χ^2 distribution. When the same models were compared, based on unadjusted degrees of freedom, neither statistics followed the χ^2 distribution. Accordingly, we decided to apply the Goodman convention.

Simulation work based on a related but more sparse data set than the one presented here, shows that the distribution of the Likelihood Ratio statistic and the corresponding G^2-statistic of the difference between models does not follow a χ^2 distribution with either of those sets of degrees of freedom (Jane Liinev, unpublished observations). Focusing on the higher tail of the observed distribution in view of testing, the difference in the Likelihood Ratio Test G^2 between models lead to somewhat conservative tests whilst the

Table 4. *Sensitivity and specificity [95% CI] of four tests for* L. infantum *infection in 151 dogs as estimated by classical validation (2 × 2) and through LCA (model II)*

	Sensitivity		Specificity	
	2 × 2 analysis	LCA	2 × 2 analysis	LCA
Clinical	0·636 [0·308–0·891]	0·45[0·232–0·668]	0·922 [0·864–0·96]	0·931 [0·888–0·975]
DAT	0·909 [0·587–0·998]	0·95 [0·855–1]	0·871 [0·804–0·922]	0·931 [0·888–0·975]
ELISA	0·818 [0·482–0·977]	0·90 [0·769–1]	0·871 [0·804–0·922]	0·931 [0·888–0·975]
IFAT	1[0·715–1]	1	0·936 [0·882–0·97]	1
Parasitology	1*	0·55[0·332–0·768]	1*	1

* By definition (reference test).

difference in Pearson χ^2 test leads to seriously liberal tests. Especially in the field of diagnostic tests, we feel that sparse data sets will be the rule rather than the exception, since the performance level of biomedical tests is usually high. Thus one can expect zero or near-zero frequencies for response patterns with mixed test results. It is not completely clear how stringent the conditions of applicability are for LCA and what can be done to circumvent the problem of these 'quasi-structural' zeros. However, when the parameters of interest appear insensitive to this choice of model, the problem of best fit is irrelevant for practical purposes. Under the opposite scenario the presumed χ^2 distribution may lead to tests that are too liberal, and it is worthwhile to simulate the true small sample distribution of the chosen test statistic to decide on the best model. Alternatively one could opt to use the AIC as a distribution free indicator for the model choice. In summary, LCA offers a worthwhile, methodological advance in the field of diagnostic test evaluation, provided that more clarity can be created on the strategy for model comparison in sparse data sets.

ACKNOWLEDGMENTS

We wish to thank the staff of the Laboratoire d'Epidémiologie et d'Ecologie Parasitaire, Institut de Pasteur de Tunis, for the collection of data and the laboratory analysis, as well as Dominique Le Ray and Diane Jacquet of the Parasitology Department of the Institute of Tropical Medicine, Antwerp, for the DAT results. We are indebted to Yinsheng Qu and Jacques Hagenaars for appreciated comments and to Paul Vine for editorial assistance. This study was carried out with the financial support of the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen project no. 1·5.480·98 (to Dr M. Boelaert), the TDR-OMS-PNUD project ID: 890266 and the EC/STD3 project T53-CT93–0253 (to Professor R. Ben Ismail and Dr P. Ready).

APPENDIX

Let A, B, C, D, and E represent the five observed test variables: respectively clinical status, ELISA, DAT, IFAT and parasitology. The five variables are dichotomous, the test result being either negative or positive, coded as 0 or 1, and represented by the indexes a, b, c, d, e.

For example π_a^A denotes the probability that an individual will show test result a on test A.

Let X designate a latent (unobserved) variable with two mutually exclusive and exhaustive levels, indexed as 't' corresponding to the true (but unknown) infection status of the animal: not infected ($t = 0$) and infected ($t = 1$).

For example π_t^X denotes the probability that an individual will be at latent class t and π_1^X the proportion of infected dogs in the study group (prevalence rate).

The unrestricted two latent class model can then be expressed as:

$$\pi_{abcde} = \sum_t \pi_{abcdet}^{ABCDEX}. \quad (1)$$

π_{abcde} denotes the probability that an individual dog will show a pattern of test results (a, b, c, d, e). In an unrestricted two latent class model, π_{abcde} is the sum of the probability that a dog is truly infected ($t = 1$) and shows test pattern (a, b, c, d, e) plus the probability that a dog is truly not-infected ($t = 0$) and shows the same test pattern (equation 1).

Furthermore, if all observed tests results are independent from each other within the latent classes,

one can specify an equation for the partially unobserved table *XABCDE*. Given basic laws of probability, the probability that a dog is in latent class t and shows a particular test pattern (a, b, c, d, e) corresponds to the probability of belonging to latent class t multiplied by the probability to show result a on test A given latent class t, multiplied by the probability to show result b on test B given t, multiplied by the probability to show result c on C given t, etc..., as written in equation 2.

$$\pi_{abcdet}^{ABCDEX} = \pi_t^X \pi_{a|t}^{A|X} \pi_{b|t}^{B|X} \pi_{c|t}^{C|X} \pi_{d|t}^{D|X} \pi_{e|t}^{E|X}. \quad (2)$$

In the above equation, $\pi_{a|t}^{A|X}$ denotes the conditional probability that a dog will show test result a given he is in latent class t.

More specifically $\pi_{0|0}^{B|X}$ denotes the probability that a truly not infected dog tests negative on the ELISA test. This probability is more commonly known as the specificity of the ELISA test. In the same way, $\pi_{1/1}^{B|X}$ denotes the probability that a truly infected dog tests positive on ELISA, a parameter more commonly known as the sensitivity of ELISA.

Solving equation 2 means we have to estimate 11 parameters: 2 parameters for each test (sensitivity and specificity, the false negative and false positive rates being the complements of the former) and 1 parameter for the proportion of truly infected dogs, given the fact that the proportion of non-infected dogs is the complement of the latter.

Maximum likelihood estimates of the parameters can be obtained using the Estimation-Maximization iterative algorithm provided in the LEM package. To explore whether multiple maxima in the likelihood function exist, a set of random starting values was used.

The freeware LEM can be downloaded from the website of the Tilburg University: http:cwis.kub.nl/~fsw_1/mto/mto3.htm.

REFERENCES

1. Rindskopf D, Rindskopf W. The value of latent class analysis in medical diagnosis. Stat Med 1986; **5**: 21–7.
2. Formann AK, Kohlmann T. Latent class analysis in medical research. Stat Meth Med Res 1996; **5**: 179–211.
3. Hadgu A, Qu Y. A biomedical application of latent class models with random effects. Appl Statist 1998; **47**: 603–16.
4. Faraone SV, Tsuang MT. Measuring diagnostic accuracy in the absence of a 'gold standard'. Am J Psychiatry 1994; **151**: 650–7.
5. Goodman LA. The analysis of systems of qualitative variables when some of the variables are unobservable. Part I – a modified latent structure approach. Am J Soc 1974; **79**: 1179–259.
6. Heinen T. Latent class and discrete latent trait models. 1st ed. Thousand Oaks: Sage Publications, 1996: 300.
7. Tesh RB. Control of zoonotic visceral leishmaniasis: is it time to change strategies? Am J Trop Med Hyg 1995; **52**: 287–92.
8. Dye C. The logic of visceral leishmaniasis control. Am J Trop Med Hyg 1996; **55**: 125–30.
9. El Harith A, Slappendel RJ, Reiter I, et al. Application of a direct agglutination test for detection of specific anti-Leishmania antibodies in the canine reservoir. J Clin Microbiol 1989; **27**: 2252–7.
10. de Korte PM, El Harith A, Dereure J, Huigen E, Faucherre V, van der Kaay HJ. Introduction of an improved direct agglutination test for the detection of *Leishmania infantum* infection in southern France. Parasitol Res 1990; **76**: 526–30.
11. Neogy AB, Vouldoukis I, Silva OA, et al. Serodiagnosis and screening of canine visceral leishmaniasis in an endemic area of Corsica: applicability of a direct agglutination test and immunoblot analysis. Am J Trop Med Hyg 1992; **47**: 772–7.
12. Mancianti F, Meciani N. Specific serodiagnosis of canine leishmaniasis by indirect immunofluorescence, indirect hemagglutination, and counterimmunoelectrophoresis. Am J Vet Res 1988; **49**: 1409–11.
13. Berrahal F, Mary C, Roze M, et al. Canine leishmaniasis: identification of asymptomatic carriers by polymerase chain reaction and immunoblotting. Am J Trop Med Hyg 1996; **55**: 273–7.
14. Ashford DA, Badaro R, Eulalio C, et al. Studies on the control of visceral leishmaniasis: validation of the Falcon assay screening test–enzyme-linked immunosorbent assay (FAST–ELISA) for field diagnosis of canine visceral leishmaniasis. Am J Trop Med Hyg 1993; **48**: 1–8.
15. Dye C, Killick Kendrick R, Vitutia MM, et al. Epidemiology of canine leishmaniasis: prevalence, incidence and basic reproduction number calculated from a cross-sectional serological survey on the island of Gozo, Malta. Parasitol 1992; **105** (Pt 1): 35–41.
16. Dye C, Vidor E, Dereure J. Serological diagnosis of leishmaniasis: on detecting infection as well as disease. Epidemiol Infect 1993; **103**: 647–56.
17. Quilici M, Dunan S, Ranque J. L'immuno-fluorescence dans les leishmanioses. Méd Tropicale 1968; **28**: 37–43.
18. Ho M, Leeuwenburg J, Mbugua G, Warnashi A, Voller A. An enzyme-linked immunosorbent assay (ELISA) for field diagnosis of visceral leishmaniasis. Am J Trop Med Hyg 1983; **32**: 943–6.
19. Boelaert M, el Safi SH, Jacquet D, et al. Operational validation of the direct agglutination test (DAT) for diagnosis of visceral leishmaniasis. Am J Trop Med Hyg 1999; **60**:129–34.
20. Qu Y, Tan M, Kutner MH. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. Biometrics 1996; **52**: 797–810.

21. Hagenaars JA. Latent structure models with direct effects between indicators. Local dependence models. Soc Meth Res 1988; **16**: 379–405.
22. Akaike H. Factor analysis and AIC. Psychometrika 1987; **52**: 317–32.
23. Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika 1974; **61**: 215–31.
24. Staquet M, Rozencweig M, Lee YJ, Muggia FM. Methodology for the assessment of new dichotomous diagnostic tests. J Chron Dis 1981; **34**: 599–610.
25. Thibodeau L. Evaluating diagnostic tests. Biometrics 1981; **37**: 801–4.
26. Valenstein P. Evaluating diagnostic tests with imperfect standards. Am J Clin Path 1990; **93**: 252–8.
27. Killick KR, Killick KM, Pinelli E, et al. A laboratory model of canine leishmaniasis: the inoculation of dogs with Leishmania infantum promastigotes from midguts of experimentally infected phlebotomine sandflies. Parasite 1994; **1**: 311–8.
28. Hasibeder G, Dye C, Carpenter J. Mathematical modelling and theory for estimating the basic reproduction number of canine leishmaniasis. Parasitol 1992; **105** (Pt 1): 43–53.
29. Hutchinson TP, Tang KY. The value of latent class analysis in medical diagnosis. Stat Med 1987; **6**: 529–31.