

# Accuracy of Clinical Signs in the Diagnosis of Pulmonary Tuberculosis: Comparison of Three Reference Standards Using Data from a Tertiary Care Centre in Rwanda

Mugabekazi Julie<sup>1</sup>, Boelaert Marleen<sup>2</sup>, Sarushi Joseph<sup>1</sup>, Seruyange Eric<sup>1</sup>, Basinga Paulin<sup>3</sup>,  
Musamakweli Andre<sup>4</sup>, Van der Stuyft Patrick<sup>2</sup>, Bisoffi Zeno<sup>5</sup>, Moreira Juan<sup>1</sup>, Van Den Ende Jef<sup>1,2,\*</sup>

<sup>1</sup>Centre Hospitalier de Kigali, Rwanda; <sup>2</sup>Institute of Tropical Medicine, Antwerp, Belgium; <sup>3</sup>School of Public Health, National University of Rwanda; <sup>4</sup>National University of Rwanda and <sup>5</sup>Centro per le Malattie Tropicali, Negrar, Verona, Italy

**Abstract:** *Objective:* To determine the prevalence of TB, and the diagnostic sensitivity and specificity of major disease characteristics in a tertiary hospital setting in Rwanda, relative to three reference standards.

*Study Design and Setting:* A prospective study was conducted in which 300 consecutive patients with cough of at least 2-weeks duration were evaluated at a tertiary healthcare facility. We compared the estimates of TB prevalence and the diagnostic accuracy of fever, haemoptysis, sputum smear microscopy, radiological signs, and HIV infection as generated by a latent class analysis (LCA) with those given by culture and by a composite reference standard (CRS), which relied on bacteriological confirmation and/or cavities.

*Results:* LCA estimated the prevalence of TB at 44%. The most sensitive characteristics were fever (90%) and HIV infection (86%), but both lacked specificity. The most specific characteristics were microscopy (99%), X-Ray cavities (97%) and apical infiltrates (93%). When culture was taken as a reference standard, the prevalence was 38%; for the CRS, it was 45%. For both, the diagnostic sensitivity and specificity were comparable to those obtained with LCA.

*Conclusion:* Three reference standards produced comparable diagnostic sensitivities and specificities using major symptoms and signs of pulmonary TB; only LCA allowed estimating the diagnostic characteristics of culture. Both LCA and CRS estimated the probability of disease higher than culture alone.

**Key Words:** Tuberculosis, prevalence, sensitivity, specificity, diagnostic accuracy, latent class analysis.

## INTRODUCTION

In the last century, gradual regression of tuberculosis (TB) was observed in all developed and in most developing countries, thanks to better hygiene, improved nutrition and specific drugs [1, 2]. Since the start of the HIV pandemic in the early eighties a recrudescence of TB has been noted, as the risk of overt disease is closely related to the CD4 count [3-6]. Every year, more than eight million new cases are reported worldwide, and more than three million die. More than 95% are found in developing countries, and more than 80% are young adults. In 1993 the World Health Organization (WHO) declared tuberculosis a worldwide emergency [1, 2].

The diagnosis of TB remains an eternal problem. With the advent of the HIV pandemic clinical and radiological aspects of TB became even more unspecific [7-9]. Direct microscopical examination with Ziehl stain and culture lacked sensitivity, and culture results arrived late, sometimes reported months after ordering [10]. Moreover, in many hospital settings in developing countries new diagnostic techniques such as broncho-alveolar lavage (BAL), rapid culture (Bactec®) and PCR have been unavailable.

TB program managers often suggest that clinicians treat too many patients without bacteriological evidence. Nevertheless, if one would only treat sputum smear-positive (SSP) cases, more than 50% of true cases of tuberculosis would remain untreated, and this percentage of missed cases would be even higher in countries with high HIV prevalence [11]. Several studies have addressed the validity of clinical diagnosis of tuberculosis in low-income countries, and were recently reviewed with the purpose of assessing the appropriateness and usefulness of the criteria used in the diagnosis and the decision to treat [8]. Every proposed approach has its own strengths and limitations. Some authors use culture as the reference standard to assess the diagnostic sensitivity and specificity of clinical, radiographic and other predictors [10,12,13]. Although culture is widely regarded as the “gold standard” for pulmonary TB diagnosis and its specificity is assumed to be 100%, its diagnostic sensitivity remains below 100%. Other authors rely on expert review, exhaustive clinical investigation, histological features or success of treatment [14]. All of these are at best imperfect reference standards [15].

“Latent class analysis” (LCA) was suggested as a solution when a validated gold standard is lacking [16]. LCA is a statistical method developed in the social sciences and introduced in biomedicine in the early eighties. It allows for the estimation of prevalence (prior probability of disease) and

\*Address correspondence to this author at the Nationalestraat 155, 2000 Antwerp, Belgium; Tel: +32(0)3/247.64.29; Fax: +32(0)3/247.64.52; E-mail: jvdende@itg.be

sensitivity and specificity of disease characteristics in the absence of a reliable reference standard [17-19]. LCA has been used in diagnostic accuracy studies in various infectious diseases, e.g., in Chagas disease, leishmaniasis, schistosomiasis, leptospirosis, Herpesvirus 8 and pneumococcal infection [20-26].

This study intends to estimate the prior probability of TB, and the diagnostic sensitivity and specificity of major disease characteristics for TB in a group of patients admitted to a tertiary hospital setting in Rwanda. It compares 3 reference standards: 1) culture; 2) a composite reference standard (direct microscopy and/or culture and/or cavities on chest X-ray); and 3) a classification obtained with LCA.

## MATERIAL AND METHODS

This prospective study was conducted over a three months period in the public wards of internal medicine of CHK, a 600-bed hospital of national referral center in Rwanda. Three hundred consecutive patients who were hospitalized with a cough of at least 2 weeks duration were included in the study. Extra-pulmonary TB (including miliary TB) and patients already under TB treatment were excluded.

The following data were collected: age and sex; history of fever; night sweats; haemoptysis; direct microscopy of sputum with Ziehl stain; culture of sputum on Löwenstein Jensen medium; HIV enzyme linked serum assay (ELISA) (VIRONOSTICA® bioMérieux, Marcy-l'Etoile) with confirmation by a second ELISA if positive (MUREX DIAGNOSTICS®, Abbott, Dartford); a Western Blot test (HIV BLOT 2.2® Genelabs Diagnostics, Singapore) in case of a discordant result; chest x-ray read by a radiologist and an internist or by three internists on a consensus base, all blinded to previously obtained clinical information. Direct microscopy was considered positive if two out of three specimens contained at least one + mark of acid-fast bacilli; if only one sputum was positive, or if two were questionable, another series of three sputa was performed.

All clinical and laboratory testing was part of routine examination. For HIV testing, specific informed consent was asked and counseling offered. Approval for this study was obtained from the relevant authorities of the Ministry of Health of Rwanda.

The diagnostic sensitivity (Se) and specificity (Spe) of clinical signs and symptoms were assessed as follows: first, a classical bivariate analysis was performed with culture as the reference standard. Second, we compared disease characteristics with a composite reference standard, which considered a patient with a positive culture, and/or at least two positive smears, and/or thick walled cavities on chest X-ray as "a case of TB". Only patients who were negative on all three criteria were considered as "no-TB". Third, we estimated Se and Spe of disease characteristics with an LCA strategy [17,27,28]

In patients for whom results from at least 3 diagnostic TB tests are available, LCA can distinguish two subgroups: "patients with pulmonary TB" and "patients without pulmonary-TB". The true disease status of these persons is considered as a "latent variable" with two mutually exclusive and exhaustive categories, "TB" and "no-TB". Observed diagnostic test results in this study including clinical, bacteriological and

radiological findings do not permit the measurement of disease status directly, because they are all (imperfect) indicators of the underlying latent variable "disease status".

In a cohort, different patterns of findings can be present: if we limit the example to tuberculosis and to three findings, some patients will present fever, cavities, and a positive smear, others cavities and a positive smear but no fever, others fever and a positive smear but no cavities, and so on. The observed numbers of patients in each pattern form a series, a "constellation", which can not be solved directly mathematically, but iterative application of algorithms can deduce the probability of each test pattern for the hypothetical disease, the prevalence of the hypothetical disease, and the diagnostic sensitivity and specificity. Different models can be tested for "fit": the calculated frequency of each pattern, based on the data estimated by the model, is compared with the frequency observed in reality. A low p-value means bad modeling, since there is a significant difference between the observed frequency (in the cohort) and the computed frequency (based on the proposed prevalence, Se and Spe).

In basic latent class models, the observed or manifest variables are assumed to be independent conditional on latent class. In advanced models this condition is relaxed [27,29]. Conditional independence of the manifest variables is usually examined by inspecting the residual correlation between pairs of tests after fitting the basic LCA model. If significant, a more complex LCA model should be fitted to the data that includes a term describing this direct dependence between the two tests. Basic and complex models are then compared and if the complex model accounting for conditional dependence does not provide a significantly better fit to the data, the basic LCA model will be preferred [26]

We fitted several series of latent class models to our data with the LATENT GOLD package (V 2.0.18, Statistical Innovations, Belmont, MA). We included in the analysis only variables that showed  $Se + Spe > 1$  in the comparison with culture (direct microscopy, cavities, apical infiltrates, HIV infection, haemoptysis and fever). We identified the LCA model providing the best fit to the data by comparing the difference in likelihood statistic ( $L^2$ ), the Bayesian Information Criterion (BIC) and Akaike's Information Criterion (AIC) [30]. The best LCA model provided the parameters of interest, and an approximate 95% CI was computed for Se and Spe as the interval lying within  $\pm 1.96 \times$  standard error of the estimate.

## RESULTS

Of 300 patients 175 were female, 125 male. Median age was 36 years (15 to 87). Direct sputum microscopy was positive in 88, culture in 115 and HIV Elisa in 216 patients. Of 88 patients with positive direct microscopy, 13 were culture negative. Of 115 patients with a positive culture, 40 had a negative direct sputum examination. Cavities were present in 34 and apical infiltrations in 73 patients. There were no missing data on any of the tests or signs for the 300 patients.

Classical contingency table analysis showed a significant correlation of direct microscopy, cavities, apical infiltrates and HIV infection with culture. Reticulo-nodular infiltrates were inversely correlated with positive culture, but this finding was excluded for further analysis because of possible bias through miliary TB (Table 1).

**Table 1. Prevalence of TB, Sensitivity (%), Specificity (%) and Odds Ratio (95% C.I.) of Disease Characteristics Compared to Culture as Reference Standard (n=300)**

	Culture Positives n=115	Sensitivity	Culture Negatives n=185	Specificity	Odds Ratio
Fever or nightsweats	102	89	161	13	1.17 (CI : 0.54-2.57)
Haemoptysis	23	20	29	84	1.34 (CI : 0.70-2.58)
Apical infiltrates	45	39	28	85	3.60 (CI : 2.00-6.52)
Cavities	25	22	9	95	5.43 (CI : 2.28-13.25)
Bilateral infiltrates	28	24	56	70	0.74 (CI : 0.42-1.31)
Basal infiltrates	34	30	67	64	0.74 (CI : 0.43-1.26)
Reticulonodular Infiltrates	7	06	35	81	0.28 (CI : 0.11-0.69)
HIV infection	100	87	116	37	3.97 (CI : 2.04-7.79)
Ziehl stain	75	65	13	93	24.81 (CI : 11.90-52.60)

**Table 2. Features of Different LCA Models Fitted to the Data. Numbers of Classes, Manifest Variables and Dependencies, and Goodness of Fit for the 5 Models**

Model	Number of Latent Classes	Number of Manifest Variables	Number of Dependencies between Pairs of Tests Controlled for	L <sup>2</sup>	df	p-value	BIC	AIC
1	2	7	0	83.06	112	0.98	-555	-140
2	2	7	3	71.80	109	0.99	-549	-146
3	3	7	0	77.48	104	0.98	-515	-130
4	2	5	0	27.84	20	0.11	-86	-12
5	2	5	1	23.51	19	0.21	-84	-14

Models 1-3 are based on direct microscopy, cavities, apical infiltrates, haemoptysis, fever, HIV infection and culture. In model 4 and 5 we remove findings with lower discriminative power: haemoptysis (low sensitivity) and fever (low specificity). In model 2 and 5 we control for conditional independence of the manifest variables.

L<sup>2</sup>: the likelihood ratio chi-squared statistic is used to assess how well the model fits the data. It indicates the amount of the relationship between the variables that remains unexplained by a model; the larger the value, the poorer the model fits the data. As a rule of thumb, a good fit is provided by a model when the L<sup>2</sup> for that model is not substantially larger than the degrees of freedom (df). However, for model comparison, a formal test of the difference in L<sup>2</sup> between the models should be performed, or criteria as BIC or AIC should be used.

BIC: Bayesian Information Criterion for model comparison. Lowest BIC corresponds to best model fit.

AIC: Akaike's Information Criterion.

Table 2 displays an overview of well-fitting LCA models. **Model 1** is a basic two-Latent Class model with seven disease characteristics. **Model 2** includes the 7 characteristics of model 1 and includes moreover direct dependence between cavities and unilateral apical infiltrates, haemoptysis and cavities, and unilateral apical infiltrates and fever. This model provided significant better fit to the data than model 1. **Model 3** is a three-latent class model that fitted the data well, but the interpretation was not meaningful. It separated patients with cavities from patients with upper lobe consolidation, a clinically less relevant distinction. **Model 4** was constructed in an attempt to remove the less discriminant signs haemoptysis (low sensitivity) and fever (low specificity) from the model. While it resulted in very similar parameter estimates of Se and Spe, model fit was poor. **Model 5**, controlling for dependence between cavities and unilateral apical infiltrates, did not improve the fit of the model 4.

LCA-Model 2 was identified as the best model. It showed a TB prevalence of 44% and good diagnostic sensitivity for fever, HIV infection, direct microscopy and culture. On the other hand haemoptysis, cavities, apical infiltrates, direct microscopy and culture showed good specificity.

Table 3 compares estimates for Se, Spe and prevalence obtained by bivariate analysis with culture and composite reference standard and by the LCA approach. Most values are similar, only the prevalence of TB and the specificity estimate for direct microscopy were substantially different (although non-significant for the former).

Table 4 shows the different observed combinations of disease characteristics as well as the expected frequencies predicted by Model-2. Of 128 (2<sup>7</sup>) possible patterns, 55 were observed. For each pattern, the post-test probability was

**Table 3. Parameter Estimate Provided by Different Methods. Prevalence, Sensitivities and Specificities for the Preferred Model of LCA, the Model with Culture Alone and the Model with the Composite Reference Standard**

	Best LCA model (model 2)		Classical analysis with culture as reference standard		Classical analysis with composite reference standard	
	Se	Spe	Se	Spe	Se	Spe
<b>Fever or nightsweats</b>	91 (85-96)	14 (9-20)	89 (81-94)	13 (8-19)	89 (82-94)	13 (9-20)
<b>Haemoptysis</b>	20 (13-28)	86 (80-91)	20 (13-28)	84 (78-89)	22 (15-30)	86 (80-91)
<b>Ziehl stain</b>	66 (56-76)	<u>99</u> (98-100)	65 (56-74)	<u>93</u> (88-96)	64 (56-73)	n.a.
<b>Cavities</b>	21 (14-29)	97 (94-100)	22 (15-30)	95 (91-98)	25 (18-33)	n.a.
<b>Apical infiltrates</b>	42 (33-52)	90 (85-95)	39 (30-49)	85 (79-90)	39 (31-48)	87 (82-92)
<b>HIV infection</b>	88 (82-94)	40 (33-48)	87 (80-93)	37 (30-45)	85 (78-91)	39 (32-47)
<b>Culture</b>	84 (77-92)	98 (93-100)	n.a.	n.a.	84 (77-90)	n.a.
<b>Prevalence</b>	<b><u>44</u> (37-51)</b>		<b><u>38</u> (33-44)</b>		<b><u>45</u> (40-51)</b>	

n.a.: not applicable; confidence intervals between brackets.  
Values that are different between models are underlined.

computed based on estimated parameters. All patients with patterns yielding over 80% of post-test probability were bacteriologically confirmed. Fourteen HIV positive patients had only fever and a positive culture. Six patients with cavities were classified as low probability (<50%).

## DISCUSSION

Our study yielded consistent estimates for the diagnostic accuracy of clinical and radiological signs and symptoms with three different reference standards. Both a latent class and a composite reference standard approach suggested that the prevalence of TB in this group of patients was approximately 44%, and thus a relative 16% higher than if estimated by culture alone (38%). The high prevalence of TB can be explained by the tertiary care level and by the frequent co-infection HIV-TB. Together with the high HIV prevalence, the tertiary care level probably explains also the frequency of microscopy negative cases (false negative rate 34% with the best LCA model), since most sputum smear-positive patients are treated at the level of the health centre or the district hospital.

The higher TB prevalence estimated by the latent class as well as composite reference standard approach compared to the culture as reference standard is not surprising, since "closed focus" pulmonary TB exists, and since the diagnostic sensitivity of culture is below 100% [10,12,31]. During the preparation of the sample, decontamination diminishes the bacterial load considerably; bacterial overgrowth, defective culture media, excessive delay between sputum collection and inoculation and a too short reading period are other causes of false negative cultures [32,33]. Therefore, and not surprisingly, the specificity of direct microscopy was estimated as higher with the LCA than in the classical approach compared to culture.

We acknowledge that the patient group studied is subject to selection bias and we do not want to claim that the esti-

mates for Se, Spe and prevalence obtained in this study are valid beyond a tertiary-care level in a region with a high HIV prevalence. Patients admitted to the study ward in Kigali were already filtered by the health system, and are not representative for those admitted to a district hospital in Rwanda e.g. Especially the specificity estimates will be affected by this bias, but the considerable presence of sputum smear-negative cases in our study group may also affect the sensitivity estimates of certain disease characteristics. By excluding all patients already under treatment, we excluded the majority of multidrug resistant TB, though these are still rather rare in Rwanda (3.9% of new cases) [34]. The exclusion of miliary TB and TB pleurisy was dictated by the fact that we focused on pulmonary TB *sensu stricto*.

As for the choice of variables, we did not include the response to a treatment with an antibiotic, since most patients already received several courses of antibiotics at the referring level. Moreover, the value of this "clinical test" has been challenged [35].

LCA led in this study to broadly similar sensitivity and specificity estimates of disease characteristics as two alternative reference standards: a composite reference standard, as well as an external reference standard (culture). Notwithstanding the above design limitations (e.g., selection of patients at reference level), our findings are not dissimilar from those of other authors. A good sensitivity of fever and HIV infection, and a high specificity of haemoptysis and cavities have been found also in a study done in Burundi and Tanzania in 1997, also at a reference level, with culture as gold standard [14]. Direct microscopy had a moderate, and culture a good sensitivity, also in the study by Von Gottberg *et al.* [36]. The non-negligible sensitivity of direct microscopy at our reference level was somehow unexpected. It suggests that human error could have been a contributing factor or that patients became smear positive between first attendance and referral.

**Table 4. Model 4, the Posterior Probability for the TB Class, the Expected, Actual and Cumulative Number of Patients Per Combination of Disease Characteristics**

Fever or Nightsweats	Haemoptysis	Cavities	Apical Infiltrates	HIV Infection	Culture	Ziehl Stain	Estimated Frequency	Observed Frequency	Cumulative Numbers	Posterior Probability of TB
Y	Y	Y	Y	Y	Y	Y	0.9	2	2	1.00
Y	Y	N	Y	Y	Y	Y	3.7	3	5	1.00
N	Y	N	Y	Y	Y	Y	0.7	1	6	1.00
Y	N	Y	N	Y	Y	Y	6.3	8	14	1.00
N	N	N	Y	Y	Y	Y	2.9	3	17	1.00
Y	N	Y	Y	Y	Y	Y	1.8	2	19	1.00
N	N	Y	N	Y	Y	Y	0.5	1	20	1.00
Y	Y	Y	N	Y	Y	Y	3.2	4	24	1.00
Y	N	N	Y	Y	Y	Y	15.7	14	38	1.00
Y	N	N	Y	N	Y	Y	2.2	5	43	1.00
Y	Y	Y	N	N	Y	Y	0.5	1	44	1.00
Y	Y	N	N	Y	Y	Y	4.9	4	48	1.00
N	N	N	Y	N	Y	Y	0.4	2	50	1.00
Y	N	N	N	Y	Y	Y	21.0	20	70	1.00
N	N	N	N	Y	Y	Y	1.8	2	72	1.00
Y	Y	N	N	N	Y	Y	0.7	1	73	1.00
Y	N	N	N	N	Y	Y	2.9	1	74	1.00
N	N	N	N	N	Y	Y	0.2	1	75	1.00
Y	Y	N	Y	Y	N	Y	0.7	2	77	0.99
Y	N	Y	N	Y	N	Y	1.2	1	78	0.99
Y	Y	Y	N	Y	Y	N	1.7	2	80	0.99
Y	N	N	Y	Y	N	Y	2.9	1	81	0.99
N	Y	N	Y	Y	N	Y	0.1	1	82	0.99
Y	Y	N	Y	Y	Y	N	1.9	2	84	0.99
Y	N	Y	N	Y	Y	N	3.3	4	88	0.99
Y	N	N	Y	Y	Y	N	8.2	9	97	0.98
N	N	N	Y	Y	Y	N	1.6	1	98	0.98
Y	Y	N	N	Y	N	Y	0.9	2	100	0.97
Y	N	N	Y	N	N	Y	0.4	1	101	0.96
Y	N	N	N	Y	N	Y	4.1	5	106	0.95
Y	Y	N	N	Y	Y	N	2.7	3	109	0.94
Y	N	N	Y	N	Y	N	1.2	1	110	0.93
N	N	Y	N	N	Y	N	0.0	1	111	0.92
Y	N	N	N	Y	Y	N	12.0	14	125	0.91
N	N	N	N	Y	Y	N	1.1	1	126	0.86
Y	Y	Y	Y	Y	N	N	0.1	1	127	0.73
Y	N	N	N	N	Y	N	2.3	2	129	0.67
Y	N	Y	Y	Y	N	N	0.3	1	130	0.64
Y	N	Y	Y	N	N	N	0.1	1	131	0.27
Y	Y	N	Y	Y	N	N	1.5	1	132	0.24
Y	N	Y	N	Y	N	N	2.8	1	133	0.22
Y	N	N	Y	Y	N	N	8.8	9	142	0.17
N	N	Y	N	Y	N	N	0.4	1	143	0.14
N	N	N	Y	Y	N	N	2.5	5	148	0.11
Y	Y	Y	N	N	N	N	0.5	1	149	0.08
Y	Y	N	Y	N	N	N	0.8	1	150	0.06
Y	N	Y	N	N	N	N	1.6	2	152	0.05
Y	Y	N	N	Y	N	N	10.0	10	162	0.05
Y	N	N	Y	N	N	N	5.2	4	166	0.04
Y	N	N	N	Y	N	N	65.4	66	232	0.03
N	Y	N	N	Y	N	N	1.3	2	234	0.03
N	N	N	N	Y	N	N	8.9	7	241	0.02
Y	Y	N	N	N	N	N	6.6	8	249	0.01
Y	N	N	N	N	N	N	43.4	43	292	0.01
N	N	N	N	N	N	N	5.9	8	300	0.00

Posterior probability of TB: the calculated posterior probability for patients with each pattern of belonging to the TB class, based on estimated prevalence, se and sp. The columns "expected frequency" and "observed frequency" allow comparison between the calculated frequency within each pattern, based on the data estimated by the model, with the expected, or observed in reality.

The last two columns show the post-test probability of a pattern, and the cumulative number of patients exceeding this post-test probability. All patients having over 80% of post-test probability had bacteriological confirmation. Fourteen HIV positive patients had only fever and a positive culture. Six patients with cavities were classified as low probability (<50%).

LCA has been used widely in veterinary medicine, but in human infectious disease the literature is still scarce. Five studies compared LCA results with another reference standard, and 3 estimated higher prevalences of disease with LCA [21,22,25]. Pirard *et al.* evaluated screening tests for Chagas disease and report a lower prevalence of infection with LCA, compared to the composite reference standard (i.e. positivity in all tests) [26]. Butler evaluated tests for pneumococcal disease and found no difference between the estimates by LCA and those based on a reference standard [23]. Whether LCA will yield or not similar estimates compared to classical methods is highly dependent on the nature and degree of misclassification by the reference standard it is compared with.

LCA might be a useful tool that provides insight in the problem of misclassification by imperfect reference standards, when validating diagnostic signs and symptoms. The composite reference standard evaluated in our study reached similar disease prevalence and sensitivity and specificity estimates as LCA, and was promoted by Alonzo and Pepe as more transparent and more reliable than LCA [37]. However, one might question the classification by the CRS of six patients with negative bacteriology and positive radiological findings as “certain” TB. LCA generated low post-test probabilities for these six patients; although clinicians would treat several of them, we can not consider them as “reference” TB cases.

The advantage of LCA compared to the composite reference standard is that, in our study, LCA allowed for the estimation of the specificity of “culture”, “smear microscopy”, and “cavities on X-ray” in a non-deterministic way, whereas the CRS considered them as 100 % specific by definition. Most interestingly, LCA and not CRS, allowed us to examine the performance of culture as a reference test for TB diagnosis in this data set. More generally, LCA produces estimates that take into account the existing uncertainty surrounding the performance of the so-called reference standard.

Clinical decisions regarding TB status should be made taking into account the final post-test probability after exploring all findings, and comparing this probability with the therapeutical threshold. LCA as such has no direct role in patient-by-patient clinical case management of TB, but it expands our toolbox in clinical research, as it is one of the few statistical techniques available to address the issue of prevalence, sensitivity and specificity estimation when no gold standard exists to do so. As far as it allows us to flag (and correct for some of) the misclassification bias that cripples many of our diagnostic accuracy studies, we feel LCA is a useful methodological addition.

## CONCLUSION

This study shows that the latent class approach in diagnostic accuracy study gave consistent estimates of sensitivities and specificities of symptoms and signs, when compared to the classical culture or a composite reference standard. However, both LCA and the composite reference standard suggest a higher disease prevalence than culture alone. The superiority of LCA relies in its ability to examine the performance of culture, the classical reference test.

## REFERENCES

- [1] WHO Report 2002: Global Tuberculosis Control. Surveillance, planning, finance. Geneva: World Health Organisation, 2002.
- [2] Frieden TR, Sterling TR, Munsiff SS, *et al.* Tuberculosis. *Lancet* 2003; 362: 887-899.
- [3] Glynn JR. Resurgence of tuberculosis and the impact of HIV infection. *Br Med Bull* 1998; 5: 579-593.
- [4] Cantwell MF, Binkin NJ. Impact of HIV on tuberculosis in sub-Saharan Africa: a regional perspective. *Int J Tuberc Lung Dis* 1997; 1: 205-214.
- [5] Batungwanayo J, Taelman H. Impact of immunodeficiency virus infection on tuberculosis in Kigali, Rwanda: One year study of 377 consecutive cases. *Int J Infect Dis* 1996; 22-27.
- [6] Lawn SD, Butera ST, Shinnick TM. Tuberculosis unleashed: the impact of human immunodeficiency virus infection on the host granulomatous response to *Mycobacterium tuberculosis*. *Microbes Infect* 2002; 4: 635-646.
- [7] Batungwanayo J, Taelman H, Dhote R, *et al.* Pulmonary tuberculosis in Kigali, Rwanda. Impact of human immunodeficiency virus infection on clinical and radiographic presentation. *Am Rev Respir Dis* 1992; 146: 53-56.
- [8] Siddiqi K, Lambert ML, Walley J. Clinical diagnosis of smear-negative pulmonary tuberculosis in low-income countries: the current evidence. *Lancet Infect Dis* 2003; 3: 288-296.
- [9] Lee MP, Chan JW, Ng KK, Li PC. Clinical manifestations of tuberculosis in HIV-infected patients. *Respirology* 2000; 5: 423-426.
- [10] Apers L, Mutsvangwa J, Magwenzi J, *et al.* A comparison of direct microscopy, the concentration method and the Mycobacteria Growth Indicator Tube for the examination of sputum for acid-fast bacilli. *Int J Tuberc Lung Dis* 2003; 7: 376-381.
- [11] Colebunders R, Bastian I. A review of the diagnosis and treatment of smear-negative pulmonary tuberculosis. *Int J Tuberc Lung Dis* 2000; 4: 97-107.
- [12] Crampin AC, Floyd S, Mwaungulu F, *et al.* Comparison of two versus three smears in identifying culture-positive tuberculosis patients in a rural African setting with high HIV prevalence. *Int J Tuberc Lung Dis* 2001; 5: 994-999.
- [13] Kanaya AM, Glidden DV, Chambers HF. Identifying pulmonary tuberculosis in patients with negative sputum smear results. *Chest* 2001; 120: 349-355.
- [14] Samb B, Henzel D, Daley CL, *et al.* Methods for diagnosing tuberculosis among in-patients in eastern Africa whose sputum smears are negative. *Int J Tuberc Lung Dis* 1997; 1: 25-30.
- [15] Barnes PF, Cave MD. Molecular epidemiology of tuberculosis. *N Engl J Med* 2003; 349: 1149-1156.
- [16] Hadgu A. The discrepancy in discrepant analysis. *Lancet* 1996; 348: 592-593.
- [17] Goodman L. 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: 1179-1259.
- [18] Rindskopf D, Rindskopf W. The value of latent class analysis in medical diagnosis. *Stat Med* 1986; 5: 21-27.
- [19] Formann AK, Kohlmann T. Latent class analysis in medical research. *Stat Methods Med Res* 1996; 5: 179-211.
- [20] Bajani MD, Ashford DA, Bragg SL, *et al.* Evaluation of four commercially available rapid serologic tests for diagnosis of leptospirosis. *J Clin Microbiol* 2003; 41: 803-809.
- [21] Boelaert M, Rijal S, Regmi S, *et al.* A comparative study of the effectiveness of diagnostic tests for visceral leishmaniasis. *Am J Trop Med Hyg* 2004; 70: 72-77.
- [22] Booth M, Vounatsou P, N goran EK, *et al.* The influence of sampling effort and the performance of the Kato-Katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Cote d'Ivoire. *Parasitology* 2003; 127: 525-531.
- [23] Butler JC, Bosshardt SC, Phelan M, *et al.* Classical and latent class analysis evaluation of sputum polymerase chain reaction and urine antigen testing for diagnosis of pneumococcal pneumonia in adults. *J Infect Dis* 2003; 187: 1416-1423.
- [24] Langhi DM, Bordin JO, Castelo A, *et al.* The application of latent class analysis for diagnostic test validation of chronic *Trypanosoma cruzi* infection in blood donors. *Braz J Infect Dis* 2002; 6: 181-187.
- [25] Pellett PE, Wright DJ, Engels EA, *et al.* Multicenter comparison of serologic assays and estimation of human herpesvirus 8 seroprevalence among US blood donors. *Transfusion* 2003; 43: 1260-1268.

- [26] Pirard M, Iihoshi N, Boelaert M, *et al.* The validity of serologic tests for *Trypanosoma cruzi* and the effectiveness of transfusional screening strategies in a hyperendemic region. *Transfusion* 2005; 45: 554-561.
- [27] Hadgu A, Qu Y. A biomedical application of latent class models with random effects. *Appl Stat* 1998; 47: 603-616.
- [28] Heinen T. *Latent Class and Discrete Latent Trait Models*. Thousand Oaks: Sage Publications, 1996.
- [29] Hagenaars JA. Latent Structure Models with Direct Effects Between Indicators. *Local Dependence Models. Sociological Methods & Research* 1988; 16: 379-405.
- [30] Agresti A. *Categorical Data Analysis*. Second ed. New York: John Wiley & Sons, 1990.
- [31] Ginesu F, Pirina P, Sechi LA, *et al.* Microbiological diagnosis of tuberculosis: a comparison of old and new methods. *J Chemother* 1998; 10: 295-300.
- [32] Thornton CG, MacLellan KM, Brink TL, *et al.* Novel method for processing respiratory specimens for detection of mycobacteria by using C18-carboxypropylbetaine: blinded study. *J Clin Microbiol* 1998; 36: 1996-2003.
- [33] Lipsky BA, Gates J, Tenover FC, Plorde JJ. Factors affecting the clinical value of microscopy for acid-fast bacilli. *Rev Infect Dis* 1984; 6: 214-222.
- [34] Umubyeyi AN, Vandebriel G, Gasana M, *et al.* Results of a national survey on drug resistance among pulmonary tuberculosis patients in Rwanda. *Int J Tuberc Lung Dis* 2007; 11: 189-194.
- [35] Siddiqi K, Walley J, Khan MA, Shah K, Safdar N. Clinical guidelines to diagnose smear-negative pulmonary tuberculosis in Pakistan, a country with low-HIV prevalence. *Trop Med Int Health* 2006; 11: 323-331.
- [36] von Gottberg A, Sacks L, Machala S, Blumberg L. Utility of blood cultures and incidence of mycobacteremia in patients with suspected tuberculosis in a South African infectious disease referral hospital. *Int J Tuberc Lung Dis* 2001; 5: 80-86.
- [37] Alonzo TA, Pepe MS. Using a combination of reference tests to assess the accuracy of a new diagnostic test. *Stat Med* 1999; 18: 2987-3003.