Thorax Online First, published on May 31, 2006 as 10.1136/thx.2005.054908

The current evidence on diagnostic accuracy of commercial based nucleic acid amplification tests for the diagnosis of pulmonary tuberculosis.

A meta-regression analysis

Stefania Greco[1,2], Enrico Girardi[3], Assunta Navarra[3], Cesare Saltini[2]

Affiliations:

[1] Dip.to di Malattie Polmonari, Azienda Ospedaliera San Camillo-Forlanini; [2] Dipartimento di Medicina Interna, Università di Roma "Tor Vergata"; [3] Dip.to di Epidemiologia, INMI L. Spallanzani-IRCCS; Roma, Italy

Key words: nucleic acid amplification techniques, pulmonary tuberculosis, diagnostic tests, meta-analysis

Word count: 3644

Corresponding author:

Stefania Greco, MD

Dip.to di Malattie Polmonari

Azienda Ospedaliera San Camillo-Forlanini

P.za Carlo Forlanini 1

00151 Roma

stgreco@scamilloforlanini.rm.it

ABSTRACT

BACKGROUND

Even if commercial nucleic acid amplification tests (NAATs) have become the most frequently used molecular tests for laboratory diagnosis of pulmonary tuberculosis (TB), published studies report variable estimates of their diagnostic accuracy. We analyzed the accuracy of commercial NAATs for pulmonary TB diagnosis separately on smear-positive and smear-negative respiratory samples, using culture as reference standard.

METHODS

English-language studies reporting data sufficient for calculating sensitivity and specificity of commercial NAATs on smear-positive and/or smear-negative respiratory samples were included. Meta-regression was used to analyse associations with reference test quality, TB prevalence, sample and test type. Predictive values for different levels of pre-test probability were quantified using Bayes' approach.

RESULTS

Sixty-three journal articles, published between 1995 and 2004, met the inclusion criteria. Median sensitivity and specificity were 0,96 and 0,85 among smear positive and 0,66 and 0,98 among smear negative samples. The number of culture media used as reference test, the inclusion of bronchial samples and the TB prevalence were found to influence the reported accuracy. The test type had no impact on diagnostic odds ratio, but seemed to be correlated with sensitivity or specificity, probably via a threshold effect.

CONCLUSIONS

Currently, commercial NAATs can be confidently used 1) to exclude TB in patients with smear-positive samples in which environmental mycobacteria infection is suspected and 2) to confirm TB in a proportion of smear-negative cases. The methodological characteristics of primary studies determine considerable and meaningful changes in the reported diagnostic accuracy.

INTRODUCTION

In spite of their theoretical ability to detect even a single mycobacterial cell, nucleic acid amplification tests (NAATs) are not sufficiently reliable to replace conventional diagnostic methods for pulmonary tuberculosis (TB). Both inherent test characteristics and errors in testing procedures may account for their inaccuracy[1]. As for microscopy and culture, the key factor in determining NAAT false negatives is the density of mycobacteria in the specimen, since it can result in the absence of organisms in the small volumes sampled for the test. Furthermore, the presence in respiratory secretions of enzymes capable of inhibiting amplification reactions accounts for an additional 3% to 25% of false negative results[2]. On the other hand, false positives arise most often from contamination of negative samples with either organisms or target DNA from samples containing large numbers of mycobacteria or from amplicons contaminating the laboratory room[2,3].

To overcome these problems, the industry developed automated commercial systems which were made more robust by means of the use of standardized procedures and reagents for sample processing, amplification and detection. These procedures, which allowed different steps of the process to take place in a single sealed tube, were intended to reduce the risk of contamination. At the same time, the use of larger sample volumes or the introduction of internal amplification controls to detect inhibitors was adopted to cut down on false negative rates.

Notwithstanding these precautionary measures, published studies show a considerable heterogeneity in the results obtained with commercial NAATs. The U.S. Centers for Disease Control (CDC) recommend that commercial NAATs be used besides microscopy to improve diagnostic certainty (pending culture results and/or patient's response to therapy) and that clinicians rely upon clinical judgment in the interpretation of results. According to the CDC, the diagnosis of pulmonary TB can be presumed in smear positive (AFB+) patients with a positive NAAT result and in smear negative (AFB-) patients with two subsequent positive NAAT results. An environmental mycobacterial disease can be hypothesized when a negative NAAT is obtained from an AFB+ and inhibitor-free sample, while, as about 20% of TB cases can be attributed to infection by AFB- patients[4], two negative NAAT from two separate AFB- samples are needed to exclude contagiousness[5].

Two previous meta-analyses on the accuracy of NAAT for the diagnosis of pulmonary TB, analyzing mostly home-grown PCR-based tests, found a substantial variability in both sensitivity and specificity, due to the different threshold set by each investigator and to differences in study design and quality[6,7]. To our knowledge, the diagnostic accuracy of commercial NAAT separately on both AFB+ and AFB- respiratory samples has never been systematically reviewed. This study is designed to assess their performance in the context of a careful analysis of the impact of the type of test as well as of the methodological and clinical characteristics of published studies on the accuracy estimates.

METHODS

We searched Medline until July 1st 2005 and Embase until March 1st 2005, using a search strategy designed to identify studies evaluating commercial NAATs use for pulmonary TB diagnosis. We screened the titles and abstracts of the identified citations and scrutinized the references listed in the retrieved articles, considering any citation that did not obviously fail the inclusion criteria.

After a preliminary analysis of a sample of articles, we considered eligible for inclusion in our meta-analysis studies that 1) examined commercial NAAT diagnostic performance on respiratory samples (<5% of non respiratory samples was tolerated), 2) used Mycobacterium tuberculosis (MTB) culture of the same sample as reference standard for the diagnosis of pulmonary TB, 3) reported primary data sufficient for separately calculating both sensitivity and specificity in AFB+ and/or in AFB- specimens and 4) were written in English.

Reasons for article exclusion were: 1) reporting sensitivity and specificity "revised" by means of discrepant analysis as the only study results; in the case of studies re-testing the samples on the basis of discrepant analysis, only the initial "unrevised" results were considered. 2) Possible duplicate publication: when an author or a research group published more than one study, the existence of overlapping study populations was ascertained by checking sample recruitment sites and/or periods, or, if these were not available, contacting authors for clarification. If this was not provided, only the article reporting on the largest number of samples was included. 3) Application of commercial NAATs on gastric aspirates (>5% of total study sample) and 4) analysis of previous versions of commercial NAATs.

Two investigators independently evaluated studies for inclusion and abstracted relevant data. Disagreements were reconciled by consensus.

DATA EXTRACTION AND QUALITY ASSESSMENT

Data were abstracted using two separate data sheets: one for AFB+ and one for AFB- samples. Information recorded were: descriptive data (author name, journal, publication year), type of respiratory sample, prevalence of MTB culture positive samples, testing procedures for commercial NAATs, culture and AFB staining and commercial NAATs sensitivity and specificity.

According to established methodological standards for evaluation of diagnostic tests[8], four aspects of study quality were examined: cohort assembly (population of recruitment, method of sample selection, data collection modality), technical quality of reference test (the use of at least two different culture media was considered a more reliable reference test), blinding and study population features (clinical/demographic characteristics, pulmonary TB severity and other diagnoses in non- pulmonary TB patients). The original studies in which data were collected (or primary studies) were classified according to whether each characteristic was present, absent or unknown. In five multicenter studies, the participating laboratories used different AFB staining and/or culture procedures: these items were scored as unknown. Three studies included a separate description of a subgroup of patients on antituberculous therapy: these data were not included in our analysis and the studies were scored as not including patients under treatment.

STATISTICAL ANALYSIS

All statistical analyses were separately performed for AFB+ and AFB- samples. For each study, we classified commercial NAAT results as true positives (TP), false negatives (FN), false positives (FP) and true negatives (TN) as determined by comparison with MTB culture results. Then, we calculated the true positive rate

(TPR=TP/TP+FN=sensitivity), the true negative rate (TNR=TN/FP+TN=specificity), their odds (odds TPR=TPR/1-TPR and oddsTNR=TNR/1-TNR) and the diagnostic odds-ratio, i.e. the ratio of the odds of a positive NAAT among MTB culture positive compared to MTB culture negative samples (DOR=oddsTPR/oddsFP rate). Thus, DOR values of >1 indicate good test performance, while DOR values of <1 a test more frequently positive on control subjects (DOR=1 means that the test has not discriminating ability).

The potential problems in odds calculations associated with sensitivities and/or specificities of 100% were solved by adding 0.5 to zero values[9]. In articles where two or more different commercial NAATs were analyzed on the same samples, both extraction of data and calculation of accuracy measures were performed by considering them as separate studies.

To delineate the impact of study characteristics on diagnostic accuracy estimates, we fitted a multivariate random-effect regression model using DOR as dependent variable and study characteristics as explanatory variables ("Metareg" in Stata 8). Since each commercial NAAT fixes a well defined numeric value as positivity criterion, we took into account the threshold differences between studies by simply adding the test type as covariate in the regression model[9].

However, clinical interpretation of DOR is not easy, as the same values can relate to different combinations of sensitivity and specificity[10]. The use of fixed thresholds allowed us to explore the impact of the study characteristics (including the different thresholds chosen) on sensitivity and specificity separately; thus, we constructed two further regression models using, as dependent variables, oddsTPR and oddsTNR, respectively. For all the models, the dependent variables were included after logarithmic transformation.

As explanatory variables, we added to the regression models the clinical and methodological characteristics of primary studies. Since unreported items can reflect true methodological flaws or poor reporting of a methodologically sound study, we included only variables with a percentage of unreported items of <15%. As it is known that sensitivity and specificity vary with disease prevalence when an imperfect standard is used to evaluate a test, we added to the models the proportion of MTB culture positive samples (among AFB+ or AFB- samples), as a proxy of the true pulmonary TB prevalence[11,12].

The within-study variance was considered by taking weights equal to the inverse of the variance of the appropriate proportions; the between–study variance was estimated using restricted maximum likelihood estimate (REML)[13].

We assessed the possibility of publication bias by evaluating funnel plot for asymmetry, Begg's adjusted rank correlation test and Egger's regression asymmetry test ("Metabias" in Stata 8).

Finally, we applied Bayes theorem to assess the changes in probability of pulmonary TB determined by the use of commercial NAATs.

RESULTS

STUDY DESCRIPTION AND SYNTHESIS

The study selection process, that is entirely reported in the Appendix, lead to the inclusion of 63 journal articles, published between 1995 and 2004[14-76]. Since 8 articles analyzed 2 different commercial NAATs, 71 studies on the whole were available for analysis. The commercial NAATs evaluated were: Roche Amplicor MTB (25 studies), its entirely automated version, Cobas Amplicor MTB (10), E-MTD (14), BDProbeTecET (12), LCx (10). Overall, the 63 articles examined 51,160 samples: 5,729 MTB culture positive and 45,431 MTB culture negative. The median number of samples per study was 410 (IQR 247 to 662), with a median pulmonary TB prevalence of 0,14 (IQR 0,07-0,3).

Fifty-six articles analyzed both sensitivity and specificity of commercial NAATs on AFB+ samples. They included 3848 MTB culture positive and 1,535 MTB culture negative samples, with a median pulmonary TB prevalence of 0.77 (IQR 0.6-0.86). Five articles reviewed 2 commercial NAATs each, thus 61 studies were available. As shown in Figure 1a, sensitivity values were homogeneously elevated (0,96, 95% CI 0,956-0,968), whilst specificity was lower and extremely variable (0.85, 95% CI 0,84-0,87).

Table I: Pooled values* (95% confidence intervals) of DOR, sensitivity and specificity of the 5 commercial NAAT

Test	NAA method	AFB+			AFB-		
		DOR	sensitivity	specificity	DOR	sensitivity	specificity
Amplicor	PCR	117 (56-246)	0.96 (0.94-0.97)	0.83 (0.8-0.86)	77 (51-115)	0.61 (0.57-0.65)	0.97 (0.968-0.974)
Cobas Amplicor	PCR	99 (56-173)	0.96 (0.95-0.97)	0.74 (0.68-0.8)	220 (144-335)	0.64 (0.59-0.69)	0.993 (0.992-0,994)
BDP	SDA	181 (39-834)	0.98 (0.96-0.99)	0.89 (0.84-0.93)	96 (53-175)	0.71 (0.66-0.76)	0.97 (0.964-0.974)
E-MTD	TMA	314 (99-995)	0.97 (0.95-0.98)	0.96 (0.93-0.97)	157 (48-510)	0.76 (0.7-0.8)	0.97 (0.966-0.974)
LCx	LCR	42 (12-142)	0.96 (0.94-0.98)	0.71 (0.64-0.78)	71 (38-132)	0.57 (0.5-0.64)	0.98 (0.978-0.985)

PCR, polymerase chain reaction; SDA, strand displacement amplification; TMA, trascription-mediated amplification; LCR, ligase chain reaction; DOR, diagnostic odds ratio. *random effect model

The 60 articles examining commercial NAATs sensitivity and specificity on AFB- samples included 1,704 MTB culture positive and 43,852 MTB culture negative samples (median pulmonary TB prevalence 0.042, IQR 0.02-0.1). Eight articles reviewed 2 commercial NAATs each, bringing the total number of studies up to 68. The inspection of the forest plot in Figure 1b indicates a high specificity but a clear heterogeneity in sensitivity values. Pooled sensitivity and specificity were 0.66 (95% IC 0,63-0,68) and 0.98 (95% IC 0,978-0,981), respectively. Pooled values of DOR, sensitivity and specificity for each test type as well as their respective nucleic acid amplification techniques are reported in Table I.

The analysis of clinical and methodological characteristics of the primary studies (Table II) demonstrated that many studies did not comply with the published guidelines for conducting and reporting diagnostic test evaluation. As regard to MTB culture [most frequently Lowenstein-Jensen (68%) and Bactec 12B (52%)], we found that 11% of primary studies did not provide any description of the reference test used to assess pulmonary TB diagnosis, while approximately one quarter used only one culture medium. Even if more than half of the studies declared the enrollment of patients with suspected pulmonary TB, they often included samples from patients on anti-tuberculous treatment. Clinical spectrum of both pulmonary TB and comparative

groups was rarely illustrated and only 9 primary studies applied either single or double blinding for test interpretation.

Table II. Analysis of the methodological characteristics of the primary studies included in the meta-analysis

Characteristic	Number of studies (%)		
	All (n=63)	AFB+ samples (n=56)	AFB- samples (n=60)
Type of respiratory specimen			
Sputum	8 (13)	6 (11)	8 (13)
Bronchial secretions	1(2)	0	1(2)
Mixed respiratory secretions	54 (86)	50 (89)	50 (83)
AFB method			
Fluorescence	45 (71)	43 (77)	42 (70)
Carbolfuchsin	7 (11)	3 (5)	7 (12)
Unreported	11 (17)	10 (18)	11 (18)
Quality of reference test			
At least 2 culture media	41 (65)	38 (68)	38 (63)
One culture media	15 (24)	11 (20)	15 (25)
Unreported	7 (11)	7 (13)	7 (12)
Commercial NAAT used			
Amplicor	25	22	22
Cobas Amplicor	10	8	10
BDProbeTecET	12	10	12
E-MTD	14	12	14
LCx	10	9	10
Population of recruitment			
MTB culture or suspected PTB	29 (46)	27 (48)	27 (45)
High PTB suspicion	3 (5)	3 (5)	3 (5)
Suspected PTB or treatment monitoring	10 (16)	9 (16)	10 (17)
Other	5 (8)	4 (7)	4 (7)
Screening	3 (5)	3 (5)	3 (5)
Unreported	13 (21)	10 (18)	13 (22)
On anti-TB treatment			
Yes	31 (49)	29 (52)	31 (52)
No	15 (24)	12 (21)	14 (23)
Unreported	20 (32)	15 (27)	18 (30)
Method of sample selection			
Consecutive or random selection	20 (32)	17 (30)	20 (33)
Case control	5 (8)	5 (9)	4 (7)
Consecutive/case control	1 (2)	1 (2)	1 (2)
Unreported	37 (59)	33 (59)	35 (58)
Data collection modality			
Prospective	8 (13)	7 (13)	8 (13)
Retrospective	5 (8)	4 (7)	5 (8)
Prospective/retrospective	2 (3)	2 (4)	2 (3)
Unreported	48 (76)	43 (77)	45 (75)
Independence of observation			
Any blinding	10 (16)	7 (13)	9 (15)
Unreported	53 (84)	49 (88)	51 (85)
Clinical/demographic characteristics			
Reported	3 (5)	3 (5)	3 (5)
Unreported	60 (95)	53 (95)	57 (95)
Distribution of TB severity			
Results of quantitative culture reported	7 (11)	7 (13)	7 (12)
Unreported	56 (89)	49 (88)	53 (88)
Other diagnoses in the control group			
Reported	1 (2)	1 (2)	1 (2)
Unreported	62 (98)	55 (98)	59 (98)

EFFECT OF STUDY CHARACTERISTICS ON COMMERCIAL NAAT DIAGNOSTIC ACCURACY

The characteristics of the studies analyzing AFB+ and AFB- samples, are reported, respectively, in the last two columns of Table II. Those included in the meta-regression models as potential sources of heterogeneity were: quality of reference test, specimen type, commercial NAAT type and pulmonary TB prevalence. In Table

IIIa and IIIb, the resulting parameter estimates of these variables are presented as relative odds. Relative odds indicate the diagnostic performance of commercial NAATs in studies with that characteristic, relative to their performance in studies without that characteristic.

Table IIIa reports the relative odds for DOR, oddsTPR and oddsTNR of commercial NAATs on AFB+ samples. The studies using at least two MTB culture media and those including bronchial specimens yielded DOR values approximately 8 times higher than those using one culture media and 6 times higher than those analyzing sputum specimens only, mainly due to an effect on oddsTNR. OddsTNR values were inversely correlated with pulmonary TB prevalence and were significantly lower in studies of the LCx test with respect to studies analyzing E-MTD.

With regard to AFB- samples (Table IIIb), relative DOR of studies using at least two MTB culture media were more than two times higher than those of the studies using only one medium, mainly due to the increase in oddsTNR. The inclusion of bronchial specimens was also associated with increased oddsTNR values. In

Table III. Effect of the study characteristics on estimates of DOR, sensitivity and specificity, as determined by meta-regression analysis

a) AFB+ SAMPLES

Study characteristic	relative DOR (95% CI)	p	relative oddsTPR (95% CI)	p	relative oddsTNR (95% CI)	p
At least 2 culture media used	8.13 (2.59 - 25.49)	0.000	1.64 (0.90 - 3.00)	0.106	2.82 (1.04 - 7.68)	0.042
PTB prevalence	0.09 (0.01 - 1.72)	0.111	0.61 (0.12 - 3.10)	0.551	0.08 (0.01 - 0.96)	0.046
Inclusion of bronchial samples	6.67 (1.36 - 32.63)	0.019	0.66 (0.22 - 1.97)	0.495	5.40 (1.55 -18.79)	0.008
Amplicor	0.87 (0.23 - 3.32)	0.838	1.08 (0.56 - 2.09)	0.812	0.54 (0.17 - 1.73)	0.297
Cobas Amplicor	0.52 (0.11 - 2.43)	0.405	1.22 (0.56 - 2.6)	0.620	0.43 (0.11 - 1.72)	0.234
BDProbeTecET	0.84 (0.17- 4.06)	0.831	2.20 (0.91 - 5.30)	0.080	0.57 (0.15 - 2.25)	0.424
LCx	0.24 (0.05 - 1.07)	0.062	1.06 (0.48 - 2.32)	0.893	0.20 (0.05 - 0.75)	0.017

b) AFB- SAMPLES

Study characteristic	relative DOR (95% CI)	p	relative oddsTPR (95% CI)	p	relative oddsTNR (95% CI)	p
At least 2 culture media used	2.26 (1.13 – 4.53)	0.021	0.97 (0.65 - 1.47)	0.895	2.77 (1.63 - 4.73)	0.000
PTB prevalence	0.02 (0.00 – 0.66)	0.028	0.45 (0.10 - 2.16)	0.322	0.01 (0.00 - 0.23)	0.003
Inclusion of bronchial samples	1.79 (0.68 – 4.70)	0.238	0.63 (0.35 - 1.16)	0.138	2.58 (1.24 - 5.37)	0.011
Amplicor	0.91 (0.36 – 2.27)	0.839	0.47 (0.28 - 0.79)	0.005	1.90 (0.93 - 3.89)	0.081
Cobas Amplicor	1.67 (0.58 – 4.82)	0.342	0.63 (0.35 - 1.13)	0.119	2.48 (1.07 - 5.76)	0.034
BDProbeTecET	0.89 (0.34 – 2.32)	0.815	0.79 (0.45 - 1.38)	0.404	1.18 (0.56 - 2.48)	0.670
LCx	0.61 (0.22 – 1.68)	0.338	0.44 (0.24 - 0.81)	0.008	1.20 (0.55 - 2.59)	0.649

Results are expressed as relative odds and 95% confidence intervals.
The coding used in multiple regression analysis was: sample type, inclusion of bronchial specimens= 1, sputum only= 0; MTB culture, at least two media used=1, no information on culture media used=1 (not reported in the table), one culture medium used =0; cNAAT type, Amplicor=1, Cobas Amplicor=1, BDP=1, LCx=1, E-MTD=0
DOR, diagnostic odds ratio; TPR, sensitivity; TNR, specificity

comparison with studies analyzing E-MTD, those using LCx or Roche Amplicor MTB provided lower oddsTPR, while those using Cobas Amplicor MTB yielded higher odds TNR. An inverse correlation between pulmonary TB prevalence and both DOR and oddsTNR values was also found.

Evaluation of publication bias showed that the Egger's test was significant both for studies on AFB+ samples (regression coefficient 1,14, p=0.011) and for AFB- samples (regression coefficient 0,97 p=0.022). The visual inspection of the two funnel plots also showed some asymmetry. Conversely, the Begg's test was not significant (see Appendix).

POST-TEST PROBABILITY OF PULMONARY TB

The changes in pulmonary TB likelihood after commercial NAAT performance are depicted, per all pre-test probabilities, in Figure 2 (a, AFB+ samples; b, AFB- samples). The top curves portray the positive predictive values, i.e. the probabilities of pulmonary TB after obtaining a positive commercial NAAT result; the bottom curves represent the inverse of the negative predictive values, i.e. the probabilities of pulmonary TB after a negative commercial NAAT result. For example, using E-MTD on an AFB- sample from a patient in which previous diagnostic information (history taking, clinical examination, imaging etc) indicated a probability of pulmonary TB of about 30%, a negative result would reduce the likelihood of pulmonary TB to about 10%, while a positive one would increase it to about 90%.

DISCUSSION

Since they require lower technical skills and shorter assay time with respect to the less expensive home grown tests, commercial NAATs have become the most frequently used molecular tests for laboratory diagnosis of pulmonary TB[77]. In this meta-analysis, we 1) calculated pooled estimates of their sensitivity and specificity (see Table I), 2) demonstrated that their reported accuracy is influenced by primary study characteristics and 3) analyzed to what degree or under what conditions they add information to the diagnostic work-up of pulmonary TB.

The reference test used for diagnosing pulmonary TB was shown to have the largest impact on accuracy, both for AFB+ and AFB- samples. Since the incorporation of one or more additional units of medium is known to reduce the false negatives of culture[78], it could, as a consequence, have determined an "artificial" improvement in commercial NAAT specificity, that is estimated on samples classified by culture as MTB-free.

The small number of studies using liquid media as the only reference test did not allow us to evaluate the independent effect on accuracy of their higher MTB recovery rates with respect to solid media[78].

The imperfect sensitivity of culture could also explain the variation of the specificity (and the DOR) of commercial NAATs with TB prevalence. At low prevalence, in fact, the number of samples containing MTB, but wrongly classified by culture as MTB-free, is likely to be small and commercial NAAT (pseudo) false positives likely to occur less frequently; on the contrary, at high prevalence, the higher number of (pseudo) false positives deteriorates specificity[12].

The higher accuracy in studies including bronchial samples, already reported in a previous meta-analysis on PCR-based NAATs[6], was mainly due to an increase in specificity. However, since the reported culture yield in bronchial samples varies from 12 to 87%[79,80], it is difficult to explain these data on the basis of the proportional agreement of positive and negative results between the two tests. Studies focused on diagnostic performance of both culture and commercial NAATs on different bronchial samples may help to clarify this issue in the future.

Although only the Amplicor (or Cobas Amplicor) and the E-MTD are currently approved by the United States FDA for clinical use, the test type did not seem to explain the heterogeneity of DOR in meta-regression. Interestingly, the studies evaluating E-MTD on AFB- samples yielded higher sensitivities and lower specificities with respect to those using Roche Amplicor MTB or Cobas Amplicor MTB (Table IIIb). The higher E-MTD sensitivity, the only FDA-approved NAAT for use on AFB- samples, could be due to kit features, such as the use of ribosomal RNA as target sequence (about two thousand copies in each MTB cell)[81], but our results suggest that E-MTD could employ a lower positivity criterion with respect to other commercial NAATs and that the differences observed could be partly due to a "threshold effect". The accuracy of E-MTD appeared to be higher than that of the recently withdrawn Abbott LCx, while no differences were noticeable with BDProbeTecET.

Regarding the diagnostic value of commercial NAATs in the evaluation of a patient with suspected pulmonary TB, we observed that, because of their very high sensitivity on AFB+ samples, commercial NAATs can be confidently used to "rule out" pulmonary TB in AFB+ patients (Figure 2a). Thus, particularly in settings where opportunistic infections are a concern, a negative inhibitor-free commercial NAAT in patients with AFB+

smears and suggestive radiographic abnormalities should direct suspicion towards an environmental mycobacteria pulmonary disease.

The more limited gain in likelihood of pulmonary TB after a positive result on AFB+ samples (particularly for some commercial NAAT, see Figure 2), seems to limit their use as confirmatory tests in these cases. Possibly, the elevated false positive rates of a number of studies are related to the inclusion of samples from patients under treatment. These studies tried to correct the errors deriving from the enrollment of an inadequate study population by applying discrepant analysis, a statistical ploy that attempts to correct sensitivity and specificity of a "new" test (that is supposed to be more accurate than the reference standard it is compared with) by involving an additional, more reliable test (clinical diagnosis of pulmonary TB). This procedure, by correcting the errors hidden among conflicting results of the "new" test and the standard and by disregarding concordant errors, leads to an overestimation of the test accuracy[82]. Thus, we elected to include only the "uncorrected" results, hence discussing the possible effects on accuracy of the presence of samples from treated patients. The unavailability of treatment data from 27% of primary papers prevented us from drawing conclusions by means of meta-regression. Nevertheless, the pooled specificity calculated on the subgroup of studies clearly stating the exclusion (206 samples) and, respectively, the inclusion (707 samples) of treated patients were 0.97 (95% IC 0.93-0.99) and 0.76 (95% IC 0.73-0.79), indicating the inclusion of treated patient samples as the main cause for reduced specificity on AFB+ samples.

In the case of a negative microscopy result, commercial NAATs are not sensitive enough to exclude the diagnosis of pulmonary TB and further diagnostic work-up remains mandatory in these patients. By contrast, their high specificity endows them with the ability to "rule in" pulmonary TB in about two thirds of the patients who will be recognized as MTB culture-positive only 2-8 weeks later (Figure 2b): based on the degree of suspicion, the clinician is allowed to initiate treatment or, having already begun it, is made more comfortable with its continuation. Furthermore and with regard to the risk assessment of infectivity, commercial NAATs, due to their higher sensitivity in comparison with microscopy, could guide the choice of which AFB- patients are to be segregated, especially in facilities where HIV-infected or other immunocompromised individuals are managed[83]**.** This use, however, is of limited value in patients already started on therapy, as the numbers of viable mycobacteria in the sputum are known to dramatically fall in the first few days of treatment[84].

This meta-analysis has limitations. First, our estimates of sensitivity and specificity of commercial NAATs are hindered by the poor quality of primary studies, a common problem in diagnostic meta-analyses. Furthermore, although the Egger's test may reveal artifactual correlations between DOR and its variance regardless of publication bias[85-86], the possibility that the studies included in our meta-analysis are a biased set cannot be ruled out. In spite of these drawbacks, we think that summary estimates of test performance are a more accurate guide for the physician than are the results of any one of the primary studies.

Second, our decision to use the specimen as the unit of analysis could have affected accuracy because of the possible inclusion of multiple paucibacillary specimens from AFB- patients. Nevertheless, we speculated that the use of the patient as unit of analysis could have determined even wider accuracy variations, since the number of specimens per patient varied both within and between primary studies and repetitive testing is known to improve sensitivity.

Third, we examined commercial NAAT accuracy in comparison with culture, without addressing the issue of microscopy- and culture-negative pulmonary TB cases, that are diagnosed on clinical grounds only. During the systematic review of literature we found only 6 studies and one FDA premarket approval application document confronting this problem[34,35,37,81,87-89]. Out of 69 specimens (52 patients), only 7 provided at least one specimen that tested positive with commercial NAATs, corresponding to a pooled sensitivity of 10% (95%IC 4-20%). Thus, our estimates of sensitivity on AFB- samples are probably inflated with respect to what can be seen in clinical settings.

Based on this systematic review, the clinical use of commercial NAATs should be limited to the exclusion of TB diagnosis in AFB+ patients with suspected NTM infection and to the confirmation of TB in a percentage of those providing AFB- samples. Further studies using rigorous methods -including careful control for treatment and use of single specimen per each patient- would be highly desirable to appreciate the operating characteristics of the commercial NAATs better. Their accuracy on the different bronchial specimens and on samples from patients with culture negative pulmonary TB are also important issues that remain to be addressed.

Competing interests: none declared

REFERENCES

1. Noordhoek GT, Van Embden JDA, Kolk AHJ. Questionable reliability of the polimerase chain reaction in the detection of Mycobacterium tuberculosis. *N Eng J Med* 1993; 329:2036
2. Richeldi L, Barnini S, Saltini C. Molecular diagnosis of tuberculosis. *Eur Respir J* Suppl. 1995;20:689-700.
3. Noordhoek GT, Mulder S, Wallace P, *et al*. Multicentre quality control study for detection of Mycobacterium tuberculosis in clinical samples by nucleic amplification methods. *Clin Microbiol Infect*. 2004;10(4):295-301
4. Behr MA, Warren SA, Salomon H, *et al*. Transmission of Mycobacterium tuberculosis from patients smear-negative for acid-fast bacilli. *Lancet* 1999;353:444–9
5. From the Centers for Disease Control and Prevention. Update: Nucleic acid amplification tests for tuberculosis. *JAMA* 2000;284(7):826
6. Sarmiento OL, Weigle KA, Alexander J, *et al*. Assessment by meta-analysis of PCR for diagnosis of smear-negative pulmonary tuberculosis. *J Clin Microbiol* 2003;41(7):3233-40
7. LL Flores, M. Pai, JM Colford, LW Riley. In-house nucleic acid amplification tests for the detection of Mycobacterium tuberculosis in sputum specimens: meta-analysis and meta-regression.BMC Microbiology 2005, 5:55
8. Bossuyt P, Reitsma JB, Bruns DE, *et al*. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* 2003;49(1):7-18
9. Irwig L, Macaskill P, Glasziou P, *et al*. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol* 1995;48:119-30
10. Deeks JJ. Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;323:157-62
11. Buck A, Gart J. Comparison of a screening test and a reference test in epidemiologic studies. *Am J Epidemiol* 1966;83:586-92
12. Boyko EJ. Meta-analysis of Pap test accuracy. *Am J Epidemiol* 1995;141(7):680-9
13. Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002;21:1559-73
14. Alcala L, Ruiz-Serrano MJ, Hernangomez S, *et al*. Evaluation of the upgraded amplified Mycobacterium tuberculosis direct test (Gen-Probe) for direct detection of Mycobacterium tuberculosis in respiratory and non-respiratory specimens. *Diagn Microbiol Infect Dis* 2001;41(1-2):51-6
15. Ausina V, Gamboa F, Gazapo E, *et al*. Evaluation of the semiautomated AbbotLCX Mycobacterium tuberculosis assay for direct detection of Mycobacterium tuberculosis in respiratory specimens. *J Clin Microbiol* 1997;35:1996-2002
16. Barrett A, Magee JG, Freeman R. An evaluation of the BDProbeTec ET system for the direct detection of Mycobacterium tuberculosis in respiratory samples. *J Med Microbiol* 2002;51(10):895-8
17. Beavis KG, Lichty MB, Jungking DL, *et al*. Evaluation of AMPLICOR PCR for direct detection of Mycobacterium tuberculosis from sputum specimens. *J Clin Microbiol* 1995;33:2582-6

18. Bennedsen J, Thomsen VO, Pfyffer GE, *et al*. Utility of PCR in diagnosing pulmonary tuberculosis. *J Clin Microbiol* 1996;34(6):1407-11
19. Bergmann JS, Keating WE, Woods GL. Clinical evaluation of the BDProbeTec ET system for rapid detection of Mycobacterium tuberculosis. *J Clin Microbiol* 2000;38(2):863-5
20. Bergmann JS, Woods GL. Clinical evaluation of the Roche AMPLICOR PCR Mycobacterium tuberculosis test for detection of M. tuberculosis in respiratory specimens. *J Clin Microbiol* 1996;34(5):1083-5
21. Bergmann JS, Yuoh G, Fish G, *et al*. Clinical evaluation of the enhanced Gen-Probe Amplified Mycobacterium tuberculosis Direct Test for rapid diagnosis of tuberculosis in prison inmates. *J Clin Microbiol* 1999;37:1419-25
22. Bodmer T, Mockl E, Muhlemann K, *et al*. Improved performance of Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test when 500 instead of 50 microliters of decontaminated sediment is used. *J Clin Microbiol* 1996;34(1):222-3
23. Bogard M, Vincelette J, Antinozzi R, *et al*. Multicenter study of a commercial, automated polymerase chain reaction system for the rapid detection of Mycobacterium tuberculosis in respiratory specimens in routine clinical practice. *Eur J Clin Microbiol* Infect Dis 2001;20(10):724-31.
24. Cartuyvels R, De Ridder C, Jonckheere S, *et al*. Prospective clinical evaluation of Amplicor Mycobacterium tuberculosis PCR test as a screening method in a low-prevalence population. *J Clin Microbiol* 1996;34(8):2001-3
25. Chedore P, Jamieson FB. Routine use of the Gen-Probe MTD2 amplification test for detection of Mycobacterium tuberculosis in clinical specimens in a large public health mycobacteriology laboratory. *Diagn Microbiol Infect Dis* 1999;35(3):185-91
26. Cleary TJ, Roudel G, Casillas O, *et al*. Rapid and specific detection of Mycobacterium tuberculosis by using the Smart Cycler instrument and a specific fluorogenic probe. *J Clin Microbiol* 2003;41(10):4783-6
27. Coll P, Garrigo M, Moreno C, *et al*. Routine use of Gen-Probe Amplified Mycobacterium Tuberculosis Direct (MTD) test for detection of Mycobacterium tuberculosis with smear-positive and smear-negative specimens. *Int J Tuberc Lung Dis* 2003;7(9):886-91
28. Dalovisio JR, Montenegro-James S, Kemmerly SA, *et al*. Comparison of the Amplified M.tuberculosis (MTB)direct test, Amplicor MTB PCR and IS6110-PCR for detection of MTB in respiratory specimens. *Clin Infect Dis* 1996;23:1099-106
29. D'Amato RF, Wallman AA, Hochstein LH, *et al*. Rapid diagnosis of pulmonary tuberculosis by using Roche AMPLICOR Mycobacterium tuberculosis PCR test. *J Clin Microbiol* 1995;33(7):1832-4
30. De La Calle IJ, De La Calle MAJ, Rodriguez-Iglesias M. Evaluation of the BDProbeTec et system as screening tool in the direct detection of mycobacterium tuberculosis complex in respiratory specimens. *Diagn Microbiol Infect Dis* 2003;47(4):573-8
31. Denis O, Devaster JM, Vandenberg O, *et al.* Evaluation of ligase chain reaction for direct detection of Mycobacterium tuberculosis in respiratory specimens. *Zentralbl Bakteriol* 1998;288(1):59-65

32. Dilworth JP, Goyal M, Young DB, *et al*. Comparison of polymerase chain reaction for IS6110 and Amplicor in the diagnosis of tuberculosis. *Thorax* 1996;51(3):320-2
33. Gallina M, Troupioti P, Rocco G, *et al*. Predicting culture results for Mycobacterium tuberculosis complex. Amplified mycobacterium tuberculosis direct test and acid-fast bacilli microscopy. *Chest* 2000;118(1):28-32
34. Gamboa F, Fernandez G, Padilla E, *et al*. Comparative evaluation of initial and new versions of the Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test for direct detection of Mycobacterium tuberculosis in respiratory and nonrespiratory specimens. *J Clin Microbiol* 1998;36(3):684-9
35. Gamboa F, Manterola JM, Lonca J, *et al*. Comparative evaluation of two commercial assays for direct detection of Mycobacterium tuberculosis in respiratory specimens. Eur *J Clin Microbiol* Infect Dis 1998;17(3):151-7
36. Gurkan O, Acican T, Gulbay B. Evaluation of an amplified Mycobacterium tuberculosis direct test in clinical specimens. *Int J Tuberc Lung Dis* 2002;6(6):538-41
37. Ichiyama S, Iinuma Y, Tawada Y, *et al*. Evaluation of Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test and Roche PCR-microwell plate hybridization method (AMPLICOR MYCOBACTERIUM) for direct detection of mycobacteria. *J Clin Microbiol* 1996;34(1):130-3
38. Iinuma Y, Senda K, Fujihara N, *et al*. Comparison of the BDProbeTec ET system with the Cobas Amplicor PCR for direct detection of Mycobacterium tuberculosis in respiratory samples. *Eur J Clin Microbiol Infect Dis* 2003;22(6):368-71
39. Iwamoto T, Sonobe T, Hayashi K. Loop-mediated isothermal amplification for direct detection of Mycobacterium tuberculosis complex, M. avium, and M. intracellulare in sputum samples. *J Clin Microbiol* 2003;41(6):2616-22
40. Jan IS, Hsueh PR, Teng LJ, *et al*. Evaluation of an automatic polymerase chain reaction assay for identification of Mycobacterium tuberculosis in respiratory specimens. *J Formos Med Assoc* 1998;97(3):204-9
41. Johansen IS, Thomsen VO, Johansen A, *et al*. Evaluation of a new commercial assay for diagnosis of pulmonary and nonpulmonary tuberculosis. *Eur J Clin Microbiol Infect Dis* 2002;21(6):455-60
42. Jonsson B, Ridell M. The Cobas Amplicor MTB test for detection of Mycobacterium tuberculosis complex from respiratory and non-respiratory clinical specimens. *Scand J Infect Dis* 2003;35(6-7):372-7
43. Kearns AM, Freeman R, Steward M, *et al*. A rapid polymerase chain reaction technique for detecting M tuberculosis in a variety of clinical specimens. *J Clin Pathol* 1998;51(12):922-4
44. Kim SY, Park YJ, Kang SJ, *et al*. Comparison of the BDProbeTec ET system with the roche COBAS AMPLICOR System for detection of Mycobacterium tuberculosis complex in the respiratory and pleural fluid specimens. *Diagn Microbiol Infect Dis* 2004;49(1):13-8
45. Lebrun L, Mathieu D, Saulnier C, Nordmann P. Limits of commercial molecular tests for diagnosis of pulmonary tuberculosis. *Eur Respir J* 1997;10(8):1874-6

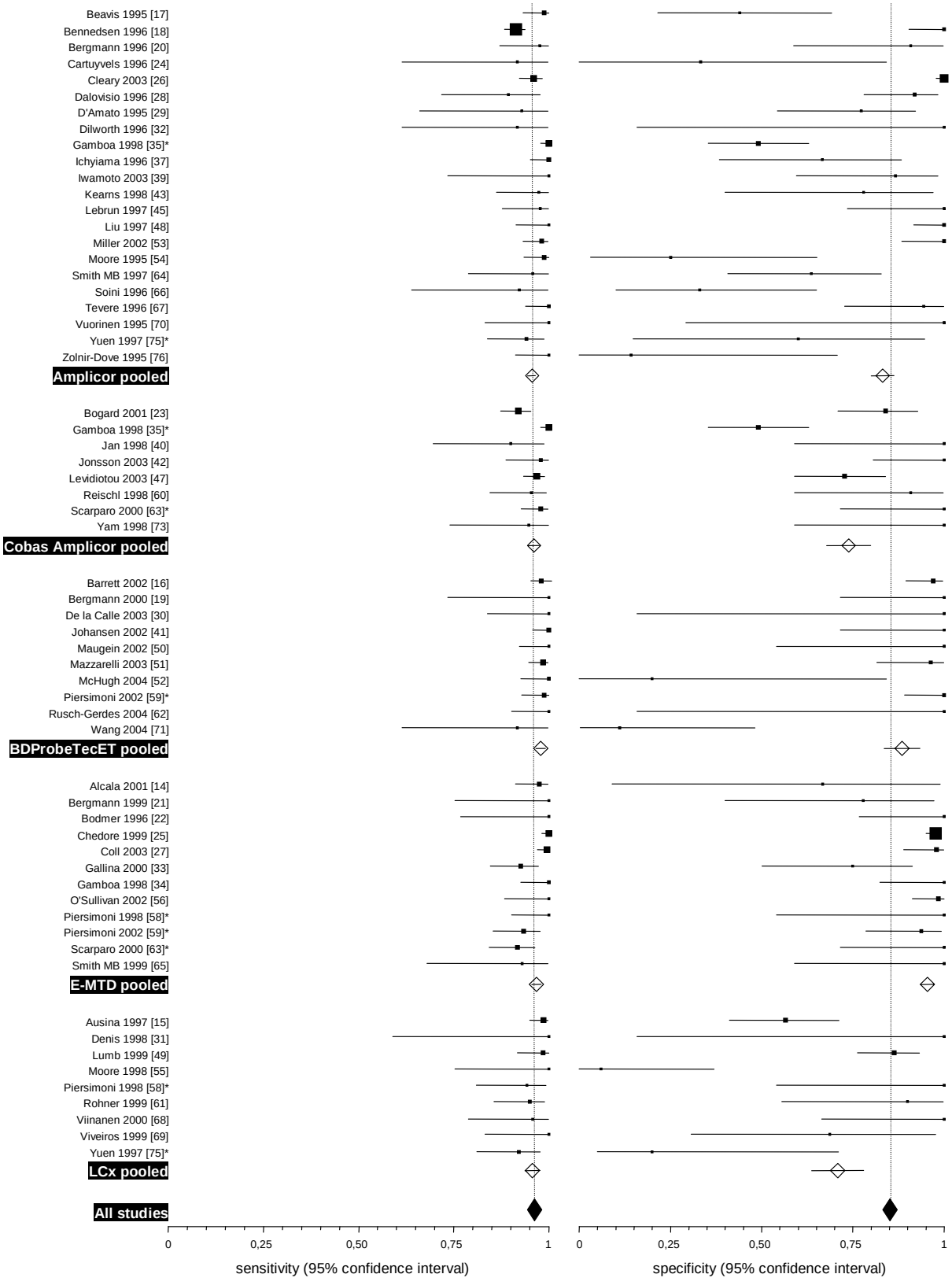
46. Lemaitre N, Armand S, Vachee A, *et al*. Comparison of real-time PCR method and the Gen-Probe Amplified Mycobacterium tuberculosis direct test for detection of Mycobacterium tuberculosis in pulmonary and nonpulmonary specimens. *J Clin Microbiol* 2004;42(9):4307-9
47. Levidiotou S, Vrioni G, Galanakis E, *et al*. Four-year experience of use of the Cobas Amplicor system for rapid detection of Mycobacterium tuberculosis complex in respiratory and nonrespiratory specimens in Greece. Eur *J Clin Microbiol* Infect Dis 2003;22(6):349-56
48. Liu D, Jones SL, Baird R, *et al*. Use of Roche Amplicor and multiplex PCR for diagnosis of human mycobacterial infections. *Opportunistic Pathogens* 1997;9(1):51-5
49. Lumb R, Davies K, Dawson D, *et al*. Multicenter evaluation of the Abbott LCx Mycobacterium tuberculosis ligase chain reaction assay. *J Clin Microbiol* 1999;37(10):3102-07
50. Maugein J, Fourche J, Vacher S, *et al*. Evaluation of the BDProbeTec ET DTB assay(1) for direct detection of Mycobacterium tuberculosis complex from clinical samples. *Diagn Microbiol Infect Dis* 2002;44(2):151-5
51. Mazzarelli G, Rindi L, Piccoli P, *et al*. Evaluation of the BDProbeTec ET system for direct detection of Mycobacterium tuberculosis in pulmonary and extrapulmonary samples:a multicenter study. *J Clin Microbiol* 2003;41(4):1779-82
52. Mc Hugh TD, Pope CF, Ling CL, *et al*. Prospective evaluation of BDProbeTec strand displacement amplification (SDA) system for diagnosis of tuberculosis in non-respiratory and respiratory samples. *J Med Microbiol* 2004;53:1215-9
53. Miller N, Cleary T, Kraus G, *et al*. Rapid and specific detection of Mycobacterium tuberculosis from acid-fast bacillus smear-positive respiratory specimens and BacT/ALERT MP culture bottles by using fluorogenic probes and real-time PCR. *J Clin Microbiol* 2002;40(11):4143-7
54. Moore DF, Curry JI. Detection and identification of Mycobacterium tuberculosis directly from sputum sediments by AMPLICOR PCR. *J Clin Microbiol* 1995;33:2686-91
55. Moore G, Curry JI. Detection and identification of Mycobacterium tuberculosis directly from sputum sediments by ligase chain reaction. *J Clin Microbiol* 1998;36:1028-31
56. O'Sullivan CE, Miller DR, Schneider PS, *et al*. Evaluation of Gen-Probe amplified mycobacterium tuberculosis direct test by using respiratory and nonrespiratory specimens in a tertiary care center laboratory. *J Clin Microbiol* 2002;40(5):1723-7
57. Piersimoni C, Callegaro A, Nista D, *et al*. Comparative evaluation of two commercial amplification assays for direct detection of Mycobacterium tuberculosis complex in respiratory specimens. *J Clin Microbiol* 1997;35(1):193-6
58. Piersimoni C, Callegaro A, Scarparo C, *et al*. Comparative evaluation of the new Gen-Probe Mycobacterium tuberculosis Amplified Direct Test and the semiautomated Abbott LCx Mycobacterium tuberculosis assay for direct detection of Mycobacterium tuberculosis complex in respiratory and extrapulmonary specimens. *J Clin Microbiol* 1998;36:3601-4
59. Piersimoni C, Scarparo C, Piccoli P, *et al*. Performance assessment of two commercial amplification assays for direct detection of Mycobacterium tuberculosis complex from respiratory and extrapulmonary specimens. *J Clin Microbiol* 2002;40(11):4138-42.

60. Reischl U, Lehn N, Wolf H, *et al*. Clinical evaluation of the automated COBAS AMPLICOR MTB assay for testing repiratory and nonrespiratory specimens. *J Clin Microbiol* 1998;36:2853-60
61. Rohner P, Jahn EI, Ninet B, *et al*. Rapid diagnosis of pulmonary tuberculosis with the LCx Mycobacterium tuberculosis assay and comparison with conventional diagnostic techniques. *J Clin Microbiol* 1998;36(10):3046-7
62. Rusch-Gerdes S, Richter E. Clinical evaluation of the semiautomated BDProbeTec ET System for the detection of Mycobacterium tuberculosis in respiratory and nonrespiratory specimens. *Diagn Microbiol Infect Dis* 2004;48(4):265-70
63. Scarparo C, Piccoli P, Rigon A, *et al*. Comparison of enhanced Mycobacterium tuberculosis amplified direct test with Cobas Amplicor Mycobacterium tuberculosis assay for direct detection of Mycobacterium tuberculosis complex in respiratory and extrapulmonary specimens. *J Clin Microbiol* 2000;38:1559-62
64. Smith MB, Bergmann JS, Harris SL, *et al*. Evaluation of Roche Amplicor MTB assay for the detection of mycobacterium tuberculosis in sputum specimens from prison inmates. *Diagn Microbiol Infect Dis* 1997;27:113-6
65. Smith MB, Bergmann JS, Onoroto M, *et al*. Evaluation of the enhanced amplified Mycobacterium tuberculosis direct test for direct detection of Mycobacterium tuberculosis complex in respiratory specimens. *Arch Pathol Lab Med* 1999;123(11):1101-3
66. Soini H, Agha SA, El-Fiky A, *et al*. Comparison of amplicor and 32-kilodalton PCR for detection of Mycobacterium tuberculosis from sputum specimens. *J Clin Microbiol* 1996;34(7):1829-30
67. Tevere VJ, Hewitt PL, Dare A, *et al*. Detection of Mycobacterium tuberculosis by PCR amplification with pan-Mycobacterium primers and hybridization to an M. tuberculosis-specific probe. *J Clin Microbiol* 1996;34(4):918-23
68. Viinanen AH, Soini H, Marjamaki M, *et al*. Ligase chain reaction assay is clinically useful in the discrimination of smear-positive pulmonary tuberculosis from atypical mycobacterioses. *Ann Med* 2000;32(4):279-83
69. Viveiros M, Pinheiro S, Moreira P, *et al*. Evaluation of a commercial ligase chain reaction assay for the diagnosis of pulmonary and extra-pulmonary tuberculosis. *Int J Tuberc Lung Dis* 1999;3(6):508-14
70. Vuorinen P, Miettinen A, Vuento R, *et al*. Direct detection of Mycobacterium tuberculosis complex in respiratory specimens by Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test and Roche Amplicor Mycobacterium Tuberculosis Test. *J Clin Microbiol* 1995;33(7):1856-9
71. Wang JY, Lee LN, Chou CS, *et al*. Performance assessment of a nested-PCR assay (the rapid BAP-MTB) and the BDProbeTec ET system for detection of Mycobacterium tuberculosis in clinical specimens. *J Clin Microbiol* 2004;43 (10):4599-603
72. Wang SX, Tay L. Evaluation of three nucleic acid amplification methods for direct detection of Mycobacterium tuberculosis complex in respiratory specimens. *J Clin Microbiol*. 1999;37(6):1932-4
73. Yam WC, Yuen KY, Seto WH. Direct detection of Mycobacterium tuberculosis in respiratory specimens using an automated DNA amplification assay and a single tube nested polymerase chain reaction (PCR). *Clin Chem Lab Med* 1998;36(8):597-9

74. Yee YC, Gough A, Kumarasinghe G, *et al*. The pattern of utilisation and accuracy of a commercial nucleic acid amplification test for the rapid diagnosis of Mycobacterium tuberculosis in routine clinical practice. *Singapore Med J* 2002;43(8):415-20
75. Yuen KY, Yam WC, Wong LP, *et al*. Comparison of two automated DNA amplification systems with a manual one-tube nested PCR assay for diagnosis of pulmonary tuberculosis. *J Clin Microbiol* 1997;35(6):1385-9
76. Zolnir-Dovc M, Poljak M, Seme K, *et al*. Evaluation of two commercial amplification assays for detection of Mycobacterium tuberculosis complex in respiratory specimens. *Infection* 1995;23(4):216-21
77. Centers for Disease Control and Prevention. CDC M. tuberculosis nucleic acid amplification testing performance evaluation program. CDC, Atlanta GA;2004. Available at http://www.phppo.cdc.gov/mpep/pdf/mtb/0401naa.pdf (accessed September 10, 2005)
78. Pfyffer G, Welscher HM, Kissling P, *et al.* Comparison of the Mycobacteria Growth Indicator Tube (MGIT) with radiometric and solid culture for recovery of acid-fast bacilli. J Clin Microbiol 1997; 35 (2): 364–8
79. Kennedy DJ, Lewis WP, Barnes PF. Yield of bronchoscopy for the diagnosis of tuberculosis in patients with human immunodeficiency virus infection. *Chest* 1992 Oct;102(4):1040-4
80. Baughman RP, Dohn MN, Loudon RG, *et al*. Bronchoscopy with bronchoalveolar lavage in tuberculosis and fungal infections. *Chest* 1991;99:92-7
81. U.S. Food and Drug Administration. Center for Devices and radiological Health. Gen-Probe® Amplified™ Mycobacterium Tuberculosis Direct (MTD) Test -- P940034/S008. Part 2, pages 15-25. Available at: http://www.fda.gov/cdrh/pdf/P940034S008b.pdf (accessed September 10, 2005)
82. Hadgu A. Discrepant analysis: a biased and an unscientific method for estimating test sensitivity and specificity. *J Clin Epidemiol* 1999;52(12):1231-7
83. Joint Tuberculosis Committee of the British Thoracic Society. Control and prevention of tuberculosis in the United Kingdom: code of practice 2000. Thorax 2000; 55: 887-901
84. Jindani A, Aber VR, Edwards EA, *et al.* The early bactericidal activity of drugs in patients with pulmonary tuberculosis. Am Rev Respir Dis. 1980;121(6):939-49
85. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58 :882-93
86. Begg CB. Systematic reviews of diagnostic studies require study by study examination : first for heterogeneity and then for sources of heterogeneity. *J Clin Epidemiol* 2005;58:865-6
87. Jouveshomme S, Cambau E, Trystram D, *et al*. Clinical utility of an amplification test based on ligase chain reaction in pulmonary tuberculosis. *Am J Respir Crit Care Med* 1998;158(4):1096-101
88. Lim TK, Mukhopadhyay A, Gough A, et al. Role of clinical judgment in the application of a nucleic acid amplification test for the rapid diagnosis of pulmonary tuberculosis. Chest. 2003;124(3):902-8

89. Al Zahrani K, Al Jahdali H, Poirier L, et al. Yield of smear, culture and amplification tests from repeated sputum induction for the diagnosis of pulmonary tuberculosis. Int J Tuberc Lung Dis 2001;5 (9):855-60

Figure 1a. Studies on AFB+ samples



Legend to Figure 1
Individual study estimates of sensitivity and specificity of commercial NAAT for the diagnosis of pulmonary TB on a) AFB+ samples and b) AFB- samples. Pooled values were calculated using random effect model. Error bars represent 95% confidence intervals. Five articles on AFB+ samples and eight articles on AFB- samples analysing two commercial NAAt each are cited twice.

Figure 1b. Studies on AFB- samples

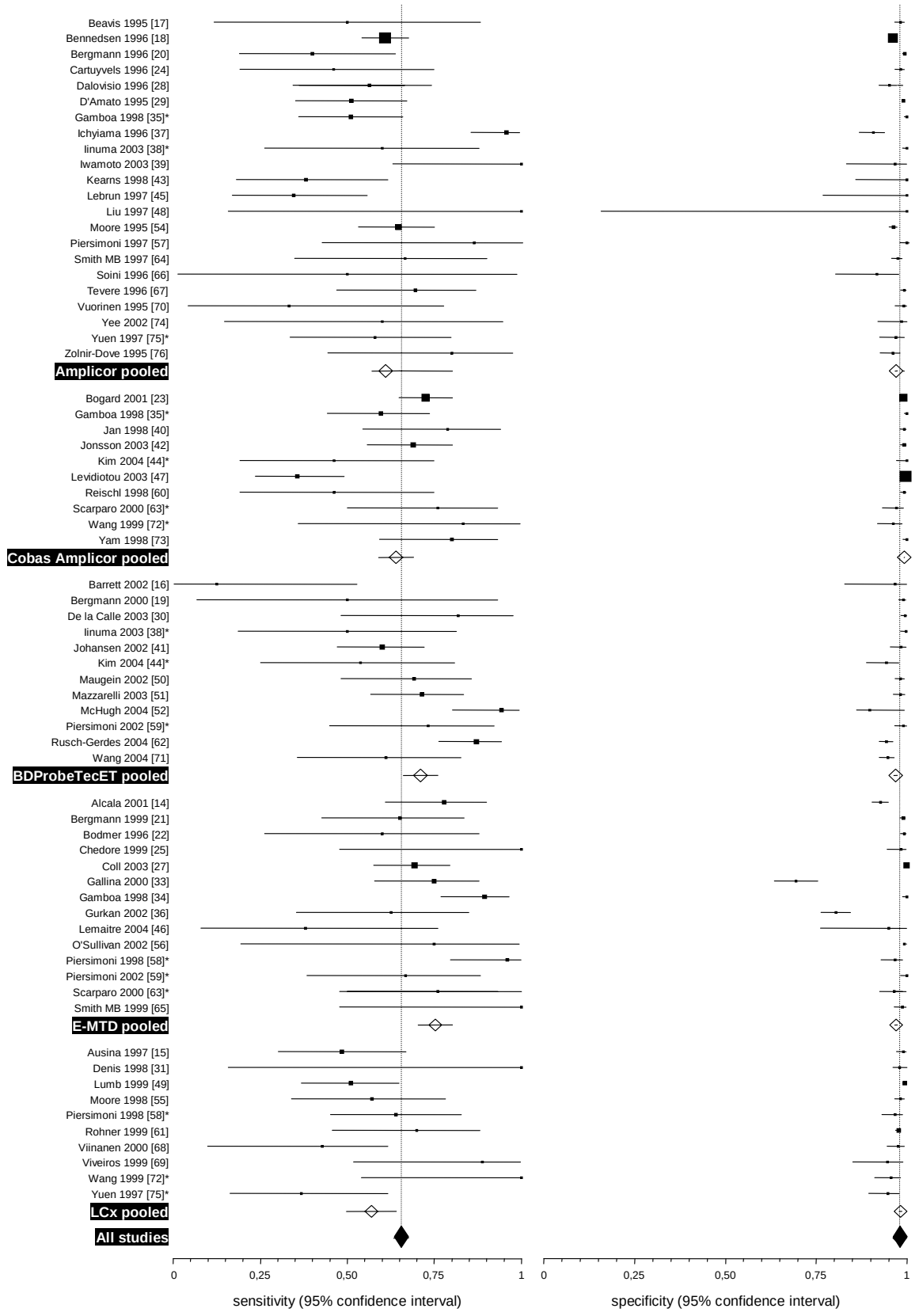
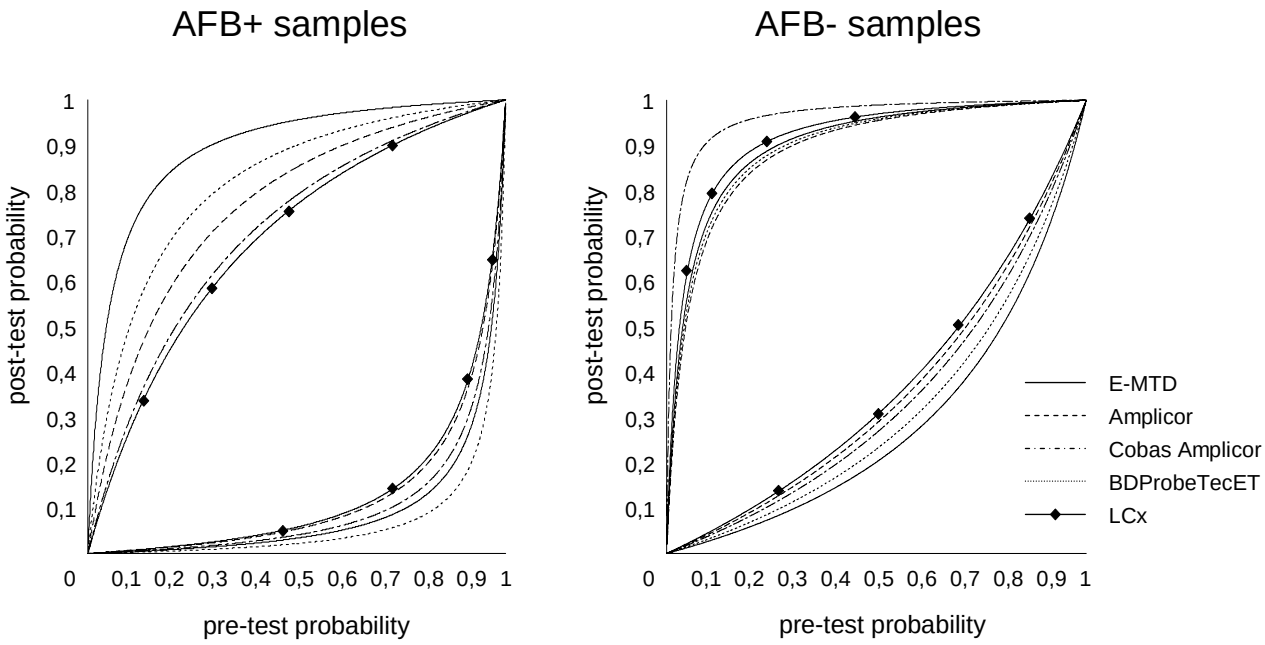


Figure 2



Legend to Figure 2.
Predictive values of pulmonary TB after carrying out commercial NAAT.
Post-test probabilities were calculated using pooled sensitivity and specificity values for each test type.
a), AFB+ samples; b), AFB- samples. Top curves: positive predictive values; bottom curves, 1-negative predictive values.