

**Diagnosing tuberculosis with urine lipoarabinomannan: systematic review and meta-analysis**

Jessica Minion<sup>1,2,3</sup>, Erika Leung<sup>1,2</sup>, Elizabeth Talbot<sup>4</sup>, Keertan Dheda<sup>5</sup>, Madhukar Pai<sup>1,2</sup>, Dick Menzies<sup>1,2</sup>

<sup>1</sup>Dept of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal, Canada

<sup>2</sup>Respiratory Epidemiology & Clinical Research Unit, Montreal Chest Institute, Montreal, Canada

<sup>3</sup>Dept of Medical Microbiology & Immunology, University of Alberta, Edmonton, Canada

<sup>4</sup>Foundation for Innovative New Diagnostics (FIND), Geneva, Switzerland

<sup>5</sup>Divisions of Pulmonary & Clinical Immunology, Dept of Medicine, University of Cape Town, Cape Town, South Africa

Abstract Word Count: 200

Word Count: 3001

Key Words: tuberculosis, diagnosis, lipoarabinomannan, urine, accuracy

Running Title: LAM systematic review and meta-analysis

Summary: The LAM urinary assay has characteristics which make it attractive for diagnosing active TB, but sensitivity is suboptimal for routine clinical use in unselected patients. These assays seem to perform better in HIV infected patients, especially those with severe immunodeficiency.

Corresponding Author:

Dr. Dick Menzies

Room K1.24 Montreal Chest Institute

3650 St. Urbain Street

Montreal, PQ, Canada, H2X 2P4

Tel: 1-514-934-1934 ext 32128

Fax: 1-514-843-2083

Email: dick.menzies@mcgill.ca

Alternate Corresponding Author:

Dr. Madhukar Pai

Dept of Epidemiology, Biostatistics & Occupational Health

McGill University  
1020 Pine Avenue West  
Montreal, PQ, Canada, H3A 1A2

Tel: 1-514-398-5422  
Fax: 1-514-398-4503  
Email: [madhukar.pai@mcgill.ca](mailto:madhukar.pai@mcgill.ca)

## **ABSTRACT**

**Background:** Lipoarabinomannan (LAM) is a potential marker of active TB. We performed a systematic review and meta-analysis regarding use of urinary LAM assays for diagnosing active TB.

**Methods:** We systematically searched for published and unpublished studies that evaluated urinary LAM for active TB diagnosis. Extracted data was pooled using bivariate random effects models and hierarchical summary receiver operating characteristic curves. Heterogeneity was explored through subgroup analysis and meta-regression. Quality was assessed according to standardized QUADAS criteria.

**Results:** In 7 studies that assessed test accuracy in microbiologically confirmed cases only, estimates of sensitivity ranged from 13% to 93%, while specificity ranged from 87 – 99%. In 5 studies that assessed accuracy in clinical and confirmed TB cases, sensitivity ranged from 8% to 80%, while specificity ranged from 88% to 99%. In five studies with results stratified by HIV status, sensitivity was 3-53% higher in HIV-positive than HIV-negative subgroups; sensitivity was highest with advanced immunosuppression.

**Conclusions:** The LAM urinary assay has several characteristics which make it attractive for diagnosing active TB, but has suboptimal sensitivity for routine clinical use. Further studies are needed to evaluate the potential value of the LAM assay in individuals with advanced HIV or for diagnosis of paediatric TB.

## **INTRODUCTION**

Despite the enormous global burden of tuberculosis (TB), case detection rates are low, posing serious hurdles for TB control (1). Diagnosis of active TB continues to rely on tests such as smear microscopy, culture and chest radiography. These tests have several limitations, are not point of care tests, and perform poorly in populations affected by the HIV epidemic (2).

As a result of recent efforts to develop new tools for TB diagnosis, several new diagnostics have been introduced and evaluated (2-4). However, an accurate point-of-care test that could be used within peripheral clinical settings with limited laboratory facilities, has not yet been successfully developed (5).

Lipoarabinomannan (LAM) is a structurally important 17.5kD heat-stable glycolipid found in the cell wall of *Mycobacterium tuberculosis* (MTB). LAM can account for up to 15% of the total bacterial weight and serves as an immunogenic virulence factor that is released from metabolically active or degrading bacterial cells during TB infection (6, 7). Detection of LAM antigens in urine has several potential advantages compared to currently used diagnostics. Urine samples are simple to collect, process, and store. There are far fewer infection control concerns compared to sputum. Urine is a particularly attractive specimen in young children who are often unable to produce sputum.

Based on early data, a prototype urinary LAM detection test was produced by Chemogen Inc (Portland, Maine, USA) and a commercial version of this test is now marketed as the Clearview™ TB ELISA (Inverness Medical Innovations Inc, Massachusetts, USA)(8). Several studies have evaluated LAM-based diagnostics, but to date there has been no systematic review on this topic. Hence, we planned a systematic review of the existing evidence base on accuracy of LAM antigen detection for diagnosis of active TB.

## **METHODS**

We followed a standard protocol for systematic reviews and meta-analyses (9), and used methods recommended by the Cochrane Diagnostic Test Accuracy Working Group (10).

### **Search Strategy**

We systematically searched 3 databases for relevant citations: PubMed, EMBASE and Web of Science (from earliest records to September 4, 2010). No language restrictions were imposed on

the search criteria. Reference lists from included studies were hand searched and experts and industry representatives were contacted to identify additional studies.

### **Eligibility Criteria**

Predetermined eligibility criteria included the use of any form of LAM detection in urine specimens in patients suspected or known to have active pulmonary TB, and the use of an accepted reference standard. Accepted reference standards included: positive culture of MTB, visualization of acid fast bacilli (AFB) from a clinical specimen (including histopathology), or positive nucleic acid amplification for MTB. Studies which used clinical information to contribute to the classification of patients were also included, and analyzed according to how that clinical information was used (see Outcome Measures and Subgroups).

### **Study Selection**

Screening of titles and abstracts was performed by one reviewer (JM) and selection of full text was performed by two independent reviewers (JM and EL). Articles retrieved for full text review along with reasons for exclusion are available from the authors.

### **Data Extraction**

We created and piloted a data extraction form with a subset of eligible studies. Based on experience gained from the pilot extractions, the form was finalized. All studies included in the final review were extracted independently by 2 reviewers (JM and EL) and any disagreements were resolved by consensus.

### **Assessment of Study Quality**

We used the QUADAS criteria to assess the quality of diagnostic accuracy studies (11). All criteria were classified as Yes, No, or Unclear based on information available in the publication. Studies were judged according to the data used for the meta-analysis, which may not have been all data available in the publication. This would apply if the study assessed the performance of the LAM assay in TB suspects as well as healthy controls; if possible data from healthy controls were excluded for the main analysis.

### **Outcome Measures and Subgroups**

Three definitions of reference standards were used according to how studies dealt with 'Reference Indeterminate' results. These patients had negative microbiologic results (i.e. smear and culture negative), but were considered 'clinical cases' of TB on the basis of clinical and radiographic features and/or response to treatment. *Analysis A* considered cases (reference

positive) to be ONLY patients with microbiologically-confirmed TB and all other patients were considered reference negative, EXCEPT the clinical cases which were excluded from the analysis. *Analysis B* defined ONLY patients with microbiologically-confirmed TB as reference positive, and ALL others were reference negative, meaning that clinical cases were considered NOT to have active TB. *Analysis C* considered patients with microbiologically-confirmed TB AND the clinical cases as reference positive (active TB), and all other patients were considered reference negative. In summary, in the three analyses the microbiologically-negative clinical cases were excluded (*A*), included as reference negative (*B*), or included as reference positive (*C*).

Pre-specified subgroup analyses were performed according to patient HIV status, whether urine was fresh or frozen prior to LAM assay testing, and the version of the LAM assay used. Subgroup analysis was planned for different CD4 count measurements within the HIV-positive subgroup. Meta-regression was performed relating the proportion of HIV-positive patients in a study (independent variable) to the resultant LAM sensitivity (dependent variable). The proportion of specimens testing positive by the LAM assay was compared between categories of varying TB likelihood and presumed bacillary burden: smear-positive specimens; culture positive specimens; smear-negative specimens; clinically diagnosed patients; suspects investigated for active TB; and healthy or not-at-risk controls.

### **Analysis**

Data were analyzed using STATA/IC 11.0 (Stata Corp, Texas, USA). Forest plots visually displaying sensitivity and specificity estimates and their 95% confidence intervals (CIs; using exact methods for proportions) from each study were constructed using MetaDiSc 1.4 software (12). Hierarchical summary receiver operating characteristic (HSROC) curves were analyzed to explore the influence of threshold effects and produce a global summary of test accuracy (13). Accuracy measures were pooled using bivariate random effects regression models (14), using the user-written program “metandi” and meta-regression was performed using “metareg” in STATA (15). Heterogeneity of accuracy estimates was assessed using the  $I^2$  statistic (16). Subgroups with fewer than 4 studies were combined using univariate random effects models because bivariate random effects regression models do not converge with small numbers of studies. The proportion of LAM positive specimens by likelihood category were also combined using univariate random effects models.

## **RESULTS**

Selection of included studies is summarized in Figure 1. We identified 3141 citations from database searches and an additional 11 through other sources. 1559 unique articles were left after excluding duplicate articles. Screening of titles and abstracts identified 25 potentially relevant articles which were then retrieved for full text review. Of these, 9 studies were considered eligible for this review.

### ***Characteristics of Included Studies***

Studies that met our selection criteria are described in Table 1.

Many studies reported data which could be extracted and analyzed in different ways according to how specimens from clinical cases were classified (see Outcome Measures and Subgroups). Out of the 9 total studies, 7 reported results that allowed an analysis which excluded clinical cases (Analysis A). Seven studies reported results that allowed an analysis using strict microbiologic criteria as the reference standard (Analysis B), and 5 studies reported results that could be analysed to include the clinical cases, along with microbiologically-confirmed cases, as active TB (Analysis C).

A summary of the quality of the studies, judged using the QUADAS criteria for diagnostic studies (11), is displayed graphically in Figure 2. Seven of the studies did not report adequate information to assess all 14 criteria. The included studies met between 4 and 14 of the 14 quality indicators.

### ***Accuracy Estimates***

When clinical cases were excluded (Analysis A), 7 studies reported a wide range of sensitivity estimates (13 – 93%), while specificity estimates were less variable (87 – 99%)(Figure 3). In the 7 studies that could be analysed to consider only microbiologically-confirmed cases as active TB (Analysis B), sensitivity estimates varied from 13% to 81%, while specificity estimates ranged from 79% to 100% (Supplement Figure 1). In the analysis combining clinical and microbiologic cases as active TB (Analysis C), sensitivity ranged from 8% to 80%, while specificity ranged from 88% to 99% in 5 studies (Supplement Figure 2).

When these studies were combined there was significant heterogeneity in nearly all of the pooled estimates; hence these should be interpreted with caution (Table 2). Using bivariate random effects models, pooled sensitivity estimates ranged between 34% (14, 62) and 60% (38, 79)

depending on the definition of the reference standard (Table 2). Sensitivity was highest when clinical cases were excluded from the analysis (Analysis A), and lowest if clinical cases were considered active TB (Analysis C). Pooled specificity estimates for Analyses A, B and C were 93% (88, 96), 93% (83, 97), and 94% (87, 98) respectively.

### ***Hierarchical Summary Receiver Operating Characteristic (HSROC) Curves***

Supplement Figures 3 – 5 plot the sensitivity (True Positive Rate) against 1-specificity (False Positive Rate) and partial curves are constructed from the bivariate random effects regression models used to calculate the pooled estimates. This visually demonstrates the large variation in sensitivity estimates between studies.

### ***Subgroup Analyses and Meta-Regression***

Five studies compared the performance of the LAM assay in HIV-positive patients to HIV-negative patients. All studies found improved sensitivity in HIV-positive populations (between 2.5% and 52.8% absolute increase in sensitivity), although 4 of the 5 also found small decreases in specificity (between 4.9% and 7.3% absolute decrease in specificity, with the fifth study reporting 0.9% absolute increase). Results that were stratified by HIV status were combined, again separating data depending on their definitions of reference standards (Table 2). In all analyses, the LAM assay had higher sensitivity in HIV-positive than HIV-negative patients with little difference in specificity. The largest difference was seen when strict microbiologic criteria were used as the reference standard (Analysis B): in HIV-negative patients LAM had an overall sensitivity of 18% (10, 29), compared to 56% (40, 71) in HIV-positive patients.

Three studies reported diagnostic accuracy based on CD4 counts within HIV-positive TB suspects. These were not meta-analyzed due to differences in CD4 cut-offs and different reference standards, however in all studies the LAM assays had higher sensitivity in patients with more severe immunosuppression (Table 3).

Analysis of sub-groups in which the urine used in the assay was either fresh or previously frozen found no statistically significant differences between these groups (Table 4). Pooled sensitivity estimates from the two studies evaluating the early non-commercial version of the test were significantly higher than estimates from subsequent studies using either of the commercial assays.

Meta-regression was performed using the overall proportion of HIV-positive TB suspects in each study as an independent predictor of the estimated overall sensitivity, weighted inversely by the standard error of the sensitivity and combined using random effects (Figure 4).

### ***Influence of Clinical Status on LAM Accuracy***

We estimated the pooled sensitivity of LAM in patient groups with different definitions of disease (smear-positive TB, all culture-positive TB, smear-negative TB, clinically positive TB) or suspicion of disease (symptomatic non-TB patients, healthy non-TB patients). As shown in Figure 5, the proportion of LAM positives increased with greater bacillary burden in TB cases and was also higher in TB suspects than healthy controls.

## **DISCUSSION**

The LAM urine assay was initially seen as a potentially revolutionary diagnostic for active TB. With its potential to be used as a simple point-of-care test, lack of biosafety concerns, and use of a non-invasive, convenient patient specimen, the LAM assay was fast-tracked for commercial development. Despite very positive initial evaluations, larger and more recent studies have failed to demonstrate adequate sensitivity for TB diagnosis under routine conditions in unselected patients.

Multiple explanations may contribute to this observation. The initial study conducted by Hamasur in 2001 (17) can be considered a proof-of-principle demonstration, so higher accuracy may be expected. Tessema (18) evaluated the original non-commercialized test methods described by Hamasur (17) and found higher sensitivity than most subsequent studies using the Chemogen or Inverness assays. The differences in the sensitivity of the different versions of the LAM assay, may result from simplification of the test for widespread use. However, methodological differences between the earlier and later studies, including differences in study design, patient population (including proportion of HIV-infected patients), disease severity, specimen handling and reference standards are very plausible explanations for the large range of sensitivity estimates reported.

The Chemogen and Inverness assays use the same polyclonal antibodies, but the manufacturing process is different. Issues such as pH, viscosity, matrix composition and spraying protocols could have impacted test performance. Additionally, preliminary evaluations have found improved sensitivity with a point-of-care lateral flow dipstick prototype (Determine TB,

Inverness Medical Innovations, Massachusetts, USA) compared to the Clearview TB ELISA in an HIV-infected patient population (19). Further evaluations of this new assay will be important, since the point-of-care format has the greatest potential to impact TB patient care.

This review does not exclude the possibility that the LAM assay is useful in the diagnosis of TB. The practical advantages of the assay are important and the sensitivity of the LAM assay might be superior to microscopy in HIV-positive patients who are more likely to have smear-negative TB and to have disseminated TB. Indeed, the Clearview™ TB ELISA (Inverness Medical Innovations Inc., Massachusetts, USA) is licensed for use only as a screening test in HIV-positive TB suspects (8) so evaluations of its use in HIV-negative populations must be considered 'off-label'. Although all studies in this review that reported results by HIV subgroup found higher sensitivity in HIV-positive populations, not all found greater yield with LAM than smear microscopy. The pooled results in Figure 5 show that only 57% of smear-positive specimens were LAM positive, suggesting that LAM would not be able to replace smear microscopy. However, we also found that LAM was positive in 41% of smear-negative specimens; this is a relatively large incremental yield compared to smear microscopy. This suggests that the LAM assay and smear microscopy may detect different groups of TB patients, and might be best used in combination, as was reported by Shah and colleagues (20).

There are several possible reasons to explain the higher sensitivity of LAM assays in immunosuppressed patients. One theory cites the correlation of higher sensitivity with greater bacillary burden (Figure 5)(21, 22), and assumes relatively greater multiplication of MTB bacilli in patients with impaired immune function. Alternatively, due to the general lack of cavity formation in immunosuppressed patients, the bacteria are forced to replicate in tissue which would facilitate the diffusion of shed LAM into the circulation. A second explanation is that a larger degree of antigen-antibody complex formation in TB patients without immune suppression interferes with LAM excretion in the urine (23). Finally, HIV-related podocyte dysfunction, which is more common in advanced HIV, can increase glomerular permeability (24) and might result in increased levels of LAM in patients' urine (19).

Peter and colleagues have proposed a clinical algorithm for the diagnostic work-up of TB suspects in low-resource settings with high TB and HIV prevalence (19). In this algorithm, the LAM assay would serve as a 'rule-in' test when screening patients at high-risk of smear-negative, HIV-associated TB. This role depends upon the LAM assay having a high and reliable

specificity. Although most studies have reported high predictive values for positive LAM results, and we found pooled specificity estimates  $\geq 90\%$ , there remains unexplained variability from individual studies (range 79 – 100%). While this may reflect a shortcoming of the assay, other explanations can be considered. It is possible that sub-clinical TB is being under-diagnosed, especially in HIV-infected patients, leading to misclassification bias. It is also possible that false-positive results are due to cross-reactivity with non-tuberculous mycobacteria (NTM) or other microorganisms. Indeed, Dheda and colleagues found significant LAM positivity in cultures containing common oral flora such as Actinobacteria and Candida species (25). The contamination of urine specimens with normal flora including Candida species, the contamination of reagents and non-sterile containers with NTM, and the colonization of patients with NTM may all lead to lower predictive values of positive LAM test results.

Finally, the benefits of using a non-invasive, easily collected specimen such as urine would be greatly appreciated in the diagnosis of paediatric TB. Considering the inadequate diagnostic yield of sputum-based diagnostics in young children, further evaluations of urinary LAM assays in paediatric populations is warranted.

This review had several strengths, including a broad and inclusive search of the published literature as well as efforts to identify currently unpublished studies. Selection of included studies and extraction of data was performed by 2 independent reviewers. Additionally, rigorous statistical methods were employed using bivariate random effects models and HSROC curves – as recommended by the Cochrane Diagnostic Test Accuracy Working Group for diagnostic meta-analyses (10). This review was limited by the small number of studies reporting evaluations of the LAM assay, especially for important subgroups. Most of the pooled estimates should be interpreted with caution in view of their significant heterogeneity.

We performed three analyses depending on how specimens from clinical cases were classified. Differences between the pooled estimates using the three approaches demonstrates how the definition of reference standard can influence the results of diagnostic studies. Although a strict microbiologic definition of TB is the most objective reference standard, nevertheless a certain number of patients will have active TB, despite negative cultures (26). Use of radiologic findings, clinical features, and treatment response are more subjective, and may lead to over-diagnosis of TB, but exclusion of these clinical cases may lead to overestimates of accuracy.

In conclusion, the LAM urine assay has many characteristics which make it a potentially useful rule-in TB diagnostic, but this review found inadequate sensitivity to use the LAM assay for the diagnosis of active TB in unselected cohorts. The assay performs better in HIV-infected patients, especially those with severe immunodeficiency, but even in HIV-infected persons the sensitivity is suboptimal. Further studies are warranted to evaluate the added value of the LAM assay in the diagnosis of active TB in individuals with advanced HIV and in children, as well as to assess newer versions of this test with technical advances compared to those reviewed here.

## REFERENCES

1. World Health Organization. Global Tuberculosis Control 2009: Epidemiology, Strategy, Financing. Geneva, Switzerland: World Health Organization.  
[http://www.who.int/tb/publications/global\\_report/2009/pdf/full\\_report.pdf](http://www.who.int/tb/publications/global_report/2009/pdf/full_report.pdf).
2. Perkins MD, Cunningham J. Facing the crisis: improving the diagnosis of tuberculosis in the HIV era. *J Infect Dis*. 2007 Aug 15;196 Suppl 1:S15-27.
3. Perkins MD, Roscigno G, Zumla A. Progress towards improved tuberculosis diagnostics for developing countries. *Lancet*. 2006 Mar 18;367(9514):942-3.
4. Pai M, O'Brien R. New Diagnostics for Latent and Active Tuberculosis: State of the Art and Future Prospects. *Semin Respir Crit Care Med*. 2008;29(5):560-68.
5. Dowdy DW. Cost-effectiveness of novel diagnostic tools for the diagnosis of tuberculosis. *Int J Tuberc Lung Dis*. [cost-effectiveness model]. 2008;12(9).
6. Hunter SW, Gaylord H, Brennan PJ. Structure and antigenicity of the phosphorylated lipopolysaccharide antigens from the leprosy and tubercle bacilli. *J Biol Chem*. 1986 Sep 15;261(26):12345-51.
7. Chan J, Fan XD, Hunter SW, Brennan PJ, Bloom BR. Lipoarabinomannan, a possible virulence factor involved in persistence of *Mycobacterium tuberculosis* within macrophages. *Infect Immun*. 1991 May;59(5):1755-61.
8. Inverness Medical Innovations. Clearview TB ELISA. [cited 2010 November 8]; Available from: [http://www.clearview.com/tb\\_elisa.aspx](http://www.clearview.com/tb_elisa.aspx).
9. Pai M, McCulloch M, Gorman JD, Pai N, Enanoria W, Kennedy G, et al. Systematic reviews and meta-analyses: an illustrated, step-by-step guide. *Natl Med J India*. [Instructive]. 2004 Mar-Apr;17(2):86-95.
10. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008 Dec 16;149(12):889-97.
11. Whiting P, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology*. [diagnostic methods]. 2003;3(1):25.
12. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006;6:31.
13. Toft N, Nielsen SS. Summary receiver operating characteristics (SROC) and hierarchical SROC models for analysis of diagnostic test evaluations of antibody ELISAs for paratuberculosis. *Prev Vet Med*. 2009 Nov 15;92(3):249-55.
14. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005 Oct;58(10):982-90.
15. Harbord R, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. In: Sterne J, editor. *Meta-Analysis in Stata: An Unupdated Collection from the Stata Journal*. College Station, Texas, USA: Stata Press; 2009. p. 181 - 99.
16. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003 Sep 6;327(7414):557-60.
17. Hamasur B, Bruchfeld J, Haile M, Pawlowski A, Bjorvatn B, Kallenius G, et al. Rapid diagnosis of tuberculosis by detection of mycobacterial lipoarabinomannan in urine. *J Microbiol Methods*. 2001 May;45(1):41-52.

18. Tessema TA, Hamasur B, Bjun G, Svenson S, Bjorvatn B. Diagnostic evaluation of urinary lipoarabinomannan at an Ethiopian tuberculosis centre. *Scand J Infect Dis.* 2001;33(4):279-84.
19. Peter J, Green C, Hoelscher M, Mwaba P, Zumla A, Dheda K. Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments. *Curr Opin Pulm Med.* 2010 May;16(3):262-70.
20. Shah M, Variava E, Holmes CB, Coppin A, Golub JE, McCallum J, et al. Diagnostic Accuracy of a Urine Lipoarabinomannan Test for Tuberculosis in Hospitalized Patients in a High HIV Prevalence Setting. *J Acquir Immune Defic Syndr.* 2009 Aug 18.
21. Boehme C, Molokova E, Minja F, Geis S, Loscher T, Maboko L, et al. Detection of mycobacterial lipoarabinomannan with an antigen-capture ELISA in unprocessed urine of Tanzanian patients with suspected tuberculosis. *Trans R Soc Trop Med Hyg.* 2005 Dec;99(12):893-900.
22. Shah M, Martinson NA, Chaisson RE, Martin DJ, Variava E, Dorman SE. Quantitative analysis of a urine-based assay for detection of lipoarabinomannan in patients with tuberculosis. *J Clin Microbiol.* 2010 Aug;48(8):2972-4.
23. Chan ED, Reves R, Belisle JT, Brennan PJ, Hahn WE. Diagnosis of tuberculosis by a visually detectable immunoassay for lipoarabinomannan. *Am J Respir Crit Care Med.* 2000 May;161(5):1713-9.
24. Doublier S, Zennaro C, Spatola T, Lupia E, Bottelli A, Deregibus MC, et al. HIV-1 Tat reduces nephrin in human podocytes: a potential mechanism for enhanced glomerular permeability in HIV-associated nephropathy. *AIDS.* 2007 Feb 19;21(4):423-32.
25. Dheda K, Davids V, Lenders L, Roberts T, Meldau R, Ling D, et al. Clinical utility of a commercial LAM-ELISA assay for TB diagnosis in HIV-infected patients using urine and sputum samples. *PLoS One.* 2010;5(3):e9848.
26. Reichman LB, Hershfield ES. *Tuberculosis: A Comprehensive International Approach.* 2nd ed. Lenfant C, editor. New York: Marcel Dekker; 2000.
27. Reither K, Saathoff E, Jung J, Minja LT, Kroidl I, Saad E, et al. Low sensitivity of a urine LAM-ELISA in the diagnosis of pulmonary tuberculosis. *BMC Infect Dis.* 2009;9:141.
28. Mutetwa R, Boehme C, Dimairo M, Bandason T, Munyati SS, Mangwanya D, et al. Diagnostic accuracy of commercial urinary lipoarabinomannan detection in African tuberculosis suspects and patients. *Int J Tuberc Lung Dis.* 2009 Oct;13(10):1253-9.
29. Lawn SD, Edwards DJ, Kranzer K, Vogt M, Bekker LG, Wood R. Urine lipoarabinomannan assay for tuberculosis screening before antiretroviral therapy diagnostic yield and association with immune reconstitution disease. *AIDS.* 2009 Sep;23(14):1875-80.
30. Daley P, Michael JS, Hmar P, Latha A, Chordia P, Mathai D, et al. Blinded evaluation of commercial urinary lipoarabinomannan for active tuberculosis: a pilot study. *Int J Tuberc Lung Dis.* 2009 Aug;13(8):989-95.

**Acknowledgements**

This study was supported by funding from the TREAT TB initiative by the IUATLD, with funding support from USAID. Jessica Minion is a recipient of a Quebec Respiratory Health Training Fellowship. Madhukar Pai is a recipient of a CIHR New Investigator Award from the Canadian Institutes of Health Research. Dick Menzies is a recipient of a Chercheur National salary award from the FRSQ. KD is supported by a South African Research Chair Initiative award, a MRC Career Development Fellowship, the EU(FW7-TBsusgent) and the EDCTP (TESA and TB-NEAT). None of these funding agencies had any role in the design and conduct of this study.

**Table 1. Characteristics of Included Studies**

Reference	Author (Year)	Country	Total N	Manufacturer	Urine	% Smear +	% HIV +	HIV Stratified	Reference <sup>a</sup>	Type of Analyses Possible <sup>b</sup>		
										A	B	C
(17)	Hamasur (2001)	SWEDEN	54	NON-COMMERCIAL	FROZEN	11.1	NR	No	Culture + AFB + Clinical <sup>c</sup>	+	-	-
(18)	Tessema (2001)	ETHIOPIA	1000	NON-COMMERCIAL	FROZEN	13.9	NR	No	AFB + Clinical	+	+	-
(21)	Boehme (2005)	TANZANIA	231	CHEMOGEN	FRESH	35	69	No	LJ + CXR	+	+	+
(27)	Reither (2009)	TANZANIA	207	CHEMOGEN	FRESH	16.5	59.1	Yes	LJ + MGIT + Clinical	+	+	+
(28)	Mutetwa (2009)	ZIMBABWE	261	CHEMOGEN	FRESH	30.5	77	Yes	LJ + Clinical	+	-	-
(29)	Lawn (2009)	SOUTH AFRICA	235	CHEMOGEN	FROZEN	3.4	100	Yes	MGIT	-	+	-
(20)	Shah (2009)	SOUTH AFRICA	404	INVERNESS	FROZEN	16	85	Yes	MGIT + Pathology + Clinical	+	+	+
(30)	Daley (2009)	INDIA	200	CHEMOGEN	FROZEN	22.5	8.5	Yes	LJ + MGIT + Clinical	-	+	+
(25)	Dheda (2010)	SOUTH AFRICA	306	INVERNESS	FROZEN	20.9	27.3	Yes	MGIT + Clinical	+	+	+

<sup>a</sup>diagnostic information available to contribute to reference standard; AFB, LJ, MGIT and Pathology considered microbiologic.

<sup>b</sup>Types of Analyses:

A – Analysis excludes clinical cases (patients that were microbiologically-negative, but had strong clinical/radiologic suspicion for TB)

B – Analysis groups clinical cases with all other patients that were microbiologically-negative to be reference negative.

C – Analysis groups clinical cases with patients that were microbiologically-positive as reference positive

<sup>c</sup>Culture confirmation (not otherwise specified) in 9 cases; AFB smear and clinical diagnosis of controls

NR = not reported; AFB = acid fast bacilli present in microscopic smear; LJ = Lowenstein-Jensen culture; MGIT = Mycobacterial Growth Indicator Tube (BACTEC, 460 or 960)

**Table 2. Pooled Meta-Analysis Estimates using Bivariate Random Effects Model**

	<b>Sensitivity<sup>a</sup></b> <b>(95% CI)</b>	<b>I<sup>2</sup></b> <b>(p-value)</b>	<b>Specificity<sup>a</sup></b> <b>(95% CI)</b>	<b>I<sup>2</sup></b> <b>(p-value)</b>
<b>Analysis (A) – excluding clinical cases (n=7)</b>	60% (38, 79)	97.0% (p<0.0001)	93% (88, 96)	84.5% (p<0.0001)
HIV +ve (n=4)	51% (32, 69)	90.7% (p<0.0001)	94% (79, 98)	77.6% (p=0.004)
HIV -ve (n=4)	14% (7, 24)	53.3% (p=0.09)	97% (90, 99)	63.1% (p=0.04)
<b>Analysis B – Clinical cases considered NOT active TB (n=7)</b>	47% (26, 69)	97.3% (p<0.0001)	93% (83, 97)	94.8% (p<0.0001)
HIV +ve (n=4)	56% (40, 71)	24.1% (p<0.0001)	95% (77, 99)	92.0% (p<0.0001)
HIV -ve (n=3) <sup>b</sup>	18% (10, 29)	0.0% (p=0.88)	90% (85, 93)	69.3% (p=0.04)
<b>Analysis C – Clinical cases considered active TB (n=5)</b>	34% (14, 62)	98.4% (p<0.0001)	94% (87, 98)	85.5% (p<0.0001)
HIV +ve (n=2) <sup>b</sup>	49% (43, 54)	54.4% (p=0.14)	91% (85, 96)	69.5% (p=0.07)
HIV -ve (n=2) <sup>b</sup>	16% (8, 28)	0.0% (p=0.4)	94% (86, 98)	70.8% (p=0.06)

A– Analysis excludes clinical cases (patients that were culture negative, but had strong clinical/radiologic suspicion for TB)

B – Analysis groups clinical cases with all other patients that were culture negative to be reference negative

C – Analysis groups clinical cases with patients that were microbiologically-positive, as reference positive

<sup>a</sup>Due to imprecision of estimates, sensitivity and specificity are rounded to the nearest whole number

<sup>b</sup>pooled estimates calculated with univariate random effects models; confidence intervals are expected to be narrower

**Table 3. Summary of Data Comparing Diagnostic Accuracy in groups stratified by CD4 count in HIV co-infected TB Suspects (3 studies including 444 subjects total)**

Reference	Author (Year)	Sensitivity (95% CI)	Specificity (95% CI)
(29)	Lawn (2009) <sup>a</sup>		
	CD4 <50	67% (41, 87)	100% (94, 100)
	CD4 50-100	35% (14, 62)	100% (0.94, 1.00)
	CD4 >100	4% (0, 22)	100% (94, 100)
(25)	Dheda (2010) <sup>b</sup>		
	CD4 <200	37% (16, 62)	100% (81, 100)
	CD4 ≥200	0% (0, 18)	100% (17, 100)
(22)	Shah (2010) <sup>c</sup>		
	CD4 <50	85% (73, 93)	NA
	CD4 50-100	71% (51, 87)	NA
	CD4 101-150	56% (30, 80)	NA
	CD4 >150	51% (38, 64)	NA

<sup>a</sup>Outpatients being screened for TB prior to initiation of HAART. Reference standard uses strict microbiologic criteria (clinical cases grouped with all other patients that were culture negative to be reference negative; Analysis B)

<sup>b</sup>Outpatients with suspected TB presenting to a TB clinic. Reference standard excluded clinical cases (patients that were culture negative, but had strong clinical/radiologic suspicion for TB; Analysis A)

<sup>c</sup>Hospitalized patients with suspected TB. Reference standard uses strict microbiologic criteria (clinical cases grouped with all other patients that were culture negative to be reference negative; Analysis B)

NA = not available

**Table 4. Pooled Estimates of Test Accuracy in Subgroups Stratified by Different Factors<sup>a</sup>**

	<b>Sensitivity (95% CI)</b>	<b>I<sup>2</sup> (p-value)</b>	<b>Specificity (95% CI)</b>	<b>I<sup>2</sup> (p-value)</b>
<b>Urine Preparation (n studies)<sup>b</sup></b>				
Fresh (n=3) <sup>c</sup>	59% (53, 64)	95.4% (p<0.0001)	89% (85, 92)	0.0% (p=0.9)
Frozen (n=6)	49% (22, 77)	97.3% (p<0.0001)	97% (89, 99)	93.3% (p<0.0001)
<b>Test Version (n studies)</b>				
Non-commercial (n=2) <sup>c</sup>	82% (76, 88)	39.8% (p=0.2)	87% (85, 89)	62.3% (p=0.1)
Chemogen (n=5)	45% (26, 66)	95.1% (p<0.0001)	94% (82, 98)	88.9% (p<0.0001)
Inverness (n=2) <sup>c</sup>	40% (35, 46)	98.7% (p<0.0001)	98% (95, 99)	61.6% (p=0.1)

<sup>a</sup>Estimates calculated using bivariate random effects models, preferentially using data excluding clinical cases (Analysis A), otherwise strict microbiologic reference standards (Analysis B)

<sup>b</sup>if urine was frozen before LAM testing was performed it was considered ‘frozen’ whether processing of urine was done before or after freezing

<sup>c</sup>univariate random effects regression performed because too few studies to converge bivariate model

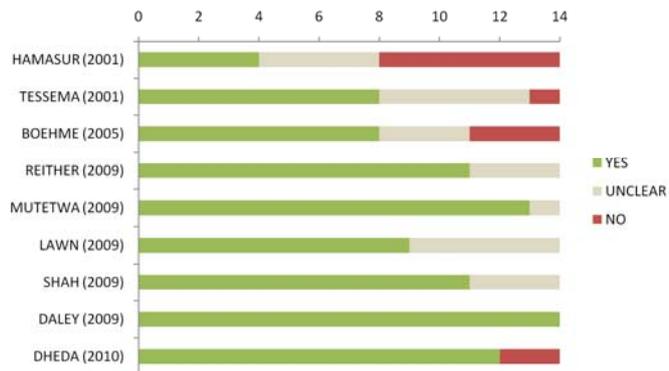
## FIGURE TITLES AND LEGENDS

**Figure 1. Selection of Included Studies from Systematic Review**



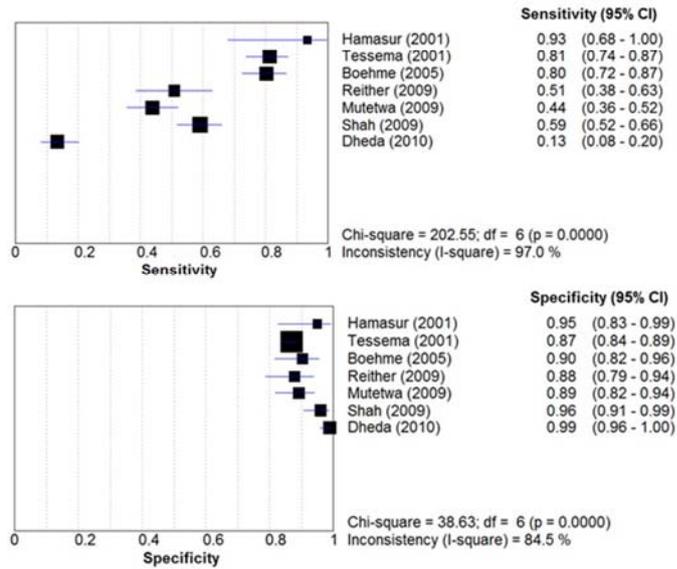
**Figure 2. Quality of Studies using QUADAS Criteria**

Note: QUADAS criteria judged on data used for meta-analysis, not necessarily all data presented in studies



**Figure 3. Forest Plots of Studies Contributing to Analysis A (n=7)**

Point estimates of sensitivity and specificity from each study reporting results eligible for analysis by excluding microbiologically-negative clinical cases are shown as closed squares. Size of the square is proportionate to the size of the study. Solid lines represent 95% confidence intervals.



#### Figure 4. Linear Regression of Sensitivity on the Proportion of HIV + Included in Studies

Open circles represent studies reporting HIV prevalence; size of the circles depend on the precision of each study estimates (i.e. the inverse of its within-study variance).

Line represents fitted values for linear regression equation:

$$\text{sensitivity} = 0.17 (SE=0.18) + 0.0042 (SE=0.0027) * \%HIV.$$

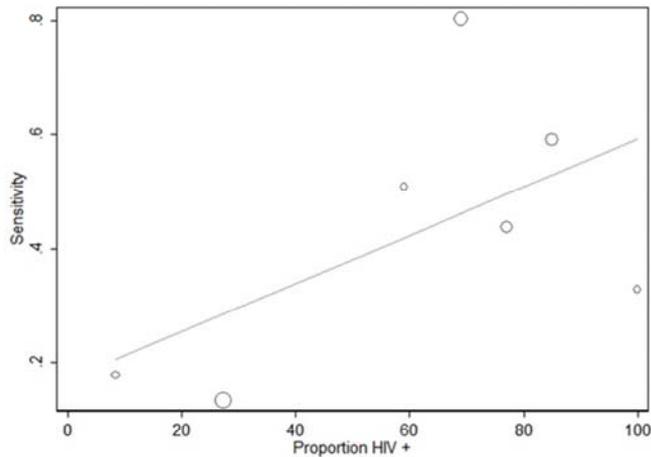
95% CIs:  $\alpha = [-0.30 - 0.64]^\dagger$ ;  $\beta = [-0.0027 - 0.011]^\dagger$ .

Logistic model (not displayed):

$$\text{logit}(\text{sensitivity}) = -1.63 (SE=0.87) + 0.021 (SE=0.013) * \%HIV.$$

95% CIs:  $\alpha = [-3.86 - 0.60]^\dagger$ ;  $\beta = [-0.012 - 0.053]^\dagger$ .

$^\dagger$ not statistically significant



**Figure 5. Proportion of Specimens LAM + by Clinical Status**

Shaded bars display the proportion of specimens LAM + within each group (labelled with pooled proportions calculated using univariate random effects models); Brackets represent upper and lower 95% confidence intervals; n=# of studies contributing data to estimate.

Smear + = positive smear and culture; Culture + = all positive cultures regardless of smear status; Smear - = negative smear with positive culture; Clinical + = negative culture with clinically diagnosed active TB; Suspect Controls = patients with signs and symptoms compatible with active TB, but TB was ruled out; Healthy Controls = participants without signs or symptoms of active TB.

