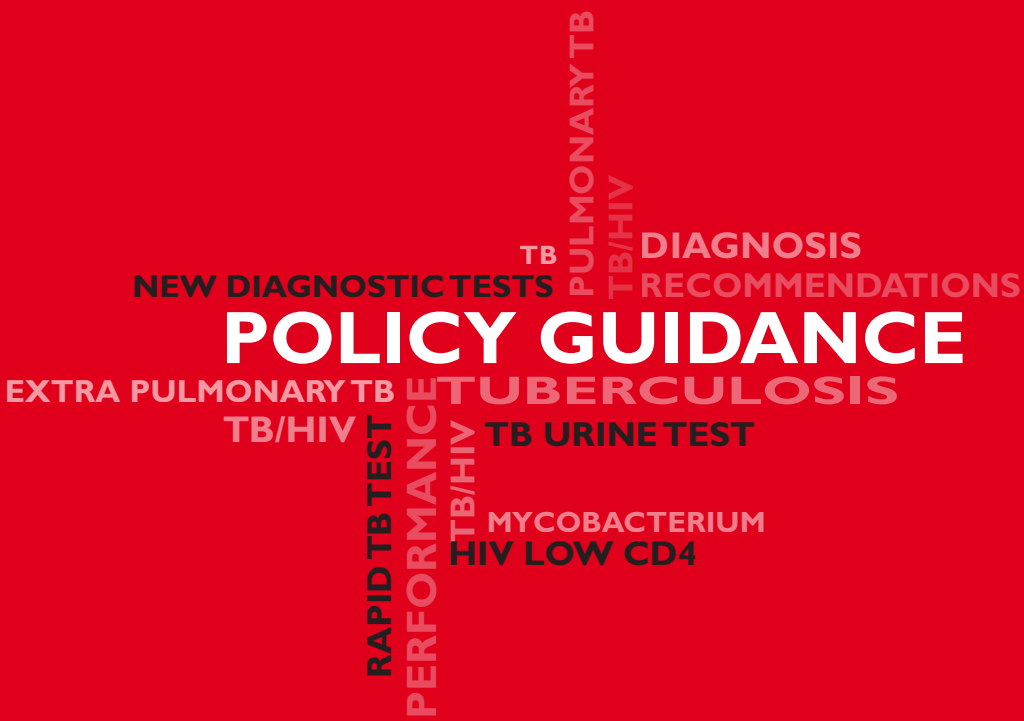
The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV



THE END TB STRATEGY

World Health Organization

The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV

Policy guidance

WHO Library Cataloguing-in-Publication Data

The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV. Policy guidance.

I.World Health Organization.

ISBN 978 92 4 150963 3

Subject headings are available from WHO institutional repository

Designed by GPS Publishing

Printed by the WHO Document Production Services, Geneva, Switzerland

WHO/HTM/TB/2015.25

Contents

Tables

Figures

Abbreviations

AFB	acid-fast bacilli
CI	confidence interval
CrI	credible interval
CRS	composite reference standard
DALY	disability-adjusted life year
DOI	Declaration of Interests
DST	drug-susceptibility testing
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HIV	human immunodeficiency virus
LAM	lipoarabinomannan
LF-LAM	Lateral flow urine lipoarabinomannan assay
MDR-TB	multidrug-resistant tuberculosis
NAAT	nucleic acid amplification test
PCR	polymerase chain reaction
PICO	Population, Intervention, Comparator, Outcome
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
TB	tuberculosis
USAID	United States Agency for International Development
WHO	World Health Organization

Acknowledgements

This document was prepared by Christopher Gilpin and Alexei Korobitsyn, with input from Karin Weyer, Haileyesus Getahun, Alberto Matteelli, Wayne van Gemert and Fuad Mirzayev (WHO Global TB Programme) and Meg Doherty (WHO HIV Department), on the basis of consensus agreed at a Guideline Development Group Meeting convened by WHO in Geneva on 1 June 2015.

The findings and recommendations from the meeting were presented to an External Review Group in August 2015. The External Review Group agreed with the recommendations made by the Guideline Development Group on the use of lateral flow urine lipoarabinomannan (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV. This document was finalized following consideration of all comments and suggestions from the participants of the Guideline Development Group (Annex 1) and the External Review Group (Annex 2).

WHO gratefully acknowledges the contributions of the Chairperson of the Guideline Development Group (Jan Brozek), the members of the Guideline Development Group, and the External Review Group. Karen Steingart, Maunank Shah and Jonny Peters (systematic reviewers for LF-LAM), David Dowdy and Colleen Hanrahan (systematic reviewers for the economic evaluations of LF-LAM) are thanked for preparing the systematic reviews and presenting their findings to the members of the Guideline Development Group.

Funding from the United States Agency for International Development is gratefully acknowledged through USAID-WHO Consolidated Grant No. GHA-G-00-09-00003 / US 2014-741.

Declarations of Interests

The members of the Guideline Development Group, commissioned experts and members of the External Review Group completed Declarations of Interests (DOIs). These were reviewed by the WHO Steering Group (Annex 3) prior to the meeting of the Guideline Development Group and prior to preparing the policy guidance. The review of each DOI assessed whether an interest had been declared and, if so, whether it was insignificant or potentially significant. If the Steering Group determined that no relevant interest had been declared or such interest was insignificant or minimal, individuals were invited to participate. A summary of DOI statements is provided in Annex 4. Claudia Denkinger declared that as a FIND employee she had provided advice to the manufacturer of LF-LAM on the specifications of other diagnostics manufactured by Alere and that she had contributed to the systematic review for LF-LAM. These declarations were deemed to be significant for a possible conflict of interest, consequently she did not contribute to the deliberations on the GRADE evaluation process or the eventual recommendations, and participated in the meeting as an observer. None of the DOIs of the GDG members were declared significant or potentially significant. Members of the systematic review team were invited to provide technical input and answer technical questions. These individuals did not participate in the GRADE evaluation process nor in the final discussions when recommendations were developed. Also, they were not involved in developing the report of the Guideline Development Group's meeting, nor in preparing the WHO policy guidance.

Executive summary

Background

Tests based on the detection of mycobacterial lipoarabinomannan (LAM) antigen in urine have emerged as potential point-of-care tests for tuberculosis (TB). LAM antigen is a lipopolysaccharide present in mycobacterial cell walls, which is released from metabolically active or degenerating bacterial cells and appears to be present only in people with active TB disease. Urine-based testing would have advantages over sputum-based testing because urine is easy to collect and store, and lacks the infection control risks associated with sputum collection.

Objectives, rationale and methods used to develop the guidance

This document provides a summary of the evidence and recommendations for the use of LF-LAM for the diagnosis and screening of active tuberculosis **in people living with HIV**.

The objectives of this policy guidance are:

- To assess the available data on the accuracy (sensitivity and specificity) of lateral flow urine lipoarabinomannan assay (LF-LAM) for screening and diagnosis[1] of active TB in HIV-infected adults, at different thresholds for test positivity, as a replacement or in combination with other diagnostic tools.
- To assess the available data on the accuracy (sensitivity and specificity) of lateral flow urine lipoarabinomannan assay (LF-LAM) for diagnosis of active TB in HIV-infected children.
- To assess data related to patient outcomes, both the association of LAM positivity and patient outcomes and the impact of LAM implementation on patient outcomes in HIV-infected patients, as a replacement or in combination with other diagnostic tools.
- To assess available data on the cost, and cost-effectiveness of LAM implementation for TB diagnosis or screening of active tuberculosis in people living with HIV compared with sputum microscopy or Xpert MTB/RIF.
- To develop WHO policy recommendations for the appropriate use of LF-LAM for the diagnosis and screening of active tuberculosis in people living with HIV.

The lateral flow urine lipoarabinomannan (LF-LAM) assay reviewed is a commercially available test to detect active TB (Alere Determine™ TB LAM Ag, Alere Inc, Waltham, MA, USA). The test is performed manually by applying 60 µL of urine to the Determine™ TB LAM Ag test strip and incubating at room temperature for 25 minutes. The strip is then inspected by eye. The intensity of any visible band on the test strip is graded by comparing it with the intensities of the bands on a manufacturer-supplied reference card. Prior to January 2014, this reference card included five bands (grade 1 representing a very low intensity band to grade 5 representing a high/dark intensity band). After January 2014, the manufacturer revised the reference bands to contain only 4 grades, such that the band intensity for the new grade 1 corresponded to the previous reference card band intensity for grade 2.

Several studies and a meta-analysis of an earlier generation LAM-ELISA test have demonstrated improved sensitivity of urinary LAM in the presence of HIV-TB co-infection, which further increases with lower CD4 counts. This finding is in contrast to traditional diagnostic methods for TB in people with HIV. Several hypotheses may explain the higher sensitivity of urine LAM detection in patients

[1] Studies that evaluated LF-LAM in participants with symptoms consistent with TB were classified as 'studies for TB diagnosis' and studies that performed systematic screening with LF-LAM in participants whether or not signs and symptoms were present as 'studies for TB screening'

with HIV-related immunosuppression, including higher bacillary burden and antigen load, greater likelihood of TB in the genitourinary tract, and greater glomerular permeability to allow increased antigen levels in urine.

In response to requests from end-users in the field for guidance on the appropriate use of the LF-LAM assay and given the potential of the assay to help reduce mortality in persons living with HIV, WHO commissioned a systematic review of the use of the LF-LAM assay for the diagnosis and screening of active TB in people living with HIV. Given the test is easy to perform, requires minimal biosafety requirements, is inexpensive and in response to its increasing use in HIV prevalent settings, it was considered necessary to develop clear guidance for which patient populations are suitable to test, to avoid having test applied inappropriately.

On June 1st 2015, the WHO Global TB Programme convened a Guideline Development Group in Geneva, Switzerland to review the evidence for the use of LF-LAM. The meeting was chaired by an expert in evidence synthesis. Recommendations were developed based on consensus among the Guideline Development Group members and were subsequently confirmed by an External Review Panel. The evidence reviewed and recommendations apply to the use of LF-LAM only as other in-house LAM based assays have not been adequately validated or used outside limited research settings. Any new or generic LAM based assay should be subject to adequate validation in the settings of intended use. The WHO policy recommendations developed from the evidence synthesis process by the Guideline Development Group are summarized below.

WHO's policy recommendations

Policy Recommendations for the use of the lateral flow urine lipoarabinomannan (LF-LAM) assay

1. **Except as specifically described below for persons with HIV infection with low CD4 counts or who are seriously ill**[2], LF-LAM **should not be used** for the diagnosis of TB **(strong recommendation, low quality of evidence)**.
2. LF-LAM **may be used** to assist in the diagnosis of TB **in HIV positive** adult *in-patients* with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count less than or equal to 100 cells/μL, or HIV positive patients who are seriously ill[2] regardless of CD4 count or with unknown CD4 count **(conditional recommendation; low quality of evidence**).

   Remarks

   a. This recommendation also applies to HIV positive adult *out-patients* with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count less than or equal to 100 cells/μL, or HIV positive patients who are seriously ill[2] regardless of CD4 count or with unknown CD4 count, based on the generalisation of data from in-patients.

   b. This recommendation also applies to HIV positive children with signs and symptoms of TB (pulmonary and/or extrapulmonary) based on the generalisation of data from adults while acknowledging very limited data and concern regarding low specificity of the LF-LAM assay in children.
3. LF-LAM **should not be used** as a screening test for TB. (**strong recommendation, low quality of evidence**).

---

[2] "seriously ill" is defined based on 4 danger signs: respiratory rate > 30/min, temperature > 39°C, heart rate > 120/min and unable to walk unaided.
World Health Organization. Improving the diagnosis and treatment of smear-negative pulmonary and extrapulmonary tuberculosis among adults and adolescents. Recommendations for HIV-prevalent and resource constrained settings. World Health Organization 2007.
Available at: http://www.uphs.upenn.edu/bugdrug/antibiotic_manual/smear_neg_and_extrapulmTb.pdf

1. Background

Key global priorities for tuberculosis (TB) care and control include improving case-detection and detecting cases earlier, including cases of smear-negative disease which are often associated with co-infection with the human immunodeficiency virus (HIV) and with young age.

In 2014, an estimated 1.2 million (13%) of the 9.6 million people who developed TB worldwide were HIV-positive[3]. The African Region accounted for 73% of the estimated number of HIV-positive incident TB cases[3]. Globally, people living with HIV are 26 times more likely to develop TB disease than those who are HIV-negative[3]. Beginning in the 1980s, the HIV epidemic led to a major upsurge in TB cases and TB mortality in many countries, especially in southern and eastern Africa. TB occurs early in the course of HIV infection and shortens patient survival if not rapidly diagnosed and treated. Many people infected with HIV in developing countries develop TB as the first manifestation of AIDS.

Tests based on the detection of mycobacterial lipoarabinomannan (LAM) antigen in urine have been developed as potential point-of-care tests for TB. Urine-based testing has advantages over sputum-based testing because urine is easy to collect and store, and lacks the infection control risks associated with sputum collection. A LAM antigen is a lipopolysaccharide present in mycobacterial cell walls, which is released from metabolically active or degenerating bacterial cells. LAM appears to be present predominately in people with active TB disease and has shown only low cross-reactivity with nontuberculous mycobacterial infections[4]. LAM is an attractive diagnostic target as urine processing requires limited infection control measures, presence of LAM in urine is indirectly related to human immune response, and its detection process is amenable to inexpensive POC platforms. Owing to suboptimal sensitivity, current urinary LAM assays are deemed unsuitable as general screening tests for TB. However, unlike traditional TB diagnostic methods, they demonstrate improved sensitivity in HIV-TB co-infection which further increases with lower CD4 counts.

Published studies have reported much higher mortality rates among HIV positive individuals with low CD4 counts who have detectable urinary LAM using the commercially available lateral flow urine LAM assay (LF-LAM) (Alere Determine™ TB LAM Ag, Alere Inc, Waltham, MA, USA) compared with LF-LAM negative individuals. The LF-LAM assay if accurate, could therefore be a useful tool to facilitate the early initiation of anti-TB treatment and help reduce mortality in this patient group.

HIV-positive patients with TB disease may be missed for the following reasons: sputum bacillary load is typically low in these patients; they may not be able to provide sufficient and high quality sputum specimens; and a substantial proportion of these patients have extrapulmonary TB without pulmonary TB. Due to high rates of mortality among this patient group, if accurate, a urinary LAM assay would be a useful tool to facilitate the early initiation of anti-TB treatment and thus people with HIV-TB co-infection who are difficult to diagnose with TB using traditional diagnostic methods may benefit.

In response to requests from end-users for guidance on the appropriate use of the LF-LAM assay, its increasing use in high-burden countries and given the potential of the assay to help reduce mortality in HIV positive individuals, WHO commissioned a systematic review of the use of the LF-LAM assay for the diagnosis and screening of active TB in people living with HIV. Given the test is easy to perform, requires minimal biosafety requirements, is inexpensive and in response to its increasing use in HIV prevalent settings, it was considered necessary to develop clear guidance for which patient populations are suitable to test, to avoid having test applied inappropriately.

---

[3] WHO 2015. World Health Organization. Global tuberculosis control: WHO report 2015. WHO/HTM/TB/2015.22

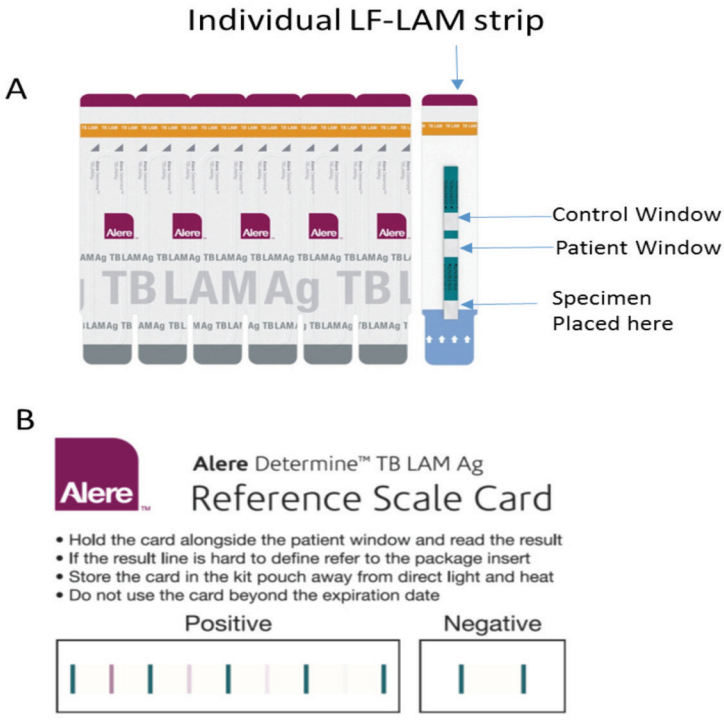
[4] Qvist T et al. BMC Infectious Diseases 2014, 14:655 http://www.biomedcentral.com/content/pdf/s12879-014-0655-4.pdf (Accessed 17 Aug 2015)

In accordance with WHO standards for assessing evidence when formulating policy recommendations, the GRADE approach (the Grading of Recommendations Assessment, Development and Evaluation, see http://www.gradeworkinggroup.org/) was used. GRADE provides a structured framework for evaluating the accuracy of diagnostic tests, and the impact on patient- and public health of new diagnostic tests. The systematic review assessed the accuracy of a commercially available LF-LAM for diagnosing and screening of active TB in adults living with HIV. The assay reviewed is a commercially available test to detect active TB (Alere Determine™ TB LAM Ag, Alere Inc, Waltham, MA, USA).

The evidence reviewed and this policy guidance apply to the use of the commercial LF-LAM assay only. Other in-house LAM based assays have not been adequately validated or used outside limited research settings. Any new or generic LAM based assay should be subject to adequate evaluation and validation in the settings of intended use as per WHO policy[5].

Figure 1. Alere Determine™ TB LAM Ag test

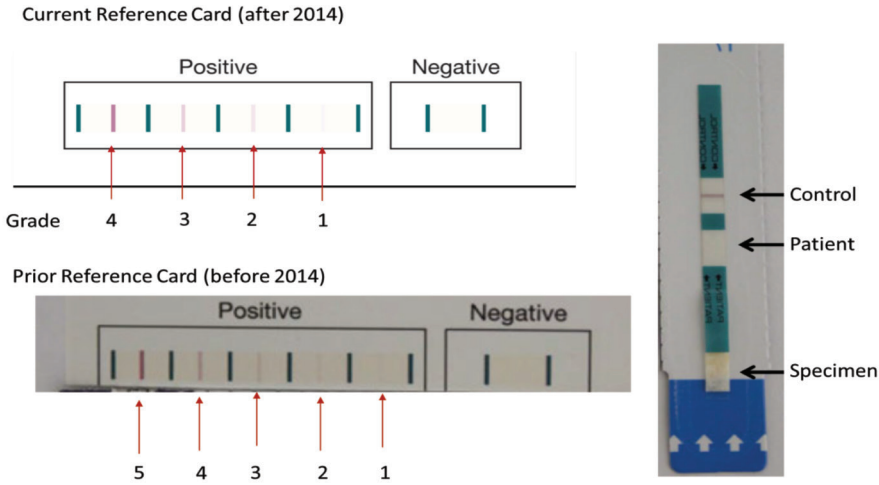
(A) Alere Determine™ TB LAM Ag tests. To the sample pad (white pad marked by the arrow symbols) 60 µL of urine is applied and visualized bands are read 25 minutes later. (B) Reference card accompanying test strips to 'grade' the test result and determine positivity (33). Copyright© [2014] [Alere Inc]: reproduced with permission.



[5] WHO 2015. Implementing tuberculosis diagnostics. Policy Framework. WHO/HTM/TB/2015.11 http://apps.who.int/iris/bitstream/10665/162712/1/9789241508612_eng.pdf?ua=1

The test is performed manually by applying 60 µL of urine to the Determine™ TB LAM Ag test strip and incubating at room temperature for 25 minutes[6] (Figure 1). The strip is then inspected by eye. The intensity of any visible band on the test strip is graded by comparing it with the intensities of the bands on a manufacturer-supplied reference card. Prior to January 2014, this reference card included five bands (grade 1 representing a very low intensity band to grade 5 representing a high/dark intensity band). After January 2014, the manufacturer revised the reference bands to contain only 4 grades, such that the band intensity for the new grade 1 corresponded to the previous reference card band intensity for grade 2 (Figure 2).

Figure 2. Grading of Determine TB Ag assay



[6] Alere. Alere Determine™ TB LAM Ag Product Information. http://www.alere.com/ww/en/productdetails/determine-tb-lam.html (accessed 16 Nov 2014).

2. Methods

2.1. Evidence synthesis

In June 2015, a Guideline Development Group was convened by WHO's Global TB Programme to assess available data on the use of LF-LAM. WHO commissioned a systematic review on the use of LF-LAM for the diagnosis and screening for TB among person with HIV as well as a review of the affordability and cost-effectiveness of LF-LAM. Both published and unpublished studies were included in these reviews, given the urgency for guidance on appropriate use of LF LAM expressed by end-users as described above.

The evaluation used the GRADE system to determine the quality of the evidence and provide information on the strength of the recommendations using a priori PICO questions agreed by the Guideline Development Group. PICO refers to the following four elements that should be included in questions that govern a systematic search of the evidence: the Population targeted by the action or intervention (in the case of systematic reviews of diagnostic test accuracy, P is the population of interest); the Intervention (I is the index test; the Comparator (C is the comparator test(s); and the Outcome (O is usually sensitivity and specificity). The PICO questions for the review are given in Box 1.

Box 1. PICO questions for systematic reviews evaluating the accuracy of the LF-LAM assay for diagnosis tuberculosis in adults and children

I. Urine LAM for Diagnosis of TB

Overall accuracy in adults and children

1a. Should LF-LAM be used to diagnose tuberculosis in adults with HIV, microbiological reference standard, grade 2?

1b. Should LF-LAM be used to diagnose tuberculosis in adults with HIV, composite reference standard, grade 2?

1c. Should LF-LAM be used to diagnose tuberculosis in children with HIV?

By health care setting

2a. Should LF-LAM be used to diagnose tuberculosis in adults with HIV in inpatient settings, grade 2?

2b. Should LF-LAM be used to diagnose tuberculosis in adults with HIV in outpatient settings, grade 2?

By CD4 threshold

3ai. Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 > 200, grade 2?

3aii. Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 200, grade 2?

3bi. Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 > 100, grade 2?

3bii. Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 100, grade 2?

3ci. Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 200 in inpatient settings, grade 2?

3cii. Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 100 in inpatient settings, grade 2?

LF-LAM and existing tests

4a. Should LF-LAM vs. sputum microscopy be used to diagnose tuberculosis in adults with HIV, grade 2?

4b. Should LF-LAM in combination with sputum microscopy be used to diagnose tuberculosis in adults with HIV, grade 2?

4c. Should LF-LAM vs. sputum Xpert MTB/RIF be used to diagnose tuberculosis in adults with HIV, grade 2?

4d. Should LF-LAM in combination with sputum Xpert be used to diagnose tuberculosis in adults with HIV, grade 2?

II. LF-LAM for Screening for TB

Overall accuracy in adults

5a. Should LF-LAM be used to screen for tuberculosis in adults with HIV, microbiological reference standard, grade 2?

5b. Should LF-LAM be used to screen for tuberculosis in adults with HIV, composite reference standard, grade 2?

By health care setting

6a. Should LF-LAM be used to screen for tuberculosis in adults with HIV in inpatient settings, grade 2?

6b. Should LF-LAM be used to screen for tuberculosis in adults with HIV in outpatient settings, grade 2?

By CD4 threshold

7ai. Should LF-LAM be used to screen for tuberculosis in adults with HIV with CD4 > 200?

7aii. Should LF-LAM be used to screen for tuberculosis in adults with HIV with CD4 ≤ 200?

7bi. Should LF-LAM be used to screen for tuberculosis in adults with HIV with CD4 > 100?

7bii. Should LF-LAM be used to screen for tuberculosis in adults with HIV with CD4 ≤ 100?

The systematic reviews were conducted according to the standards outlined by the Cochrane Collaboration in the Cochrane Handbook.[7] A comprehensive search of the following databases was performed on 2 February 2015, without date or language restrictions: Cochrane Infectious Diseases Group Specialized Register; PubMed; EMBASE; ISI Web of Knowledge; MEDION; LILACS; BIOSIS; and SCOPUS. Searches of the metaRegister of Controlled Trials (mRCT) and the search portal of the WHO International Clinical Trials Registry Platform (www.who.int/trialsearch) were also performed to identify ongoing trials and ProQuest Dissertations & Theses A&I to identify relevant dissertations.

The data were reported in the TP, FP, FN, TN format to calculate sensitivity and specificity estimates and 95% confidence intervals (CI) for individual studies. The individual study results graphically by plotting the estimates of sensitivity and specificity (and their 95% CIs) in forest plots using Review Manager. Thus for the individual studies in the coupled forest plots have 95% confidence intervals, not credible intervals.

For the meta-analyses, however, in order to obtain pooled sensitivity and specificity estimates combining multiple studies, a bivariate random-effect model using a Bayesian approach was used as the statistical method. The Bayesian approach provides 95% credible intervals (CrI). The 95% CrI is the Bayesian equivalent of the classical (frequentist) 95% CI. 95% CI for individual study estimates and 95% CrI for pooled study estimates was specified as appropriate. The 95% CrI may be interpreted as an interval that has a 95% probability of capturing the true value of the unknown parameter given the observed data and the prior information.

The choice of an optimal reference standard is critical to assess the accuracy of diagnostic tests, since the reference standard is used to determine the presence or absence of the target condition.

[7] Higgins JPT, Green S, eds. Cochrane handbook of systematic reviews for interventions, version 5.1.0. Cochrane Collaboration, 2011 (available at http:// www.cochrane-handbook.org).

The target condition was active TB disease, which included both pulmonary and extrapulmonary TB. The search strategy to identify studies for inclusion used either or both of the following reference standards:

A microbiological reference standard for TB was defined as a positive culture for *M. tuberculosis* complex or a positive TB nucleic acid amplification test (NAAT)[8] and the absence of TB was defined as a negative culture for *M. tuberculosis* complex and/or negative NAAT[8] result (if performed).

A composite reference standard that included at least one of the following components: Positive culture for *M. tuberculosis* complex, positive NAAT[8], positive acid-fast (AFB) smear, or clinical decision to start TB treatment with clinical confirmation after at least one month of treatment. Not having TB was defined as a negative culture and/or NAAT (if performed), negative AFB smear, TB treatment given, and resolution of TB signs and symptoms at follow-up.

Where feasible, meta-analyses were used to summarize the results of independent studies, and these results have been displayed as forest plots. Where meta-analysis was not feasible due to heterogeneity, the evidence has been presented in a narrative synthesis.

Grade evidence profiles were prepared to assess the accuracy of LF-LAM for the diagnosis of active TB disease in adults and children with HIV with signs or symptoms of TB, and to assess the accuracy of LF-LAM used as a screening test for active TB disease in individuals with HIV irrespective of signs or symptoms of TB.

Using the GRADE framework, calculations of test sensitivity and specificity were used as proxy measures for patient outcomes; these outcomes were based on the relative importance or impact of false-positive and false-negative results: Poor sensitivity would result in false-negative results so that patients with TB would be missed, which would have negative consequences in terms of time to treatment initiation, morbidity, mortality and transmission of disease. Poor specificity would result in false-positive results so that patients without TB would be prescribed unnecessary treatment while any alternative diagnoses would be missed.

Rates for true positives, true negatives, false positives and false negatives were calculated using pretest probabilities based on existing literature – that is, an assumed prevalence of TB of 1% among HIV infected persons irrespective of signs and symptoms for TB, an assumed prevalence of TB of 10% among HIV infected persons with symptoms suggested of TB and an assumed prevalence of TB of 30% among seriously ill, hospitalised HIV infected persons with signs and symptoms of TB.

The evaluation of the impact on patients was based on a balance among the following values:

- *true positives* – the benefit to patients from rapid diagnosis and treatment;
- *true negatives* – the benefit to patients who would be spared unnecessary treatment; the benefit of reassurance and alternative diagnosis;
- *false positives* – the likelihood of anxiety and morbidity caused by additional testing, unnecessary treatment and possible adverse effects; possible stigma associated with a diagnosis of TB; and the chance that a false positive may halt further diagnostic evaluation;
- *false negatives* – the increased risk of morbidity and mortality, delayed treatment initiation and the continued risk of transmission of TB.

[8] NAATs included; GenoType MTBDRplus (HAIN Lifesciences, Nehren, Germany); and Xpert® MTB/RIF assay (Cepheid, Sunnyvale, USA).

2.2. Guideline Development Group meeting

The WHO Steering Group (Annex 3) was responsible for scoping the guideline, drafting the PICO questions and overseeing the evidence retrieval and analyses. The Steering Group was also responsible for selecting members for the Guideline Development Group (Annex 1) and External Review Group (Annex 2), for managing declarations of interest, and for organising the Guideline Development meeting. A brief biography of each of the Guideline Development Group members was made available for public scrutiny on the WHO Global TB Programme website (http://www.who.int/tb/laboratory/policy_statements/en/) two weeks prior to the Guideline Development Group Meeting.

PICO questions were drafted by the WHO Steering Group and were presented to the Guideline Development Group for discussion and modification. The Steering Group also prepared an initial list of relevant outcomes, including desirable effects and undesirable effects, and requested the Guideline Development Group to identify any other important outcomes.

In the absence of any comprehensive survey of HIV patient preferences and values for new diagnostic tests, a Global Health Delivery online forum (http://www.ghdonline.org/new-tb-diagnostics/discussion/patient-preferences-for-new-tb-diagnostic-tests/) was convened immediately prior to the Guideline Development Group Meeting as a mechanism to solicit input from a broad range of TB community stakeholders on the benefits of a urine based diagnostic and or screening test for TB for people living with HIV. Technical partners working with affected communities were targeted for the discussion forum although a major limitation of this approach was that it was not always clear whether the respondents in the discussion were qualified to represent the views of affected patients and their communities. Nevertheless, the most common responses were grouped together as the following themes that patients may consider important in a new diagnostic test.

- The ideal test is one that is accurate, least invasive, rapid, affordable, simple to handle, and can be used in the field;
- An ideal test should be one with higher sensitivity compared with specificity;
- Tests need to have acceptable levels of false positive and negative results – the benefits should outweigh the harms;
- Patients would prefer a test with fewer false negatives than false positives;
- In low prevalence settings, patients need a test with low false positives and low false negative results;
- Patients want evidence from a test before starting empiric treatment;
- When patients know that tests are unreliable (false positives and false negatives), this may affect their decision to start or adhere to empiric treatment;
- Patients are less likely to adhere to treatment if not supported by an accurate test.

A webinar was conducted with members of the Guideline Development Group prior to the meeting to refine and finalize the proposed patient outcomes and to rate their relative importance. The following outcomes for each PICO question were determined, and the ratings of their importance were unanimously agreed:

- Critical outcomes – diagnostic accuracy as reflected by true-positive, true-negative, false-positive and false-negative results; time to treatment initiation; mortality.
- Important outcomes – disease severity, intra- and inter- reader test variability, cost.

The format for the "Evidence to Recommendations" (EtR) tables was discussed and agreed upon by the Guideline Development Group members during the webinar. The format included the following sections: description of the problem; diagnostic test accuracy; patient values and preferences; certainty of the evidence for test accuracy; benefits and harms of the test's use; resources required; equity; acceptability; feasibility and neutral language to formulate the recommendations.

Evidence to recommendations tables were developed for each of the PICO questions in order to guide the recommendation development process. These were completed during the meeting and focused on three priority groups of PICO questions for which the group considered that evidence was sufficient to make recommendations:

- Should LF-LAM be used to diagnose tuberculosis in adults with HIV, using a microbiological reference standard and grade 2 test reactions? (Annex 8: EtR table 1a)
- Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 100 cells/μL in inpatient settings, using a microbiological reference standard and grade 2 test reactions? (Annex 8: EtR table 3cii)
- Should LF-LAM be used to screen for tuberculosis in adults with HIV, using a microbiological reference standard and grade 2 test reactions? (Annex 8:EtR table 5a)

The meeting was chaired by a guideline methodologist with expertise in guideline development processes and methods. The methodologist participated in the initial planning, scoping and development of the key questions for the Guideline Development Group meeting. During the meeting, the methodologist helped the guideline development group formulate recommendations based on the evidence presented. Decisions were based on consensus, which was defined as an unanimous agreement among all Guideline Development Group members. Consensus was achieved for the three priority evidence to recommendations tables listed above. The remaining evidence to recommendations tables were compiled by the WHO Steering Group and circulated to the Guideline Development Group for their agreement following the meeting. All tables were subsequently agreed to by consensus among the Guideline Development Group members.

The full set of evidence to recommendations tables are included as a separate online Annex 8 and are available at http://www.who.int/tb/dots/laboratory/policy/en.

2.3. External Review Group

The findings and recommendations from the Guideline Development Group meeting were sent to an External Review Group of international experts in the field of TB laboratory diagnostics, which included representatives from the WHO TB Supranational Reference Laboratory Network, the WHO TB Strategic and Technical Advisory Group (STAG-TB) and the Core Group members of the Global Laboratory Initiative Working Group of the Stop TB Partnership. The External Review Group agreed with the Guideline Development Group's recommendations and with the subsequent WHO policy guidance.

3. Scope

This document provides a pragmatic summary of the evidence and recommendations on the use of LF-LAM for the diagnosis and screening of active tuberculosis, in people living with HIV. It should be read in conjunction 2015 WHO Framework for implementing TB diagnostics, which provides guidance on implementing the diagnostic tools and methods approved by WHO within the context of a country's infrastructure, resources, epidemiology of TB and HIV. These documents are available at http://www.who.int/tb/dots/laboratory/policy/en.

3.1. Target audience

This policy guidance is intended to be used by managers and laboratory directors working in TB and HIV programmes in coordination with external laboratory consultants, donor agencies, technical advisers, laboratory technicians, procurement officers for laboratory equipment, service providers in the private sector, relevant government sectors, and implementation partners that are involved in country-level strengthening of TB and HIV laboratories. Individuals responsible for programme planning, budgeting, mobilizing resources and implementing training activities for TB and TB–HIV diagnostic services may also benefit from this document.

Date of review: 2018 or earlier should significant additional evidence become available.

4. Evidence base for policy formulation

Nineteen studies were identified through the literature search (Figure 3). These included three ongoing studies that provided data on patient-important outcomes but did not contribute data on diagnostic accuracy of LF-LAM.

The remaining 16 studies evaluated LF-LAM in 6588 HIV-positive individuals, including 1789 (27%) with a microbiological diagnosis of TB. These studies either assessed the accuracy of LF-LAM for the diagnosis of active TB in people with HIV with signs or symptoms of TB, or in people with HIV irrespective of signs or symptoms of TB. All studies were performed in low-income or middle-income countries.

Of 16 studies included in the quantitative analysis, 6 were full–text articles published in peer-reviewed journals, while the other 10 were not published at the time of the systematic review. For the unpublished studies, sets of data were obtained directly from the research teams and the quality of the data was assessed using the same methodology as for published studies. Of 10 unpublished studies, included in the review, by the time this current policy guidance was issued, two were published in peer review journals, six have been published as abstracts in conference proceedings, whereas two studies remained unpublished. Authors of the two unpublished studies were contacted and they agreed that summaries of their studies are made available (Annex 6). The two unpublished studies were included in the meta-analyses and GRADE tables due to the limited number of full text published studies available.

Substantial differences for the following characteristics were noted across studies: purpose for which LF-LAM was applied (diagnosis versus screening); setting (inpatient versus outpatient); threshold used to define LF-LAM positivity (grade 1 versus grade 2); inclusion and exclusion of participants based on whether or not they could produce sputa; whether patients received an evaluation for extrapulmonary TB; and type of reference standard (microbiological versus composite).

Sensitivity and specificity for the assay were determined at both the grade 1 and grade 2 cut-off for positivity (pre-January 2014 grades). For adults, only those studies with grade 2 reactions were finally included in the GRADE assessment as per the manufacturer's instructions after January 2014. For children, the analysis used a combination of grade 1 and grade 2 reactions due to limited data. Unless specifically noted, all sensitivity and specificity analyses were performed with respect to a microbiological reference standard.

Seven of the 16 studies evaluated LF-LAM in individuals with symptoms consistent with TB and were classified as 'studies for the diagnosis of active TB'. The remaining nine studies performed systematic screening with LF-LAM in individuals irrespective of whether signs and symptoms were present and were classified as 'studies for the screening of active TB'. References for the included published and unpublished studies are available in Annex 5.

Figure 3. Selection of studies evaluating the accuracy of LF-LAM for the diagnosis of active TB: flow diagram of studies identified by the literature

The systematic review was limited to adults. However, as a supplement to the review a rapid assessment of studies evaluating LF-LAM in children was performed. One study was identified through the electronic searches on 2 February 2015. Subsequently, four studies were identified (three ongoing and one submitted abstract) through contact with the respective researchers. These studies are cited separately in Annex 5.

4.1. Quality of included studies

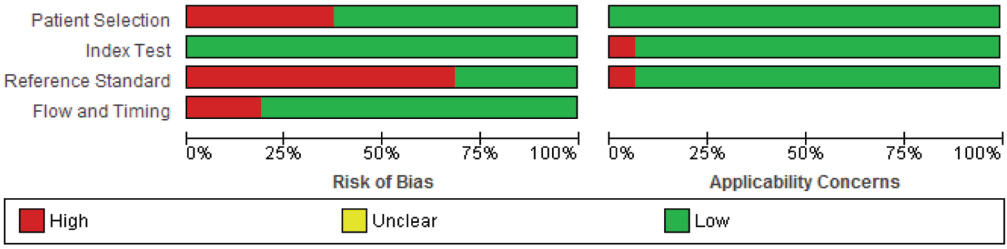
The appraisal of the studies included in the review used the QUADAS-2 tool[9], which consists of four domains: patient selection, index test, reference standard, and flow and timing. For each outcome, the quality of evidence according to GRADE was initially rated as high since all studies were cross-sectional or cohort studies prospectively enrolling patients at risk for having pulmonary or extrapulmonary TB. Downgrading of quality was based on limitations of the studies identified

[9] Whiting PF et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 2011, 155:529–536.

using six GRADE criteria: (1) study design, (2) risk of bias, (3) directness, (4) inconsistency, (5) imprecision, and (6) publication or reporting bias.

Overall quality of the 16 included studies is presented in Figure 4.

Figure 4. Risk of bias and applicability concerns graph: judgements about each domain presented as percentages across the 16 included studies[a]



[a] The appraisal of the studies, which were unpublished at the time of the systematic review, was conducted similarly to that for the published studies.

4.2. LF-LAM accuracy for the diagnosis of active TB in people living with HIV

Of the 16 included studies, seven (44%) evaluated accuracy of LF-LAM for the diagnosis of patients with signs and symptoms of active TB disease. The median CD4 cell count across these seven studies ranged from 71 to 210 cells/μL. All seven studies provided data at the grade 1 test threshold (3126 patients; 39% with TB) and six of these seven studies (86%) also provided data at grade 2 (3037 patients; 38% with TB). These six studies were included in the final GRADE evaluation.

Six (86%) of the seven studies were conducted either exclusively or largely in an inpatient setting and four out of the seven studies (57%) studies included patients from an outpatient setting.

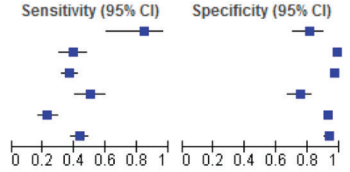
4.2.1. Overall accuracy of LF-LAM compared with a microbiological reference standard

Six studies were included in the analysis of the overall accuracy of LF-LAM (grade 2), involving 3037 HIV-infected patients, 1163 (38%) with microbiologically confirmed TB. For individual studies, sensitivity estimates ranged from 23% to 84%, and specificity estimates from 75% to 99%. The pooled sensitivity across studies was 44% (95% CrI, 31- 60%) and the pooled specificity was 92% (95% CrI, 83-96%) (Figure 5). The lowest sensitivity (23%) was observed in the study by Peter and Theron. The differences observed between this study and the other studies in this analysis included the study setting (outpatients only), a focus on pulmonary TB only (no extrapulmonary samples taken) and the exclusion of patients unable to produce sputa.

Figure 5. Forest plots of sensitivity and specificity of LF-LAM for diagnosis of active TB in people living with HIV against a microbiological reference standard

LAM_DX_MicroRef_Grd2_All

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Andrews 2014	16	14	3	60	0.84 [0.60, 0.97]	0.81 [0.70, 0.89]
Lawn 2014	53	3	83	274	0.39 [0.31, 0.48]	0.99 [0.97, 1.00]
Nakiyingi	136	21	231	609	0.37 [0.32, 0.42]	0.97 [0.95, 0.98]
Peter 2012	58	31	58	94	0.50 [0.41, 0.59]	0.75 [0.67, 0.82]
Peter and Theron 2014	41	27	140	361	0.23 [0.17, 0.29]	0.93 [0.90, 0.95]
Van Rie 2014	149	26	195	354	0.43 [0.38, 0.49]	0.93 [0.90, 0.95]



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Dx - diagnosis; MicroRef - microbiological reference standard; Grd - grade).

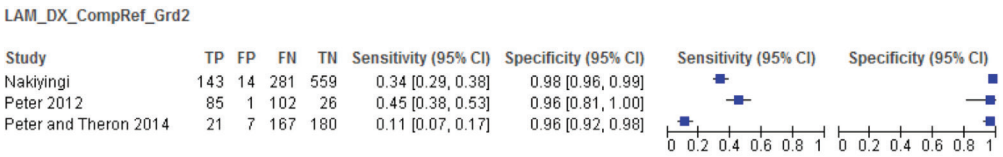
The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (black horizontal line).Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.2.2. Overall accuracy of LF-LAM compared with a composite reference standard

Three studies were included in the analysis involving 1586 HIV-infected patients, 799 (50%) with TB based on a composite diagnosis. In these studies, 270 patients (17%) were considered 'unclassifiable' using the composite reference standard (54 LAM-positive; 216 LAM-negative). For individual studies, sensitivity estimates ranged from 11% to 45% and specificity estimates ranged from 96% to 98%. The pooled sensitivity was 28% (95% CrI, 13- 51%) and the pooled specificity was 97% (95% CrI, 93-99%) (Figure 6).

Compared to using a microbiological reference standard, LF-LAM (grade 2) pooled sensitivity decreased from 44% to 28% when using a composite reference while pooled specificity increased from 92% to 97%. The increased specificity and reduced sensitivity was explained by the broad definition of the reference standard that, in addition to the microbiological criteria, included clinical criteria, i.e., decision to start treatment, and, after at least one month follow-up, a clinical diagnosis of TB.

Figure 6. Forest plot of the sensitivity and specificity of LF-LAM for diagnosis of active TB in people living with HIV against a composite reference standard



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Dx - diagnosis; CompRef - composite reference standard; Gr - grade.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

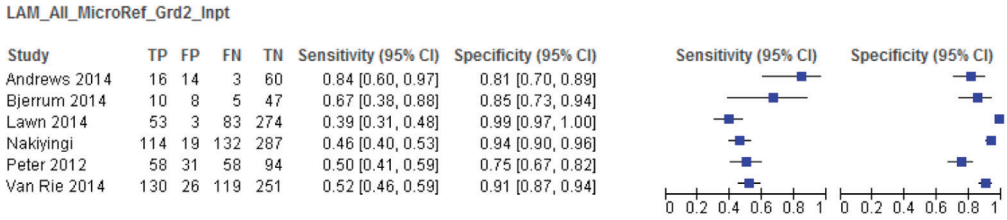
4.2.3. LF-LAM accuracy by health care setting

In comparison with studies in outpatients, studies among inpatients revealed higher pooled sensitivity and lower pooled specificity (see below). The higher sensitivity among inpatients suggests that LF-LAM has improved sensitivity in patients with higher disease severity or higher bacillary burden. The lower pooled specificity among inpatients may be associated with the misclassification of patients (especially those with extrapulmonary TB) and a suboptimal reference standard. Therefore, only studies with a microbiological reference standard were included for analysis of accuracy by health care setting.

4.2.3.1. LF-LAM for the diagnosis of active TB in adult HIV positive inpatients

Six studies were included in the analysis involving 1895 HIV-infected inpatients, 781 (41%) with TB. For individual studies, sensitivity estimates ranged from 39% to 84% and specificity estimates ranged from 75% to 99%. The pooled sensitivity across studies was 54% (95% CrI, 43-67%) and the pooled specificity was 90% (95%CrI, 79-95%). When three studies at high risk of bias for the reference standard were excluded (Andrews 2014, Bjerrum 2014 and Peter 2012), specificity rose to 95% (95%Cri, 85- 98%) with a modest decline in sensitivity to 47% (95% CrI, 32-61%). This finding suggests that specificity estimates were impacted by the application of a reference standard that may have misclassified TB patients as not having TB (Figure 7).

Figure 7. Forest plots of sensitivity and specificity of LF-LAM for diagnosis of active TB against a microbiological reference standard, HIV positive inpatients



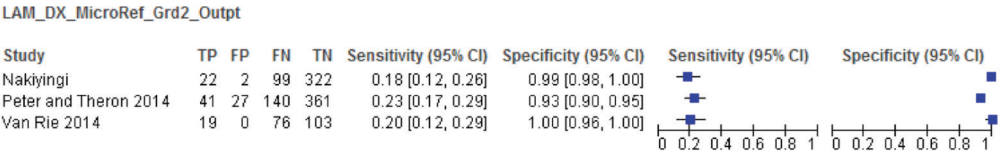
TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; MicroRef - microbiological reference standard; Gr - grade; Inpt - inpatient.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.2.3.2. LF-LAM for the diagnosis of active TB in adult HIV positive outpatients

Three studies were included in the analysis involving 1212 HIV-infected outpatients, 397 (33%) with TB. For individual studies, sensitivity estimates ranged from 18% to 23% and specificity estimates ranged from 93% to 100%. The pooled sensitivity across studies was 21% (95% CrI, 12-34%) and the pooled specificity was 97% (95%CrI, 87- 99%) (Figure 8).

Figure 8. Forest plots of the sensitivity and specificity of LF-LAM for diagnosis of active TB against a microbiological reference standard, HIV positive outpatients



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Dx - diagnosis; MicroRef - microbiological reference standard; Gr - grade; Outpt - outpatient.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.2.4. LF-LAM for the diagnosis of TB in adults living with HIV by CD4 threshold

Five studies were included in the analysis involving 2314 HIV-infected persons whose CD4 thresholds were determined, 819 (35%) with TB. A summary of pooled sensitivity and specificity estimates for LF-LAM (grade 2) stratified by CD4 threshold is given in Table 1.

A 'dose-response' relationship was observed whereby LF-LAM pooled sensitivity increased as the degree of immunodeficiency increased (that is, with decreasing CD4 count) from 15% (95% CrI, 8-27%) in patients with CD4 cell count >200 cells/μL to 49% (95% CrI, 34-66%) in patients with CD4 count ≤200 cells/μL to 56% (95% CrI, 41-70%) in patients with CD4≤100 cells/μL. Pooled specificity also varied across CD4 strata, though less than pooled sensitivity and not likely to be statistically significant (CIs overlap): 96% (95% CrI, 89-99%) for CD4>200, 92% (95% CrI, 78-97%) for CD4>100 cells/μL, and 90% (95% CrI, 81- 95%) for CD4≤100 cells/μL.

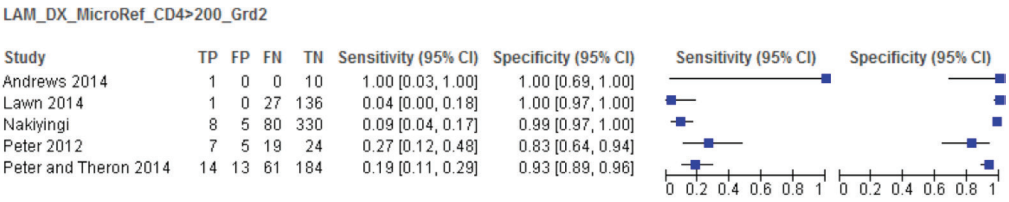
Table 1. Summary of LF-LAM diagnostic accuracy for active TB at different CD4 thresholds

CD4 Threshold	Studies (total participants, % with TB)	Pooled sensitivity (95% credible interval)	Pooled specificity (95% credible interval)
CD4 > 200	5 studies (925, 24% TB)	15% (8, 27)	96% (89, 99)
CD4 ≤ 200	5 studies (1344, 45% TB)	49% (34, 66)	90% (78, 95)
CD4 > 100	5 studies (1410, 30% TB)	26% (16, 46)	92% (78, 97)
CD4 ≤ 100	5 studies (859, 47% TB)	56% (41, 70)	90% (81, 95)

4.2.4.1. LF-LAM for the diagnosis of active TB in adults living with HIV and CD4 count greater than 200 cells/µL

Five studies were included in the analysis, involving 925 HIV infected persons whose CD4 counts were greater than 200 cells/µL, of which 218 (24%) with TB. For individual studies, sensitivity estimates ranged from 4% to 100% and specificity estimates ranged from 83% to 100%, The pooled sensitivity across studies was 15% (95%CrI, 8-27%) and the pooled specificity was 96% (95% CrI, 89-99%) (Figure 9).

Figure 9. Forest plots of sensitivity and specificity of LF-LAM for diagnosis of active TB in HIV positive patients with CD4 count > 200 cells/µL, microbiological reference standard



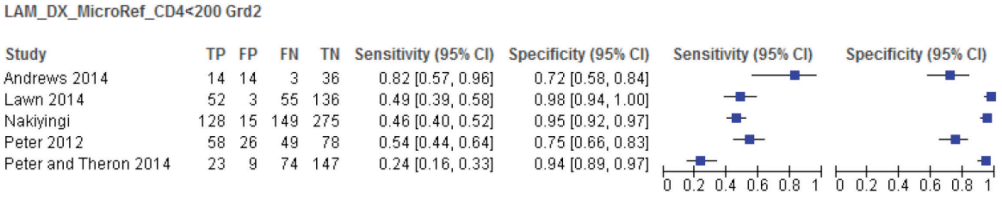
TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Dx - diagnosis; MicroRef - microbiological reference standard; Gr - grade.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.2.4.2. LF-LAM for the diagnosis of active TB in adults living with CD4 count less than or equal to 200 cells /µL

Five studies were included in the analysis involving 1344 HIV infected persons whose CD4 counts were less than or equal to 200 cells /µL, 605 (45%) with TB. For individual studies, sensitivity estimates ranged from 24% to 82% and specificity estimates ranged from 72% to 98%, The pooled sensitivity across studies was 49% (95%CrI, 34-66%) and the pooled specificity was 90% (95% CrI, 78-95%) (Figure 10).

Figure 10. Forest plots of sensitivity and specificity of LF-LAM for diagnosis of active TB in HIV positive patients with CD4 count ≤ 200 cells/µL, microbiological reference standard



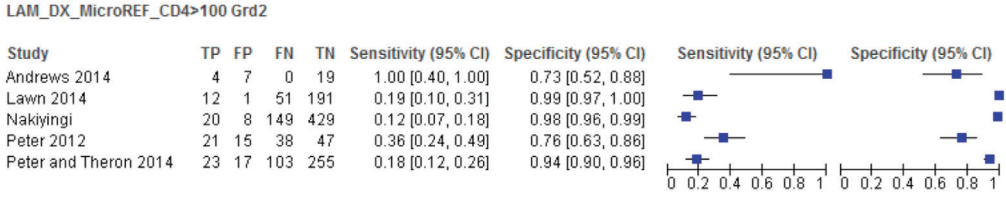
TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Dx - diagnosis; MicroRef - microbiological reference standard; Gr - grade.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.2.4.3. LF-LAM for the diagnosis of active TB in adults living with with CD4 count greater than 100 cells /µL

Five studies were included in the analysis, involving 1410 HIV infected persons whose CD4 counts were greater than 100 cells/µL, of which 421 (30%) with TB. For individual studies, sensitivity estimates ranged from 12% to 100% and specificity estimates ranged from 73% to 99%. The pooled sensitivity across studies was 26% (95%CrI, 16-46%) and the pooled specificity was 92% (95% CrI, 72-97%) (Figure 11).

Figure 11. Forest plots of sensitivity and specificity of LF-LAM for diagnosis of active TB in HIV positive patients with CD4 count > 100 cells /µL, microbiological reference standard



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Dx - diagnosis; MicroRef - microbiological reference standard; Gr - grade.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

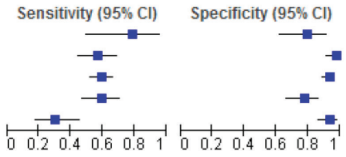
4.2.4.4. LF-LAM for the diagnosis of active TB in adults living with HIV and CD4 count less than or equal to 100 cells /µL

Five studies were included in the analysis involving 859 HIV-infected persons whose CD4 counts were less than or equal to 100, of which 402 (47%) with TB. For individual studies, sensitivity estimates ranged from 30% to 79% and specificity estimates ranged from 79% to 98%. The pooled sensitivity across studies was 56% (95%CrI, 41-70%) and the pooled specificity was 90% (95% CrI, 91-95%) (Figure 12).

Figure 12. Forest plots of sensitivity and specificity of LF-LAM for diagnosis of active TB in HIV positive patients with CD4 count ≤ 100 cells/μL, microbiological reference standard

LAM_DX_MicroRef_CD4<100 Grd2

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Andrews 2014	11	7	3	27	0.79 [0.49, 0.95]	0.79 [0.62, 0.91]
Lawn 2014	41	2	31	81	0.57 [0.45, 0.69]	0.98 [0.92, 1.00]
Nakiyingi	116	12	80	176	0.59 [0.52, 0.66]	0.94 [0.89, 0.97]
Peter 2012	44	16	30	55	0.59 [0.47, 0.71]	0.77 [0.66, 0.87]
Peter and Theron 2014	14	5	32	76	0.30 [0.18, 0.46]	0.94 [0.86, 0.98]



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Dx - diagnosis; MicroRef - microbiological reference standard; Gr grade.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.2.5. LF-LAM compared to conventional tests for the diagnosis of TB in adults

Four studies were identified (Lawn 2014, Nakiyingi 2014, Peter 2012, Peter and Theron 2014) that directly compared LF-LAM and sputum microscopy in the same patients, involving 800 HIV-infected patients with TB (36%) and 2,220 HIV-infected patients without TB.

Three studies (Lawn 2014, Nakiyingi 2014, Peter and Theron 2015) were identified that directly compared LF-LAM and Xpert MTB/RIF in the same patients, involving 1,223 HIV-infected patients, of which 433 (35,4%) with TB. Because the conventional tests required sputum, the analysis was restricted to those patients who could produce sputa, possibly introducing bias for patient selection. Results should therefore be interpreted with caution.

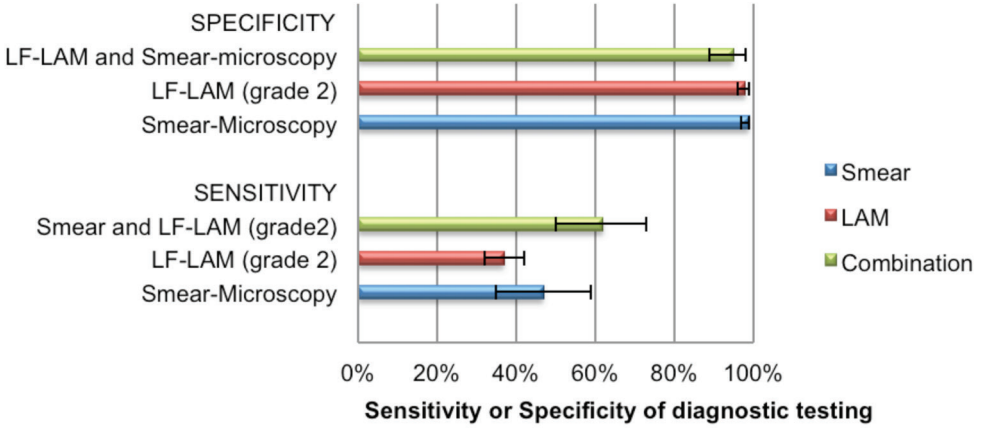
4.2.5.1. LF-LAM in comparison with microscopy for the diagnosis of active TB in adults living with HIV

The pooled sensitivity of LF-LAM was 37% (95% CrI, 32-42%) compared with 47% (95% CrI, 35-59%) for sputum microscopy (using a microbiological reference standard. The pooled specificity of LF-LAM was 95% (95% CrI, 93-97%) versus 98% (95% CrI, 93-100%) for sputum microscopy (Figure 13).

4.2.5.2. LF-LAM in combination with microscopy for the diagnosis of active TB in adults living with HIV

The pooled sensitivity of a combination of LF-LAM and sputum microscopy (either test positive, microbiological reference standard) was 62% (95% CrI, 50-73%) which was higher than for either test alone. The pooled specificity of LF-LAM combined with microscopy (microbiological reference standard) was 91% (95% CrI 73-96%) which was lower than for either test alone (Figure 13).

Figure 13. Sensitivity and specificity of LF-LAM in combination with and in comparison to sputum microscopy for TB diagnosis in adults living with HIV



Bars represent point estimates of pooled sensitivity or specificity of tests used individually or in combination. Error bars represent 95% credible intervals.

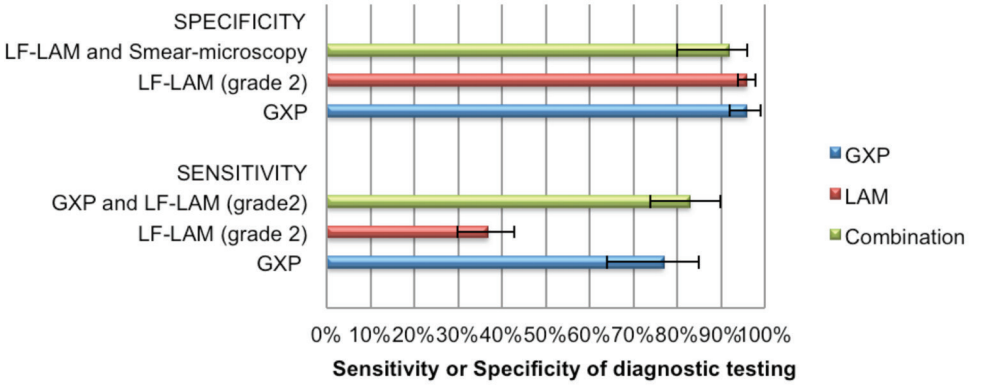
4.2.5.3. LF-LAM in comparison with Xpert MTB/RIF for the diagnosis of active TB in adults living with HIV

The pooled sensitivity of Xpert MTB/RIF was 77% (95% CrI, 64-85%) which was significantly higher than for LF-LAM at 37% (95% CrI, 30-43%), with a microbiological as reference standard The pooled specificity of LF-LAM was 96% (95% CrI, 94-98%) compared with 96% (95% CrI, 92-99%) for Xpert MTB/RIF (Figure 14). Studies were restricted to those patients who could produce sputa, which biased these studies for patient selection. Results should therefore be interpreted with caution.

4.2.5.4. LF-LAM in combination with Xpert MTB/RIF for the diagnosis of active TB in adults living with HIV

The pooled sensitivity of a combination of LF-LAM and Xpert MTB/RIF (either test positive) was 83% (95% CrI, 74-90%) which was slightly higher than for Xpert MTB/RIF alone (see 4.2.5.3) using a microbiological reference standard). The pooled specificity was 91% (95% CrI,81-96%) which was lower than for either test alone (see 4.2.5.3, Figure 14).

Figure 14. Sensitivity and specificity of LF-LAM in combination with and in comparison to Xpert MTB/RIF for TB diagnosis in adults living with HIV



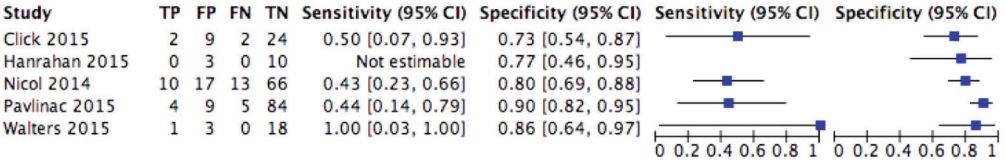
Legend: GXP, sputum Xpert MTB/RIF. Bars represent point estimates of pooled sensitivity or specificity of tests used individually or in combination. Error bars represent 95% credible intervals.

4.2.6. LF-LAM for the diagnosis of TB in children living with HIV compared with a microbiological reference standard

Five studies (three ongoing) involving 280 HIV-infected paediatric patients, 37 (12%) with TB were included. The study by Nicol 2014, contributed 23 (62%) of the TB patients. All studies were prospective cohort studies including a follow-up period of three to six months. Studies took place in both inpatient and outpatient settings. Three studies were conducted in South Africa and two in Kenya. Studies differed in the number and types of specimens collected to confirm TB.

For individual studies, sensitivity estimates ranged from 43% to 100% and specificity estimates ranged from 73% to 90%. The pooled sensitivity across studies was 47% (95% CrI, 27-69%) and the pooled specificity was 82% (95% CrI, 71-89%) (Figure 15). In a subgroup analysis of four studies involving children aged birth to four years (141 HIV-infected children, nine (6%) with TB), pooled sensitivity was 47% (95% CrI, 17-80%) and the pooled specificity was 82% (95% CrI, 68-91%). There were insufficient data to perform a meta-analysis for children aged five to 14 years.

Figure 15. Forest plots of sensitivity and specificity of LF-LAM for TB diagnosis in children living with HIV



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

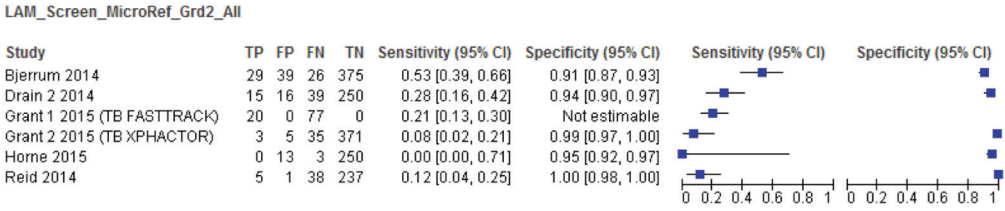
4.3. LF-LAM for screening of active TB in people living with HIV

4.3.1. Overall accuracy of LF-LAM to screen for TB in adults living with HIV compared with a microbiological reference standard

Six studies were included, involving 1847 HIV-infected patients, 290 (16%) with TB. For individual studies, sensitivity estimates ranged from 0% to 53%, and specificity estimates from 91% to 100%.

The pooled sensitivity across studies was 23% (95% CrI, 13- 37%) and the pooled specificity was 96% (95% CrI, 92-98%) (Figure16). The accuracy estimates for LF-LAM for screening contrast with those of LF-LAM for diagnosis [pooled sensitivity 44% (31% to 60%) and pooled specificity 92% (83% to 96%).

Figure 16. Forest plots of sensitivity and specificity of LF-LAM for TB screening in adults living with HIV, microbiological reference standard



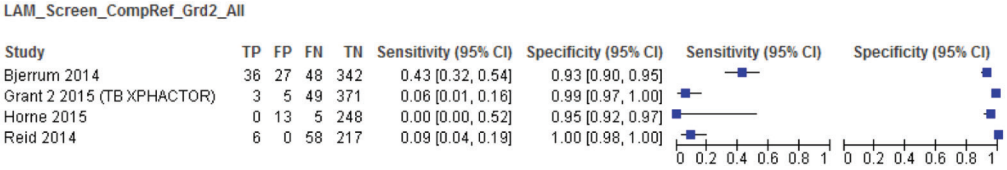
TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Screen - screening; MicroRef - microbiological reference standard; Grd - grade.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line. Values for test results (LF-LAM grade 2reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.3.2. Overall accuracy of LF-LAM to screen for TB in adults living with HIV compared with a composite reference standard

Four studies were identified that included 1428 HIV-infected patients, 205 (14%) with TB. For individual studies, sensitivity estimates ranged from 0% to 43%, and specificity estimates ranged from 93% to 100%. The pooled sensitivity across studies was 18% (95%CrI, 7- 35%) and the pooled specificity was 97% (95% CrI, 92- 99%) (Figure 17). Compared to using a microbiological reference standard [pooled sensitivity 23% (95% CrI, 13-37%)], pooled sensitivity decreased using a composite reference standard whilespecificity remained essentially unchanged (97% composite versus 96% microbiological).

Figure 17. Forest plots of sensitivity and specificity of LF-LAM for TB screening in adults living with HIV, composite reference standard



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Screen - screening; CompRef - composite reference standard; Grd - grade.

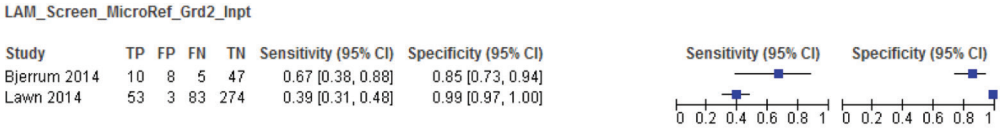
The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.3.3. LF-LAM for screening for TB in adults living with HIV by health care setting

4.3.3.1. LF-LAM for screening for TB in adult HIV positive inpatients

Two studies were identified that included 483 HIV-infected inpatients, 151 (31%) with TB. The sensitivities for the two studies were 39% and 67% and corresponding specificities were 99% and 85%. The pooled sensitivity was 52% (95% CrI, 29-76%) and the pooled specificity was 94% (95% CrI, 74-98%) (Figure 18).

Figure 18. Forest plots of sensitivity and specificity of LF-LAM for TB screening, HIV positive inpatients, microbiological reference standard



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Screen - screening; MicroRef - microbiological reference standard; Grd - grade; Inpt - inpatients.

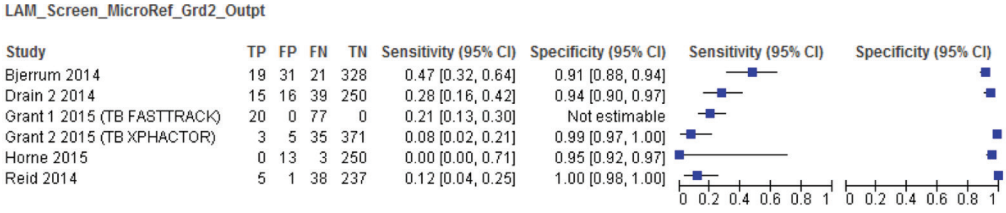
The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.3.3.2. LF-LAM for screening for TB in adult HIV positive outpatients

Six studies were identified that included 1777 HIV-infected inpatients, 275 (15%) with TB. In the individual studies, sensitivity estimates ranged from 0% to 47% and specificity estimates ranged from 91% to 100%, The pooled sensitivity across studies was 22% (95% CrI, 13-35%) and the pooled specificity of 96% (95% CrI, 93-98%) (Figure 19).

Compared with LF-LAM for TB screening among inpatients living with HIV, pooled sensitivity among outpatients decreased considerably (22% versus 52%), while specificity was comparable (96% versus 94%).

Figure 19. Forest plots of sensitivity and specificity of LF-LAM for TB screening HIV positive outpatients, microbiological reference standard



TP - true positive; FP - false positive; FN - false negative; TN - true negative; CI - confidence interval; Screen - screening; MicroRef - microbiological reference standard; Grd - grade; Outptt - outpatients.

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). Values for test results (LF-LAM grade 2 reactions only) are the number of each type of result (true positive, false positive, false negative, true negative).

4.3.4. LF-LAM for screening for TB in adults living with HIV by CD4 threshold

Four studies were identified that only performed an analysis based on LF-LAM Grade 1 results, stratified by CD4 levels. The LF-LAM grade 1 represents a very low intensity band and is no longer recommended use by the manufacturer. The data were insufficient to conduct CD4 stratified analyses for LF-LAM for TB screening at grade 2, hence no GRADE tables were created for PICO question 7.

4.4. LF-LAM and patient outcomes

Patient outcomes in addition to diagnostic accuracy could not be systematically addressed given limited data. Nonetheless, any available data on patient outcomes were summarized from the included studies, including data from secondary analyses.

Six of the nineteen included studies provided data on the association of LF-LAM and mortality, either in peer-reviewed publications, related reports or in unpublished data (Balcha 2014, Lawn 2012a, Nakiyingi 2014, Peter and Theron 2014, Bjerrum 2014, Drain 2015). These studies were conducted in both inpatient and outpatient settings. Data on patient outcomes were largely restricted to post-hoc analyses in these studies. Nonetheless, available data consistently suggested higher disease severity among LF-LAM positive TB patients than LF-LAM negative TB patients. All six studies showed a consistent finding of increasing mortality with LF-LAM positivity, despite considerable variability in the length of follow-up, method of TB diagnosis, and provision of treatment. As LF-LAM was not used to decide TB treatment initiation in any of the studies, this association may be the result of LF-LAM positive individuals having delayed or missed TB diagnoses prior to LF-LAM testing.

A study by Balcha 2014 reported significantly higher mortality (20% versus 3%, p < 0.001) in LF-LAM positive than LF-LAM negative patients. Similarly, another study by Manabe 2014[10] (secondary analysis by Nakiyingi 2014) reported higher mortality (28% versus 13%, p = 0.035) in LF-LAM positive TB patients and additionally found higher mortality in LF-LAM positive patients without microbiological evidence of TB (34% versus 19% for LF-LAM positive and LF-LAM for negative patients, respectively). The study by Lawn 2012, found that among 23 TB patients who were LF-LAM positive, five patients died (22%) compared to zero deaths among 36 TB patients who were LF-LAM negative. Another study (Lawn 2013[11]) reported that LF-LAM sensitivity was 100% among TB patients who died compared to 25% among TB patients who were alive at 90 days (p = 0.002). Peter and Theron reported mortality of 25% (9/32) LF-LAM positive and 11% (40/361) LF-LAM negative patients. Unpublished study 2 (Reid et al.) reported that among 469 patients 40% of those who were LAM positive died versus 13% of those who were LF-LAM negative (p < 0.001). Among 55 TB patients, 52% of LAM positive patients died compared to only 12% of those who were LF-LAM negative (p = 0.002).

Bjerrum 2014, reported that among 38 TB patients who received TB treatment, 22% (4/18) of those who were LF-LAM positive died compared to only 5% (1/20) of those who were LF-LAM negative. In this study, LF-LAM results were unavailable to clinicians for clinical decision making. Another post-hoc analysis (Peter 2013) reported that among inpatients, LF-LAM-positive TB patients missed by empirical early treatment had lower CD4 counts and higher median illness severity scores, compared to patients who received early treatment based on clinical decision making. A study by Drain 2015 reported LF-LAM responses over time, showing that among patients receiving TB therapy, having a positive LF-LAM test at the two-month visit was associated with an adjusted hazard ratio (HR) of 5.58 for mortality (median follow-up time of 49 months) compared to patients with a negative LF-LAM test at the two-month visit. Participants with a positive LF-LAM at six months had an adjusted HR of 42.1 for mortality during study follow-up. No difference (adjusted HR 1.41, p = 0.49) in mortality was found comparing baseline LF-LAM results.

4.5. Cost and cost-effectiveness of using LF-LAM for the diagnosis of active TB

A systematic review of economic evaluations of LF-LAM for diagnosis of active tuberculosis (TB) in HIV-infected individuals was performed. Two studies were identified, both evaluating populations in sub-Saharan Africa and with a focus on inpatient populations with CD4 counts less than 100 cells/μL. References for the included and excluded studies are given in Annex 6. Both studies found the addition of LF-LAM to existing diagnostic strategies based on sputum smear microscopy or Xpert MTB/RIF to be highly cost-effective across a wide array of sensitivity analyses, although addition of costs related to future HIV care in one study caused cost-effectiveness ratios to become substantially less favorable.

Incremental cost-effectiveness ratios (without inclusion of HIV care costs) ranged from $21 to $265 per disability-adjusted life year (DALY) averted for the addition of LF-LAM to existing diagnostic algorithms based on sputum smear microscopy in Uganda, from $10 to $3,162 per DALY averted when added to Xpert MTB/RIF-based algorithms in Uganda, and from $135 to $8,707 for addition of LF-LAM to existing diagnosis (including smear microscopy or Xpert MTB/RIF, culture and/or clinical judgment) in South Africa. The most important drivers of cost-effectiveness were the specificity of LF-LAM, the prevalence of active TB in the target population, the life expectancy of TB survivors, and the costs of TB and HIV treatment.

[10] Annex 5 -Related report 1
[11] Annex 5 -Related report 2

While the ability of a positive LF-LAM result to avert mortality was varied in both studies, it was not an important driver of overall cost-effectiveness, as TB and HIV treatment costs dwarfed the cost of LF-LAM itself.

It is important to note that the economic data to support or refute the use of LF-LAM are extremely sparse, consisting at present of two studies, performed largely via modelling with similar techniques, in similar settings, and with overlapping authorship teams. The study populations were dominated by hospitalized patients; thus, findings cannot be generalized to outpatient settings, where the prevalence of TB (and the probability of rapid death if TB remains untreated) is generally lower, and the sensitivity of LF-LAM alone for TB screening is poor.

Consideration of additional patient populations such as those with CD4 count <50 cells/μL, those with presumptive TB who cannot produce sputum, critically ill patients with presumptive extrapulmonary TB, or patients being screened in combination with other tests (e.g., chest X-ray) could further inform the cost-effectiveness of LF-LAM. While findings in these two studies appeared robust to an array of sensitivity analyses in the settings studied, further evaluations by independent groups carried out in alternative settings are necessary before making definitive conclusions about the cost-effectiveness of LF-LAM more broadly.

4.6. Inter- and intra-reader variability between test readers in studies, included in the review

Concerning inter-reader variability, there was a high degree of agreement; studies reported on inter-reader variability in the form of a Kappa statistic or percent concordance between multiple readers. Four studies of TB diagnosis (Lawn 2014; Nakiyingi 2014; Peter 2012; Peter 2015) reported Kappa statistics ranging from 0.78 to 0.97, with the majority reporting values > 0.92. Two studies of TB screening (Bjerrum 2014, Lawn 2012) reported Kappa values of 0.92 to 0.97, and one study (Balcha 2014) reported 100% concordance between readers. There were limited data on intra-reader agreement; Peter 2012 reported a Kappa statistic of 0.92 to 0.96, indicating very good agreement.

5. Summary of evidence to recommendations

5.1. LF-LAM for the diagnosis of active TB

Based on the GRADE process the Guideline Development Group (GDG) determined that the overall quality of the evidence was low, largely due to patient selection bias and the limitations of the reference standards used in the different studies, many of which were still unpublished at the time of systematic review (Tables 2-16).

For TB diagnosis among symptomatic patients, overall LF-LAM pooled sensitivity was 44% and pooled specificity was 92%. If these values are applied to hypothetical cohorts of patients in different epidemiological settings (Tables 2-16), the balance between false and true diagnoses is such that the Guideline Development Group recommended that LF-LAM cannot be relied on as a stand-alone test for detecting TB. For example, in a hypothetical cohort of 1000 HIV-infected patients where 10% of those with symptoms actually have TB (typically found in high TB-HIV settings), LF-LAM will miss more patients with TB (56) than correctly diagnosing those with active TB (44).

The Guideline Development Group felt that both false-positive TB and false-negative diagnosis may contribute to harm for patient. False-negatives may contribute to large harm, which includes delay of TB diagnosis, continued TB transmission, and increased mortality. False-positives are also a concern given that they may contribute small to large harms, including possible adverse events of unnecessary treatment and delay in diagnosis of a different disease. It is important to consider false-positive diagnosis also from the patient perspective, who may be concerned with the stigma associated with a positive test result.

Considering these benefits and harms the GDG made a strong recommendation against using the test to diagnose TB in all persons with HIV (see Section 6, recommendation 1).

The evidence assessment showed increased sensitivity of LF-LAM in the sickest patients, notably among inpatients with low CD4 counts. The pooled sensitivity and specificity of LF-LAM in inpatients with CD4 thresholds ≤ 100 cells/μL were 61% and 89% respectively. When these estimates are applied to a hypothetical cohort of 1000 HIV-infected patients with CD4 count less than or equal to 100 cells/μL, where 30% of those with symptoms actually have TB, LF-LAM will correctly exclude the majority of patients without TB (623 out of 700) and correctly diagnose the majority with TB (183 out of 300). Given the need for rapid exclusion of TB in this subgroup of patients and the fact that LF-LAM does not require sputum collection, the Guideline Development Group suggested that this sub-group of patients may benefit from LF-LAM testing. However, given the test specificity, there was also need to take into account the downstream consequences of patients incorrectly diagnosed with TB. Overall, the Guideline Development Group felt that the net benefit of using the LF-LAM in the subpopulation at high risk of mortality outweighed the harms associated with false-positive diagnosis.

Considering these benefits and harms the GDG made a conditional recommendation for using test to diagnose TB in seriously ill patients with HIV (see Section 6, recommendation 2).

5.2. LF-LAM for the screening for TB

For TB screening among patients regardless of TB symptoms, LF-LAM pooled sensitivity was 23% and the pooled specificity was 96% compared with a microbiological reference standard. As such, the Guideline Development Group felt that a screening strategy based on LF-LAM alone cannot be relied on. For example, if pooled sensitivity and specificity estimates for LF-LAM are applied to a hypothetical cohort of 1000 HIV-infected individuals being screened for TB regardless of symptoms, where 1% actually have TB, LF-LAM will correctly diagnose only 2 out of 10 individuals with TB while falsely diagnosing 40 out of 990 without TB. The Guideline Development Group found no evidence that LF-LAM improve patient-important outcomes, and high proportions of false-positive and false-negative results may adversely impact patients. It is strongly recommended that these tests not be used for the screening for TB (Tables 17-20).

Considering these benefits and harms the GDG made a strong recommendation against using test to diagnose TB in all persons with HIV (see Section 6, recommendation 3).

6. WHO policy recommendations

Given the GRADE evidence assessment and considering the relative benefits and harms associated with the use of the LF-LAM assay, WHO recommends that:

1. **Except as specifically described below for persons with HIV infection with low CD4 counts or who are seriously ill,**[12] LF-LAM **should not be** used for the diagnosis of TB **(strong recommendation, low quality of evidence)**.

2. LF-LAM **may be used** to assist in the diagnosis of TB **in HIV positive** adult *in-patients* with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count less than or equal to 100 cells/μL, or HIV positive patients who are seriously ill[12] regardless of CD4 count or with unknown CD4 count **(conditional recommendation; low quality of evidence)**.

   Remarks

   a. This recommendation also applies to HIV positive adult *out-patients* with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count less than or equal to 100 cells/μL, or HIV positive patients who are seriously ill[12] regardless of CD4 count or with unknown CD4 count, based on the generalisation of data from in-patients.

   b. This recommendation also applies to HIV positive children with signs and symptoms of TB (pulmonary and/or extrapulmonary) based on the generalisation of data from adults while acknowledging very limited data and concern regarding low specificity of the LF-LAM assay in children.

3. LF-LAM **should not be used** as a screening test for TB. (strong recommendation; low quality of evidence).

[12] "seriously ill" is defined based on 4 danger signs: respiratory rate > 30/min, temperature > 39°C, heart rate > 120/min and unable to walk unaided.
World Health Organization. Improving the diagnosis and treatment of smear-negative pulmonary and extrapulmonary tuberculosis among adults and adolescents. Recommendations for HIV-prevalent and resource constrained settings. World Health Organization 2007.
Available at: http://www.uphs.upenn.edu/bugdrug/antibiotic_manual/smear_neg_and_extrapulmTb.pdf

7. Implementation considerations

Even with targeted use of the LF-LAM assay (in HIV positive adult in-patients with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count less than or equal to 100 cells/µL, or in HIV positive patients who are seriously ill regardless of CD4 count or with unknown CD4 count), the following implementation considerations apply:

- LF-LAM does not differentiate between the various species of mycobacterium and cannot be used to distinguish *M. tuberculosis* from other species. However, in areas endemic for tuberculosis the LAM antigen detected in a clinical sample is likely to be attributed to *M. tuberculosis*.
- Implementation of LF-LAM in the targeted patient groups does not eliminate the need for other diagnostic tests - Xpert MTB/RIF, culture or sputum-smear microscopy - as these tests exceed LF-LAM in diagnostic accuracy. Whenever possible, a positive LF-LAM should be followed up with a confirmation test such as Xpert MTB/RIF, line probe assay or bacteriological culture and drug-susceptibility testing.
- LF-LAM is designed to detect mycobacterial LAM antigen in human urine. Other samples (e.g. sputum, serum, plasma, CSF or other body fluids) or pooled urine specimens should not be used.
- LF-LAM test cards must be stored at 2-30°C until the expiration date. Kit components are stable until the expiration date when handled and stored as directed. Devices that have become wet or if the packaging has become damaged should not be used.

7.1. Plans for disseminating the WHO policy guidance on LF-LAM

This WHO policy guidance will be published online (http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/) and disseminated through WHO/GTB and HIV Department listserves to all WHO Regional and Country Offices, Member States, the Global Laboratory Initiative, the TB/HIV Working Group and New Diagnostics Working Groups of Stop TB Partnership, donors, technical agencies and other stakeholders.

8. Research needs

Current recommendations on the commercial LF-LAM test should not prevent or restrict further research on new TB diagnostics, especially point-of-care assays that can be used as close as possible to where patients access TB treatment. Further operational research on the LF-LAM test should focus on the following priorities:

- Evaluation of diagnostic algorithms in different epidemiological and geographical settings and patient populations;
- Conducting more rigorous studies with higher quality reference standards (including multiple specimen types, also extrapulmonary) to improve confidence in specificity estimates.
- Determine training, competency, and quality assessment needs;
- Gathering more evidence on the impact on TB treatment initiation and mortality;
- Perform country-specific cost-effectiveness and cost-benefit analyses of targeted LF-LAM use in different programmatic settings.

Table 2. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV, microbiological reference standard, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV, microbiological reference standard, grade 2?

Participants: Adults (15 years and older) infected with HIV, presumed to have TB
Setting: Mainly inpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.44 (95% CrI: 0.31 to 0.60)
Pooled specificity	0.92 (95% CrI: 0.83 to 0.96)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	6 studies 1163 patients	cross-sectional (cohort type accuracy study)	not serious[1]	not serious[2]	not serious	serious[3]	not serious	4 (3 to 6)	44 (31 to 60)	132 (93 to 180)	⨁⨁⨁◯ Moderate
False negatives (patients incorrectly classified as not having tuberculosis)								6 (7 to 4)	56 (69 to 40)	168 (207 to 120)	
True negatives (patients without tuberculosis)	6 studies 1874 patients	cross-sectional (cohort type accuracy study)	serious[4]	not serious[2]	not serious	serious[5]	not serious	911 (822 to 950)	828 (747 to 864)	644 (581 to 672)	⨁⨁◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								79 (168 to 40)	72 (153 to 36)	56 (119 to 28)	

1. QUADAS-2 was used to assess risk of bias. Two studies were considered to be at high risk of bias for patient selection; one study excluded patients who were unable to produce sputum and one study only included patients with a suspicion of disseminated TB. Evidence was not downgraded.
2. Five of the included studies were performed in inpatient settings.
3. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.
4. QUADAS-2 was used to assess risk of bias. Three studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
5. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 3. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV, composite reference standard, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV, composite reference standard, grade 2?

Participants: Adults (15 years and older) infected with HIV, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture, NAAT, smear, clinical findings

Pooled sensitivity	0.28 (95% CrI: 0.13 to 0.51)
Pooled specificity	0.97 (95% CrI: 0.93 to 0.99)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	3 studies 799 patients	cross-sectional (cohort type accuracy study)	serious[1]	not serious	not serious	serious[2]	not serious	3 (1 to 5)	28 (13 to 51)	84 (39 to 153)	⨁⨁◯◯ low
False negatives (patients incorrectly classified as not having tuberculosis)								7 (9 to 5)	72 (87 to 49)	216 (261 to 147)	
True negatives (patients without tuberculosis)	3 studies 787 patients	cross-sectional (cohort type accuracy study)	not serious[3]	not serious	not serious	not serious	not serious	960 (921 to 980)	873 (837 to 891)	679 (651 to 693)	⨁⨁⨁⨁ high
False positives (patients incorrectly classified as having tuberculosis)								30 (69 to 10)	27 (63 to 9)	21 (49 to 7)	

1. QUADAS-2 was used to assess risk of bias. A composite reference standard was considered to be at high risk of bias for determining sensitivity. The quality of the evidence was downgraded by one point..
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.
3. QUADAS-2 was used to assess risk of bias. All studies were considered to be at low risk of bias for the composite reference standard.

Table 4. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in children with HIV

PICO Question: Should LF-LAM be used to diagnose tuberculosis in children with HIV?

Participants: Children (0-15 years) infected with HIV, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.47 (95% CrI: 0.27 to 0.69)
Pooled specificity	0.82 (95% CrI: 0.71 to 0.89)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	5 studies 37 patients	cross-sectional (cohort type accuracy study)	not serious	serious[1]	not serious	serious[2]	not serious	5 (3 to 7)	47 (27 to 69)	141 (81 to 207)	⨁⨁◯◯ low
False negatives (patients incorrectly classified as not having tuberculosis)								5 (7 to 3)	53 (73 to 31)	159 (219 to 93)	
True negatives (patients without tuberculosis)	5 studies 243 patients	cross-sectional (cohort type accuracy study)	not serious	serious[3]	not serious	serious[4]	not serious	812 (703 to 881)	738 (639 to 801)	574 (497 to 623)	⨁⨁◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								178 (287 to 109)	162 (261 to 99)	126 (203 to 77)	

1. In all studies, there were few children with TB (i.e., true positives plus false negatives). For determination of sensitivity, one study contributed 62% of the TB patients. The quality of the evidence was downgraded by one point..
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point..
3. In all studies, there were few participants without TB (i.e., true negatives plus false positives). For the determination of specificity, two studies contributed most of the data. The quality of the evidence was downgraded by one point.
4. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 5. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV in inpatient setting, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV in inpatient settings, grade 2?

Participants: Adults (15 years and older) inpatients, infected with HIV, presumed to have TB
Setting: Inpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.54 (95% CrI: 0.43 to 0.67)
Pooled specificity	0.90 (95% CrI: 0.79 to 0.95)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	6 studies 781 patients	cross-sectional (cohort type accuracy study)	not serious[1]	not serious	not serious	serious[2]	not serious	5 (4 to 7)	54 (43 to 67)	162 (129 to 201)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								5 (6 to 3)	46 (57 to 33)	138 (171 to 99)	
True negatives (patients without tuberculosis)	6 studies 1114 patients	cross-sectional (cohort type accuracy study)	serious[3]	not serious	not serious	serious[4]	not serious	891 (782 to 941)	810 (711 to 855)	630 (553 to 665)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								99 (208 to 49)	90 (189 to 45)	70 (147 to 35)	

1. QUADAS-2 was used to assess risk of bias. Two studies were considered to be at high risk of bias. One study excluded patients who were unable to produce sputum and one study only included patients with a suspicion of disseminated TB. Evidence was not downgraded.
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. This was a borderline judgment. The quality of the evidence was downgraded by one point.
3. QUADAS-2 was used to assess risk of bias. Three studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
4. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 6. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV in outpatient setting, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV in outpatient settings, grade 2?

Participants: Adults (15 years and older), outpatients, infected with HIV, presumed to have TB
Setting: Outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.21 (95% CI: 0.12 to 0.34)
Pooled specificity	0.97 (95% CI: 0.87 to 0.99)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	3 studies 397 patients	cross-sectional (cohort type accuracy study)	serious[1]	not serious	not serious	not serious	not serious	2 (1 to 3)	21 (12 to 34)	63 (36 to 102)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								8 (9 to 7)	79 (88 to 66)	237 (264 to 198)	
True negatives (patients without tuberculosis)	3 studies 815 patients	cross-sectional (cohort type accuracy study)	not serious	not serious	not serious	serious[2]	not serious	960 (861 to 980)	873 (783 to 891)	679 (609 to 693)	⊕⊕⊕◯ moderate
False positives (patients incorrectly classified as having tuberculosis)								30 (129 to 10)	27 (117 to 9)	21 (91 to 7)	

1. QUADAS-2 was used to assess risk of bias. Two studies were considered to be at high risk of bias for patient selection. The quality of the evidence was downgraded by one point.
2. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 7. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV with CD4 > 200 cells/µL, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 > 200 cells/µL, grade 2?

Participants: Adults (15 years and older), infected with HIV, with CD4 > 200 cells/µL, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.15 (95% CrI: 0.08 to 0.27)
Pooled specificity	0.96 (95% CrI: 0.89 to 0.99)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	5 studies 218 patients	cross-sectional (cohort type accuracy study)	not serious[1]	not serious	not serious	serious[2]	not serious	2 (1 to 3)	15 (8 to 27)	45 (24 to 81)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								8 (9 to 7)	85 (92 to 73)	255 (276 to 219)	
True negatives (patients without tuberculosis)	5 studies 707 patients	cross-sectional (cohort type accuracy study)	serious[3]	not serious	not serious	serious[4]	not serious	950 (881 to 980)	864 (801 to 891)	672 (623 to 693)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								40 (109 to 10)	36 (99 to 9)	28 (77 to 7)	

1. QUADAS-2 was used to assess risk of bias. One study excluded patients who could not produce sputa. Evidence was not downgraded.
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. This was a borderline judgment. The quality of the evidence was downgraded by one point.
3. QUADAS-2 was used to assess risk of bias. Three studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
4. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which credible limits are assumed. The quality of the evidence was downgraded by one point.

Table 8. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV with CD4 ≤ 200 cells/µL, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 200 cells/µL, grade 2?

Participants: Adults (15 years and older), infected with HIV, with CD4 ≤ 200 cells/µL, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.49 (95% CrI: 0.34 to 0.66)
Pooled specificity	0.90 (95% CrI: 0.78 to 0.95)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	5 studies 605 patients	cross-sectional (cohort type accuracy study)	not serious[1]	not serious	not serious	serious[2]	not serious	5 (3 to 7)	49 (34 to 66)	147 (102 to 198)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								5 (7 to 3)	51 (66 to 34)	153 (198 to 102)	
True negatives (patients without tuberculosis)	5 studies 739 patients	cross-sectional (cohort type accuracy study)	serious[3]	not serious	not serious	serious[4]	not serious	891 (772 to 941)	810 (702 to 855)	630 (546 to 665)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								99 (218 to 49)	90 (198 to 45)	70 (154 to 35)	

1. QUADAS-2 was used to assess risk of bias. One study excluded patients who could not produce sputa. Evidence was not downgraded.
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. This was a borderline judgment. The quality of the evidence was downgraded by one point.
3. QUADAS-2 was used to assess risk of bias. Three studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
4. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 9. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV with CD4 > 100 cells/µL, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 > 100 cells/µL, grade 2?

Participants: Adults (15 years and older), infected with HIV, with CD4 > 100 cells/µL, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.26 (95% CrI: 0.16 to 0.46)
Pooled specificity	0.92 (95% CrI: 0.78 to 0.97)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	5 studies 421 patients	cross-sectional (cohort type accuracy study)	not serious[1]	not serious	not serious	serious[2]	not serious	3 (2 to 5)	26 (16 to 46)	78 (48 to 138)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								7 (8 to 5)	74 (84 to 54)	222 (252 to 162)	
True negatives (patients without tuberculosis)	5 studies 989 patients	cross-sectional (cohort type accuracy study)	serious[3]	not serious	not serious	serious[4]	not serious	911 (772 to 960)	828 (702 to 873)	644 (546 to 679)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								79 (218 to 30)	72 (198 to 27)	56 (154 to 21)	

1. QUADAS-2 was used to assess risk of bias. One study excluded patients who could not produce sputa. Evidence was not downgraded.
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.
3. QUADAS-2 was used to assess risk of bias. Three studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
4. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 10. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV with CD4 ≤ 100 cells/µL, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 100 cells/µL, grade 2?

Participants: Adults (15 years and older), infected with HIV, with CD4 ≤ 100 cells/µL, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.56 (95% CrI: 0.41to 0.70)
Pooled specificity	0.90 (95% CrI: 0.81 to 0.95)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	5 studies 402 patients	cross-sectional (cohort type accuracy study)	not serious[1]	not serious	not serious	serious[2]	not serious	6 (4 to 7)	56 (41 to 70)	168 (123 to 210)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								4 (6 to 3)	44 (59 to 30)	132 (177 to 90)	
True negatives (patients without tuberculosis)	5 studies 457 patients	cross-sectional (cohort type accuracy study)	serious[3]	not serious	not serious	serious[4]	not serious	891 (802 to 941)	810 (729 to 855)	630 (567 to 665)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								99 (188 to 49)	90 (171 to 45)	70 (133 to 35)	

1. QUADAS-2 was used to assess risk of bias. One study excluded patients who could not produce sputa. Evidence was not downgraded.
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.
3. QUADAS-2 was used to assess risk of bias. Three studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
4. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 11. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV with CD4 ≤ 200 cells/µL in inpatient settings, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 200 cells/µL in inpatient settings, grade 2?

Participants: Adults (15 years and older), infected with HIV, with CD4 ≤ 200 cells/µL, presumed to have TB
Setting: Inpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.56 (95% CrI: 0.42 to 0.71)
Pooled specificity	0.88 (95% CrI: 0.70 to 0.95)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	4 studies 508 patients	cross-sectional (cohort type accuracy study)	not serious	not serious	not serious	serious[1]	not serious	6 (4 to 7)	56 (42 to 71)	168 (126 to 213)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								4 (6 to 3)	44 (58 to 29)	132 (174 to 87)	
True negatives (patients without tuberculosis)	4 studies 603 patients	cross-sectional (cohort type accuracy study)	serious[2]	not serious	not serious	serious[1]	not serious	871 (693 to 941)	792 (630 to 855)	616 (490 to 665)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								119 (297 to 49)	108 (270 to 45)	84 (210 to 35)	

1. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.
2. QUADAS-2 was used to assess risk of bias. Two studies were considered to be at high risk of bias for the reference standard. The qua lity of the evidence was downgraded by one point.
3. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 12. GRADE evidence profile: accuracy of LF-LAM to diagnose tuberculosis in adults with HIV with CD4 ≤ 100 cells/µL in inpatient settings, grade 2

PICO Question: Should LF-LAM be used to diagnose tuberculosis in adults with HIV with CD4 ≤ 100 cells/µL in inpatient settings, grade 2?

Participants: Adults (15 years and older), infected with HIV, with CD4 ≤ 100 cells/µL, presumed to have TB
Setting: Inpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.61 (95% CrI: 0.27 to 0.69)
Pooled specificity	0.89 (95% CrI: 0.71 to 0.89)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	4 studies 356 patients	cross-sectional (cohort type accuracy study)	not serious	not serious	not serious	serious[1]	not serious	6 (5 to 8)	61 (48 to 75)	183 (144 to 225)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								4 (5 to 2)	39 (52 to 25)	117 (156 to 75)	
True negatives (patients without tuberculosis)	4 studies 376 patients	cross-sectional (cohort type accuracy study)	serious[2]	not serious	not serious	serious[3]	not serious	881 (742 to 941)	801 (675 to 855)	623 (525 to 665)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								109 (248 to 49)	99 (225 to 45)	77 (175 to 35)	

1. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.
2. QUADAS-2 was used to assess risk of bias. Two studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
3. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 13. GRADE evidence profile: accuracy of LF-LAM vs. sputum microscopy to diagnose tuberculosis in adults with HIV, grade 2

PICO Question: Should LF-LAM vs. sputum microscopy be used to diagnose tuberculosis in adults with HIV, grade 2?

Participants: Adults (15 years and older), infected with HIV, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity LF-LAM	0.37 (95% CrI: 0.32 to 0.42)	**Pooled sensitivity sputum microscopy**	0.47 (95% CrI: 0.35 to 0.59)
Pooled specificity LF-LAM	0.95 (95% CrI: 0.93 to 0.97)	**Pooled specificity sputum microscopy**	0.98 (95% CrI: 0.93 to 1.00)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year						Test accuracy QoE
								pre-test probability of 1%		Pre-test probability of 10%		Pre-test probability of 30%		
			Risk of bias	Indirect-ness	Inconsis-tency	Impre-cision	Publication bias	LF-LAM	sputum microscopy	LF-LAM	Sputum microscopy	LF-LAM	Sputum microscopy	
True positives (patients with tuberculosis)	4 studies 617 patients	cross-sectional (cohort type accuracy study)	very serious[1]	not serious	not serious	not serious	not serious	4 (3 to 4)	5 (3 to 6)	37 (32 to 42)	47 (35 to 59)	111 (96 to 126)	141 (105 to 177)	⊕⊕○○ low
								1 fewer TP in LF-LAM		10 fewer TP in LF-LAM		30 fewer TP in LF-LAM		
False negatives (patients incorrectly classified as not having tuberculosis)								6 (7 to 6)	5 (7 to 4)	63 (68 to 58)	53 (65 to 41)	189 (204 to 174)	159 (195 to 123)	
								1 more FN in LF-LAM		10 more FN in LF-LAM		30 more FN in LF-LAM		
True negatives (patients without tuberculosis)	4 studies 998 patients	cross-sectional (cohort type accuracy study)	serious[2]	not serious	not serious	not serious	not serious	941 (921 to 960)	970 (921 to 990)	855 (837 to 873)	882 (837 to 900)	665 (651 to 679)	686 (651 to 700)	⊕⊕⊕○ moderate
								29 fewer TN in LF-LAM		27 fewer TN in LF-LAM		21 fewer TN in LF-LAM		
False positives (patients incorrectly classified as having tuberculosis)								49 (69 to 30)	20 (69 to 0)	45 (63 to 27)	18 (63 to 0)	35 (49 to 21)	14 (49 to 0)	
								29 more FP in LF-LAM		27 more FP in LF-LAM		21 more FP in LF-LAM		

1. QUADAS-2 was used to assess risk of bias. The pooled analyses were restricted to those patients who could produce sputa; therefore, we considered these studies at high risk of bias for patient selection. The quality of the evidence was downgraded by two points.
2. QUADAS-2 was used to assess risk of bias. Two studies were considered at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.

Table 14. GRADE evidence profile: accuracy of LF-LAM in combination with sputum microscopy to diagnose tuberculosis in adults with HIV, grade 2

PICO Question: Should LF-LAM in combination with sputum microscopy be used to diagnose tuberculosis in adults with HIV, grade 2?

Participants: Adults (15 years and older), infected with HIV, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.62 (95% CrI: 0.50 to 0.73)
Pooled specificity	0.91 (95% CrI: 0.73 to 0.96)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	4 studies 617 patients	cross-sectional (cohort type accuracy study)	very serious[1]	not serious	not serious	not serious[2]	not serious	6 (5 to 7)	62 (50 to 73)	186 (150 to 219)	⊕⊕○○ low
False negatives (patients incorrectly classified as not having tuberculosis)								4 (5 to 3)	38 (50 to 27)	114 (150 to 81)	
True negatives (patients without tuberculosis)	4 studies 998 patients	cross-sectional (cohort type accuracy study)	serious[3]	not serious	not serious	serious[4]	not serious	901 (723 to 950)	819 (657 to 864)	637 (511 to 672)	⊕⊕○○ low
False positives (patients incorrectly classified as having tuberculosis)								89 (267 to 40)	81 (243 to 36)	63 (189 to 28)	

1. QUADAS-2 was used to assess risk of bias. The pooled analyses were restricted to those patients who could produce sputa; therefore, these studies were considered at high risk of bias for patient selection. The quality of the evidence was downgraded by two points.
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. This was a borderline decision. No points were downgraded.
3. QUADAS-2 was used to assess risk of bias. Two studies were considered at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
4. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 15. GRADE evidence profile: accuracy of LF-LAM vs. sputum Xpert MTB/RIF microscopy to diagnose tuberculosis in adults with HIV, grade 2

PICO Question: Should LF-LAM vs. sputum Xpert MTB/RIF be used to diagnose tuberculosis in adults with HIV, grade 2?

Participants: Adults (15 years and older), infected with HIV, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity LF-LAM	0.37 (95% CrI: 0.32 to 0.42)	**Pooled sensitivity sputum microscopy**	0.47 (95% CrI: 0.35 to 0.59)	**Prevalences**	1%	10%	30%
Pooled specificity LF-LAM	0.95 (95% CrI: 0.93 to 0.97)	**Pooled specificity sputum microscopy**	0.98 (95% CrI: 0.93 to 1.00)				

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year						Test accuracy QoE
								Pre-test probability of 1%		Pre-test probability of 10%		Pre-test probability of 30%		
			Risk of bias	Indirect-ness	Inconsis-tency	Impre-cision	Publication bias	LF-LAM	sputum microscopy	LF-LAM	sputum microscopy	LF-LAM	sputum microscopy	
True positives (patients with tuberculosis)	3 studies 238 patients	cross-sectional (cohort type accuracy study)	very serious[1]	not serious	not serious	not serious	not serious	4 (3 to 4)	8 (6 to 9)	37 (30 to 43)	77 (64 to 85)	111 (90 to 129)	231 (192 to 255)	⊕⊕○○ low
								4 fewer TP in LF-LAM		40 fewer TP in LF-LAM		120 fewer TP in LF-LAM		
False negatives (patients incorrectly classified as not having tuberculosis)								6 (7 to 6)	2 (4 to 1)	63 (70 to 57)	23 (36 to 15)	189 (210 to 171)	69 (108 to 45)	
								4 more FN in LF-LAM		40 more FN in LF-LAM		120 more FN in LF-LAM		
True negatives (patients without tuberculosis)	3 studies 411 patients	cross-sectional (cohort type accuracy study)	not serious[2]	not serious	not serious	not serious	not serious	950 (931 to 970)	950 (911 to 980)	864 (846 to 882)	864 (828 to 891)	672 (658 to 686)	672 (644 to 693)	⊕⊕⊕⊕ high
								0 fewer TN in LF-LAM		0 fewer TN in LF-LAM		0 fewer TN in LF-LAM		
False positives (patients incorrectly classified as having tuberculosis)								40 (59 to 20)	40 (79 to 10)	36 (54 to 18)	36 (72 to 9)	28 (42 to 14)	28 (56 to 7)	
								0 fewer FP in LF-LAM		0 fewer FP in LF-LAM		1 fewer FP in LF-LAM		

1. QUADAS-2 was used to assess risk of bias. The pooled analyses were restricted to those patients who could produce sputa; therefore, these studies were considered at high risk of bias for patient selection. The quality of the evidence was downgraded by two points.
2. QUADAS-2 was used to assess risk of bias. One study was considered to be at high risk of bias for the reference standard. No points were downgraded.

Table 16. GRADE evidence profile: accuracy of LF-LAM in combination with sputum Xpert MTB/RIF microscopy to diagnose tuberculosis in adults with HIV, grade 2

PICO Question: Should LF-LAM in combination with sputum Xpert MTB/RIF be used to diagnose tuberculosis in adults with HIV, grade 2?

Participants: Adults (15 years and older), infected with HIV, presumed to have TB
Setting: Inpatient and outpatient

Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.83 (95% CrI: 0.74 to 0.90)
Pooled specificity	0.92 (95% CrI: 0.80 to 0.96)

Prevalences	1%	10%	30%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year			Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	Pre-test probability of 30%	
True positives (patients with tuberculosis)	3 studies 238 patients	cross-sectional (cohort type accuracy study)	very serious[1]	not serious	not serious	not serious	not serious	8 (7 to 9)	83 (74 to 90)	249 (222 to 270)	⊕⊕○○ low
False negatives (patients incorrectly classified as not having tuberculosis)								2 (3 to 1)	17 (26 to 10)	51 (78 to 30)	
True negatives (patients without tuberculosis)	3 studies 411 patients	cross-sectional (cohort type accuracy study)	not serious[2]	not serious	not serious	serious[3]	not serious	911 (792 to 950)	828 (720 to 864)	644 (560 to 672)	⊕⊕⊕○ moderate
False positives (patients incorrectly classified as having tuberculosis)								79 (198 to 40)	72 (180 to 36)	56 (140 to 28)	

1. QUADAS-2 was used to assess risk of bias. The pooled analyses were restricted to those patients who could produce sputa; therefore, these studies were considered at high risk of bias for patient selection. The quality of the evidence was downgraded by two points.
2. QUADAS-2 was used to assess risk of bias. One study was considered at high risk of bias for the reference standard. No points were downgraded.
3. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 17. GRADE evidence profile: accuracy of LF-LAM to screen for tuberculosis in adults with HIV, grade 2

PICO Question: Should LF-LAM be used to screen for tuberculosis in adults with HIV, microbiological reference standard, grade 2?

Participants: Adults (15 years and older), infected with HIV, who may or may not have had signs and symptoms compatible with TB and had not been previously evaluated for TB

Setting: Mostly outpatient
Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.23 (95% CrI: 0.13 to 0.37)
Pooled specificity	0.96 (95% CrI: 0.92 to 0.98)

Prevalences	1%	10%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year		Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	
True positives (patients with tuberculosis)	6 studies 290 patients	cross-sectional (cohort type accuracy study)	serious[1]	not serious[2]	not serious	not serious	not serious	2 (1 to 4)	23 (13 to 37)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								8 (9 to 6)	77 (87 to 63)	
True negatives (patients without tuberculosis)	6 studies 1557 patients	cross-sectional (cohort type accuracy study)	very serious[3]	not serious[2]	not serious	not serious	not serious	950 (911 to 970)	864 (828 to 882)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								40 (79 to 20)	36 (72 to 18)	

1. QUADAS-2 was used to assess risk of bias. Three studies were considered to be at high risk of bias because the studies excluded patients who were unable to produce sputum. The quality of the evidence was downgraded by one point.
2. Studies were conducted in asymptomatic patients and almost exclusively in outpatient settings.
3. QUADAS-2 was used to assess risk of bias. Five studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by two points.
4. The 95% CrI for false positives may lead to different decisions depending on which confidence limits are assumed. This was a borderline judgment. No points were downgraded.

Table 18. GRADE evidence profile: accuracy of LF-LAM to screen for tuberculosis in adults with HIV, composite reference standard, grade 2

PICO Question: Should LF-LAM be used to screen for tuberculosis in adults with HIV, composite reference standard, grade 2?

Participants: Adults (15 years and older), infected with HIV, who may or may not have had signs and symptoms compatible with TB and had not been previously evaluated for TB

Setting: Mostly outpatient
Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture, NAAT, smear, clinical findings

Pooled sensitivity	0.18 (95% CrI: 0.07 to 0.35)
Pooled specificity	0.97 (95% CrI: 0.92 to 0.99)

Prevalences	1%	10%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year		Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	
True positives (patients with tuberculosis)	4 studies 205 patients	cross-sectional (cohort type accuracy study)	serious[1]	not serious[2]	not serious	not serious	not serious	2 (1 to 3)	18 (7 to 35)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								8 (9 to 7)	82 (93 to 65)	
True negatives (patients without tuberculosis)	4 studies 1223 patients	cross-sectional (cohort type accuracy study)	not serious[3]	not serious[2]	not serious	not serious	not serious	960 (911 to 980)	873 (828 to 891)	⊕⊕⊕⊕ high
False positives (patients incorrectly classified as having tuberculosis)								30 (79 to 10)	27 (72 to 9)	

1. QUADAS-2 was used to assess risk of bias. Two studies were considered to be at high risk of bias because these studies excluded patients who were unable to produce sputum. The quality of the evidence was downgraded by one point.
2. Studies were conducted in asymptomatic patients and almost exclusively in outpatient settings. One study exclusively enrolled pregnant women.
3. QUADAS-2 was used to assess risk of bias. All studies were considered to be at low risk of bias for the composite reference standard.
4. The 95% CrI for false positives may lead to different decisions depending on which confidence limits are assumed. This was a borderline judgment. No points were downgraded.

Table 19. GRADE evidence profile: accuracy of LF-LAM to screen for tuberculosis in adults with HIV in inpatient setting, grade 2

PICO Question: Should LF-LAM be used to screen for tuberculosis in adults with HIV in inpatient settings, grade 2?

Participants: Adults (15 years and older) inpatients, who may or may not have had signs and symptoms compatible with TB and had not been previously evaluated for TB

Setting: Inpatient
Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.52 (95% CrI: 0.29 to 0.76)
Pooled specificity	0.94 (95% CrI: 0.74 to 0.98)

Prevalences	1%	10%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year		Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	
True positives (patients with tuberculosis)	2 studies 151 patients	cross-sectional (cohort type accuracy study)	serious[1]	not serious	not serious	serious[2]	not serious	5 (3 to 8)	52 (29 to 76)	⊕⊕◯◯ low
False negatives (patients incorrectly classified as not having tuberculosis)								5 (7 to 2)	48 (71 to 24)	
True negatives (patients without tuberculosis)	2 studies 332 patients	cross-sectional (cohort type accuracy study)	serious[3]	not serious	not serious	serious[4]	not serious	931 (733 to 970)	846 (666 to 882)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								59 (257 to 20)	54 (234 to 18)	

1. QUADAS-2 was used to assess risk of bias. We considered one study to be at high risk of bias because this study excluded patients unable to produce sputum. The quality of the evidence was downgraded by one point.
2. The wide 95% CrI for true positives and false negatives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.
3. QUADAS-2 was used to assess risk of bias. We considered one study to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by one point.
4. The wide 95% CrI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. The quality of the evidence was downgraded by one point.

Table 20. GRADE evidence profile: accuracy of LF-LAM to screen for tuberculosis in adults with HIV in outpatient setting, grade 2

PICO Question: Should LF-LAM be used to screen for tuberculosis in adults with HIV in outpatient settings, grade 2?

Participants: Adults (15 years and older) outpatients, who may or may not have had signs and symptoms compatible with TB and had not been previously evaluated for TB

Setting: Outpatient
Target condition: Active TB (Pulmonary and Extrapulmonary)
Reference test: *M. tuberculosis* culture or NAAT

Pooled sensitivity	0.22 (95% CI: 0.13 to 0.35)
Pooled specificity	0.96 (95% CI: 0.93 to 0.98)

Prevalences	1%	10%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease quality of evidence					Effect per 1000 patients/year		Test accuracy QoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 10%	
True positives (patients with tuberculosis)	2 studies 151 patients	cross-sectional (cohort type accuracy study)	serious[1]	not serious	not serious	serious[2]	not serious	5 (3 to 8)	52 (29 to 76)	⊕⊕⊕◯ moderate
False negatives (patients incorrectly classified as not having tuberculosis)								5 (7 to 2)	48 (71 to 24)	
True negatives (patients without tuberculosis)	2 studies 332 patients	cross-sectional (cohort type accuracy study)	serious[3]	not serious	not serious	serious[4]	not serious	931 (733 to 970)	846 (666 to 882)	⊕⊕◯◯ low
False positives (patients incorrectly classified as having tuberculosis)								59 (257 to 20)	54 (234 to 18)	

1. QUADAS-2 was used to assess risk of bias. Three studies were considered to be at high risk of bias because they excluded patients unable to produce sputum. The quality of the evidence was downgraded by one point.
2. QUADAS-2 was used to assess risk of bias. Five studies were considered to be at high risk of bias for the reference standard. The quality of the evidence was downgraded by two points.
3. The 95% CI for false positives may lead to different decisions depending on which confidence limits are assumed. This was a borderline judgment. No points were downgraded.

10. Annexes

Annex 1. Meeting participants

Guideline Development Group Members

1. Jan Brozek (Chair/GRADE methodologist)
McMaster University
Canada
brozekj@mcmaster.ca

2. Jeremiah Chakaya
Kenya Medical Research Institute
Kenya
Chakaya.jm@gmail.com

3. Gavin Churchyard
The Aurum Institute NPC
South Africa
gchurchyard@auruminstitute.org

4. Daniela Maria Cirillo
HSR San Raffaele Scientific Institute (SRL Milan)
Italy
cirillo.daniela@hsr.it

5. Diane Havlir
University of California, USA
Diane.Havlir@ucsf.edu

6. Nagalineswaran Kumarasamy
YRGCARE Medical Centre, India
kumarasamy@yrgcare.org

7. Angela Mushavi
Ministry of Health, Zimbabwe
mushavia@yahoo.co.uk

8. Ingrid Oxley Oxland
Nelson Mandela Metropolitan University
South Africa
ingridoxley@gmail.com

9. Wendy Stevens
Wits University, South Africa
wendy.stevens@nhls.ac.za

10. Satoshi Mitarai
Japan Anti-Tuberculosis Association
Japan
mitarai@jata.or.jp

11. Thato Mosidi
Chris Hani Baragwanath Hospital
South Africa
thatomosidi@gmail.com

12. Thomas M Shinnick
US Centers for Disease Control and Prevention
USA
tms1@cdc.gov

13. Francis Varaine
Médecins Sans Frontières, France
Francis.VARAINE@paris.msf.org

14. Anna Vassall
London School of Hygiene and Tropical Medicine, UK
Anna.Vassall@lshtm.ac.uk

Observers

1. Claudia Denkinger
FIND, Switzerland
Claudia.Denkinger@finddx.org

Systematic reviewers

15. David Dowdy
Johns Hopkins Bloomberg School of Public Health, USA
ddowdy1@jhmi.edu

16. Colleen Hanrahan
Johns Hopkins University, USA
chanrah1@jhmi.edu

17. Jonathan Peter–remote particpation
University of Cape Town
South Africa
Jonny.Peter@uct.ac.za

18. Maunank Shah
Johns Hopkins Bloomberg School of Public Health, USA
mshah28@jhmi.edu

19. Karen Steingart
Liverpool School of Tropical Medicine, UK
karen.steingart@gmail.com

Annex 2. Members of the External Review Group

1. Heidi Albert
FIND
Cape Town, South Africa
Heidi.Albert@finddx.org

2. Heather Alexander
US Centers for Disease Control and Prevention
United States of America
drz5@cdc.gov

3. Martina Casenghi
Médecins Sans Frontières-Access Campaign
Martina.CASENGHI@geneva.msf.org

4. Chris Coulter
Queensland Mycobacterium Reference Laboratory
Brisbane, Australia
chris.coulter@health.qld.gov.au

5. Francis Drobniewski
Imperial College
United Kingdom
f.drobniewski@imperial.ac.uk

6. Kathleen England
KNCV tuberculosis foundation, Netherlands
kathleen.england@kncvtbc.org

7. Levan Gagnidze
International Organization for Migration,
Bangkok, Thailand
lgagnidze@iom.int

8. Rumina Hassan
SRL Pakistan, Aga Khan University
Karachi, Pakistan
rumina.hasan@aku.edu

9. Nazir Ahmed Ismail
Candidate SRL, National Institute of Communicable Diseases, South Africa
naziri@ncid.ac.za

10. Paul Klatser
Royal Tropical Institute,
Amsterdam, Netherlands
p.klatser@kit.nl

11. Richard Lumb
Adelaide SRL,
Adelaide, Australia
Richard.Lumb2@health.sa.gov.au

12. Beatrice Mutayoba
Ministry of Health and Social Welfare
Tanzania
beatricemutayoba@yahoo.com

13. Alaine Nyaruhirira
Management Sciences for Health,
Pretoria, South Africa
alainenyaruhirira@hotmail.com

14. Sabira Tahseen,
National TB Reference Laboratory
Islamabad, Pakistan
sabira.tahseen@gmail.com

15. Maria Alice Telles
Management Sciences for Health, Sao Paolo, Brazil
atelles.msh@gmail.com

16. Maarten Van Cleeff
Challenge TB, KNCV tuberculosis foundation, Netherlands
maarten.vancleeff@kncvtbc.org

Annex 3 WHO Steering Group

1. Meg Doherty
WHO HQ, HIV department
dohertym@who.int

2. Alberto Matteelli
WHO HQ, Global TB Programme
matteellia@who.int

3. Haileyesus Getahun
WHO HQ, Global TB Programme
getahunh@who.int

4. Fuad Mirzayev
WHO HQ, Global TB Programme
mirzayevf@who.int

5. Christopher Gilpin
WHO HQ, Global TB Programme
gilpinc@who.int

6. Wayne van Gemert
WHO HQ, Global TB Programme
vangemertw@who.int

7. Alexei Korobitsyn
WHO HQ, Global TB Programme
korobitsyna@who.int

8. Karin Weyer
WHO HQ, Global TB Programme
weyerk@who.int

Annex 4. Declarations of Interests

None declared
Jan Brozek (Chair); Jeremiah Chakaya; Gavin Churchyard; Ingrid Oxley Oxland; Nagalineswaran Kumarasamy; Daniela Cirillo.

Declared, insignificant
Satoshi Mitarai: Received research support and paid travel from Eiken, Japan to attend the 5th National LAMP research forum in China in 2014.(USD 3000) Research coordination, data analysis and manuscript writing of the paper on LAMP, 2011. No funding received related to LF-LAM.

Beatrice Mutayoba: Public statements and positions related to her appointment as NTP Director, Tanzania.

Angela Mushavi: National PMTCT and paediatric HIV care and treatment coordinator, Zimbabwe.

Wendy Stevens: Funding received for other TB assay validations (Cepheid, Abbott, Roche, Hain, DNA genotek, Alere) generally in the form of reagents. No funding received related to LF-LAM.

Anna Vassall: Consultancy on modelling the cost-effectiveness of new diagnostics – AIGHD €3000. Results were not presented during the current meeting.

Francis Varaine: Leader of the MSF working group on TB, required to defend positions related to TB diagnostics.

Thato Mosidi: Member of the South African, country co-ordination mechanism for the Global Fund.

Diane Havlir: National Institute of Health (NIH) supports her faculty salary at the University of California San Francisco. Research support received with the provision of medications only. Truvada was supplied for an NIH study but no funds were received.

Thomas Shinnick: An employee of the United States Centres for Disease Control and Prevention (CDC). CDC supports travel and research related to his work on the laboratory services needed for

tuberculosis control; represented CDC's positions on laboratory services needed for tuberculosis diagnosis, treatment, and control. Served on the Data and Safety Monitoring Board (DSMB) organized by Otsuka for the clinical trial of delamanid – no remuneration received.

Karen Steingart: Conducted the systematic review on LF-LAM
Maunank Shah: Conducted the systematic review on LF-LAM
Jonny Peter: Conducted the systematic review on LF-LAM; LF-LAM study investigator and received 3000 test LF-LAM tests Alere USA for the study (estimated commercial value USD 10,000).

David Dowdy: Conducted the review on economic evaluations for the use of LF-LAM
Colleen Hanrahan: Conducted the review on economic evaluations for the use of LF-LAM

Declared, significant (observer status)
Claudia Denkinger: As a FIND employee, has provided consultation services to advise on specifications of other diagnostics manufactured by Alere. FIND supported a study on LAM by Keertan Dheda. Contributed to the systematic review of LAM in preparation for the Guideline Development Meeting.

Annex 5. References to studies for the review of the diagnostic accuracy of LF-LAM

Included studies

Published

1. Andrews B, Muchemwa L, Lakhi S, Bwalya M, Mabula C, Chipii G, Heimburger DC, Bernard GR. Validation of a clinical prediction score and performance of urine lipoarabinomannan test for detecting tuberculosis bacteremia in HIV-positive patients with severe sepsis. In: American Thoracic Society International Conference Abstracts. Am J Respir Crit Care Med.189:2014: A2559.

2. Balcha TT, Winqvist N, Sturegard E, Skogmar S, Reepalu A, Jemal ZH, et al. Detection of lipoarabinomannan in urine for identification of active tuberculosis among HIV-positive adults in Ethiopian health centres. Trop Med Int Health. 2014 Jun;19(6):734-42.

3. Bjerrum S, Kenu E, Lartey M, Newman MJ, Addo KK, Andersen AB, Johansen IS. Diagnostic accuracy of the rapid urine lipoarabinomannan test for pulmonary tuberculosis among HIV-infected adults in Ghana- findings from the DETECT HIV-TB study. In: 45th Union World Conference on Tuberculosis and Lung Health. International Journal Tuberculosis and Lung Disease 2014;18 Suppl 1(11): OAP-261-30.

4. Drain PK, Losina E, Coleman SM, Giddy J, Ross D, Katz JN, et al. Diagnostic accuracy of a point-of-care urine test for tuberculosis screening among newly-diagnosed HIV-infected adults: a prospective, clinic-based study. BMC Infect Dis. 2014;14:110.

5. Drain PK, Losina E, Coleman SM, Giddy J, Ross D, Katz JN, et al. Value of urine lipoarabinomannan grade and second test for optimizing clinic-based screening for HIV-associated pulmonary tuberculosis. J Acquir ImmuneDefic Syndr. 2015;68(3):274-80.

6. Drain PK, Grobler A, Gounder L, Sahid F, Wilson D, Bassett I, Moosa Y. Rapid urine lipoarabinomannan testing after two months of tuberculosis treatment independently predicts mortality in a resource-limited setting. In: 44th Union World Conference on Lung Health of the International Union Against Tuberculosis and Lung Disease. International Journal Tuberculosis and Lung Disease 2013;17 Suppl 2(12):S264.

7. Tlali M, Fielding K, Charalambous S, Karat A, Hoffmann C, Johnson S, Muravha T, Xaba N, Mathenjwa T, Dorman S, Churchyard G, Grant A. Sensitivity of the TB Determine LAM test compared to sputum culture gold standard TB in ambulant HIV positive participants enrolled in the TB fast track study in South Africa. In: 4th South African TB Conference. 2014

8. LaCourse SM, Cranmer L, Matemo D, Richardson B, Kinuthia J, John-Stewart G, Horne DJ.. Active tuberculosis case finding in HIV-infected pregnant women in Kenya reveals poor performance of symptom screening and rapid diagnostic tests.. JAIDS 2015; in press

9. Lawn SD, Kerkhoff AD, Vogt M, Wood R. Diagnostic accuracy of a low-cost, urine antigen, point-of-care screening assay for HIV-associated pulmonary tuberculosis before antiretroviral therapy: a descriptive study. Lancet Infect Dis. 2012 Mar;12(3):201-9.

10. Lawn SD, Kerkhoff A, Burton R, Schutz C, van Wyk G, Vogt M, Pahlana P, Nicol M, Meintjes G. Massive diagnostic yield of HIV-associated tuberculosis using rapid urine assays in South Africa. In: Conference on Retroviruses and Opportunistic Infections. 2014.

11. Nakiyingi L, Moodley VM, Manabe YC, Nicol MP, Holshouser M, Armstrong DT, et al. Diagnostic accuracy of a rapid urine lipoarabinomannan test for tuberculosis in HIV-infected adults. J Acquir Immune Defic Syndr Mar 26.

12. Peter JG, Theron G, van Zyl-Smit R, Haripersad A, Mottay L, Kraus S, et al. Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients. Eur Respir Journal 2012;40 (5):1211-20.

13. Peter J, Theron G, Chanda D, Clowes P, Rachow A, Lesosky M, Hoelscher et al. Test characteristics and potential impact of the urine LAM lateral flow assay in HIV-infected outpatients under investigation for TB and able to self-expectorate for diagnostic testing. BMS Infectious Diseases 2015;15(262)

14. Van Rie A, Jong E, Mkhwanazi M, Sanne I. Diagnosing TB in those hardest to diagnose: urine lipoarabinomannan for suspects of disseminated and extrapulmonary TB. Atlanta, USA: Abstract of the 20th Conference on Retroviruses and Opportunistic Infections (CROI); 2013. Abstract #841

Unpublished

1. Hanifa Y, Chihota V, Ndlovu N, Karstaedt A, Sahid F, Adonis L, Maesela C, Jawad S, Kinkel H, Stevens W, Erasmus L, Nicol M, McCarthy K, Charalambous S, Churchyard G, Fielding K, Grant A. Xpert MTB/Rif for people attending HIV care: an interventional cohort study to guide rational implementation (XPHACTOR). http://tb.lshtm.ac.uk/projects/xphactor-xpert-mtbrif-people-attending-hiv-care-interventional-cohort-study-guide-rational (accessed 9/4/2015).

2. Reid SE, Harris JB, Kaunda K, Chitambi R, Siyambango M, Henostroza G, Kruuner A**.** Diagnostic Accuracy of Urine Lipoarabinomannan (LAM) Point-of-Care Assay, Xpert MTB/RIF (GXP) and Smear Microscopy for HIV-Associated Tuberculosis (TB) in Outpatients Enrolling in HIV Care in Lusaka, Zambia

Studies with focus on children

1. Nicol MP, Allen V, Workman L, Isaacs W, Munro J, Pienaar S, Black F, Adonis L, Zemanay W, Ghebrekristos Y, Zar HJ. Urine lipoarabinomannan testing for diagnosis of pulmonary tuberculosis in children: a prospective study. Lancet Glob Health. 2014 May;2(5):e278-84

2. Click E., Improving TB diagnosis in children with and without HIV in Kenya, New Diagnostics Tools In Children, The TOTO TB Kenya Study, http://www.impaactnetwork.org/DocFiles/2014AnnualMtg/TBSC/5bSongTOTO_Kenya17Jun14.pdf, Contact Eleanor Click, eoc9@cdc.gov,

3. Pavlinac P., Pediatric Urgent Start of Highly Active Antiretroviral Treatment (HAART) (PUSH). Nested in this larger clinical trial assessing timing of anti-retroviral initiation in hospitalized HIV-infected children we are conducting intensive TB screening to determine the agreement among Xpert

MTB/RIF, LAM, and liquid culture in diagnosing TB from stool, urine, or sputum/gastric aspirate, respectively. https://clinicaltrials.gov/ct2/show/NCT02063880, Contact Patricia Pavlinac, ppav@uw.edu

4. Walters E., Innovative strategies to improve the diagnosis of intrathoracic tuberculosis in children, http://www.impaactnetwork.org/DocFiles/2014AnnualMtg/TBSC/5bEWaltersDxTB17Jun14.pdf, contact Elisabetta Walters, ewal@sun.ac.za

5. Hanrahan C, LAM for paediatric tuberculosis at a primary care clinic in Johannesburg, South Africa [personal communication]. In press.

Related reports

1. Manabe Y, Bareng AS, Nakiying, Mbabbazi, Lubega, Shah M, Dorman S. Point-of-Care lateral flow assays for tuberculosis and cryptococcal antigenuria predict death in HIV infected adults in Uganda. Plos One July 2014

2. Lawn SD, Kerkhoff AD, Vogt M, Wood R. HIV-associated tuberculosis: relationship between disease severity and the sensitivity of new sputum-based and urine-based diagnostic assays. BMC Med. 2013; 11:231.

3. Drain PK, Gounder L, Grobler A, Sahid F, Bassett IV, Moosa MY. Urine lipoarabinomannan to monitor antituberculosis therapy response and predict mortality in an HIV-endemic region: a prospective cohort study. BMJ Open. 2015; 5(4):e006833.

Excluded studies (with reasons for exclusion)

1. Achkar JM, Lawn SD, Moosa MY, Wright CA, Kasprowicz VO. Adjunctive tests for diagnosis of tuberculosis: serology, ELISPOT for site-specific lymphocytes, urinary lipoarabinomannan, string test, and fine needle aspiration. J Infect Dis. 2011;204 Suppl 4:S1130-41. Reason for exclusion: Review/editorial/commentary

2. Agha MA, El-Helbawy RH, El-Helbawy NG, El-Sheak NM. Utility of quantitative analysis of urine lipoarabinomannan in the diagnosis of tuberculosis. Egyptian Journal of Chest Diseases and Tuberculosis. 2013;62 (3):401-7. Reason for exclusion: Not LF-LAM

3. Amoudy HA, AlAsmer ABH, Abul AT, Mustafa AS. Evaluation of complex and defined antigens of Mycobacterium tuberculosis in an IgG-specific ELISA for the diagnosis of tuberculosis. Medical Principles and Practice. 1997;6(2):103-9. Reason for exclusion: Not LF-LAM

4. 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;99(12):893-900. Reason for exclusion: Not LF-LAM

5. Boyles TH. The pros and cons of urinary lipoarabinomannan testing. AIDS. 2012;26(17):2263-4; author reply 4-5. Reason for exclusion: Not LF-LAM

6. 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;161(5):1713-9. Reason for exclusion: Not LF-LAM

7. Cho S-N, Lee J-H, Lee H-Y, Won H-J, Chong Y, Chang J, et al. Detection of Mycobacterium tuberculosis antigens in sputum for the diagnosis of pulmonary tuberculosis. Journal of the Korean Society for Microbiology. 1997;32(3):285-91. Reason for exclusion: Not LF-LAM

8. Conesa-Botella A, Loembe MM, Manabe YC, Worodria W, Mazakpwe D, Luzinda K, et al. Urinary lipoarabinomannan as predictor for the tuberculosis immune reconstitution inflammatory syndrome. J Acquir Immune Defic Syndr. 2011;58(5):463-8. Reason for exclusion: Not LF-LAM

9. 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;13(8):989-95. Reason for exclusion: Not LF-LAM

10. Deng S, Yuan T, Xia J, Huang H, Cheng X, Chen M. Clinical utility of a combination of lipoarabinomannan, 38-kDa, and 16-kDa antigens as a diagnosis tool for tuberculosis. Diagn Microbiol Infect Dis. 2011;71(1):46-50. Reason for exclusion: Not LF-LAM

11. 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. Reason for exclusion: Not LF-LAM

12. Dheda K, Van-Zyl Smit RN, Sechi LA, Badri M, Meldau R, Symons G, et al. Clinical diagnostic utility of IP-10 and LAM antigen levels for the diagnosis of tuberculous pleural effusions in a high burden setting. PLoS One. 2009;4(3):e4689. Reason for exclusion: Not LF-LAM

13. Elsawy A, Redwan EM. Urine lipoarabinomannan as initial markers for active pulmonary tuberculosis. Australian Journal of Basic and Applied Sciences. 2012;6(3):751-5. Reason for exclusion: Not LF-LAM

14. Gounder CR, Kufa T, Wada NI, Mngomezulu V, Charalambous S, Hanifa Y, et al. Diagnostic accuracy of a urine lipoarabinomannan enzyme-linked immunosorbent assay for screening ambulatory HIV-infected persons for tuberculosis. J Acquir Immune Defic Syndr. 2011;58(2):219-23. Reason for exclusion: Not LF-LAM

15. 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;45(1):41-52. Reason for exclusion: Not LF-LAM

16. Hanrahan CF, Van Rie A. Urine antigen test for diagnosis of HIV-associated tuberculosis. The Lancet Infectious Diseases. 2012;12 (11):826. Reason for exclusion: Review/editorial/commentary

17. Kashino SS, Pollock N, Napolitano DR, Rodrigues V, Campos A. Identification and characterization of Mycobacterium tuberculosis antigens in urine of patients with active pulmonary tuberculosis: an innovative and alternative approach of antigen discovery of useful microbial molecules. Clinical and Experimental Immunology. 2008;153(1):56-62. Reason for exclusion: Not LF-LAM

18. Kerkhoff AD, Wood R, Vogt M, Lawn SD. Predictive value of anaemia for tuberculosis in HIV-infected patients in sub-Saharan Africa: an indication for routine microbiological investigation using new rapid assays. J Acquir Immune Defic Syndr. 2013. Reason for exclusion: Duplicate study

19. Kerkhoff AD, Wood R, Vogt M, Lawn SD. Prognostic value of a quantitative analysis of lipoarabinomannan in urine from patients with HIV-associated tuberculosis. PLoS One. 2014;9(7):e103285. Reason for exclusion: Not LF-LAM

20. Koulchin VA, Boehme C, Molokova E, Turecheck C, Hoelscher M, Perkins M, et al. Evaluation of the detection of urinary lipoarabinomannan as a tool for diagnosis of mycobacterial infections. Abstracts of the General Meeting of the American Society for Microbiology. 2007;107:174-5. Reason for exclusion: Not LF-LAM

21. Kroidl I, Clowes P, Mwakyelu J, Maboko L, Kiangi A, Rachow A, et al. Reasons for false-positive lipoarabinomannan ELISA results in a Tanzanian population. Scand J Infect Dis. 2013. Reason for exclusion: Not LF-LAM

22. Lawn SD. Point-of-care detection of lipoarabinomannan (LAM) in urine for diagnosis of HIV-associated tuberculosis: a state of the art review. BMC Infect Dis. 2012;12:103. Reason for exclusion: Review/editorial/commentary

23. 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;23(14):1875-80. Reason for exclusion: Not LF-LAM

24. Lawn SD, Kerkhoff AD, Burton R, Meintjes G. Underestimation of the incremental diagnostic yield of HIV associated tuberculosis in studies of the Determine TB-LAM Ag urine assay. AIDS. 2014;28(12):1846-8. Reason for exclusion: Review/editorial/commentary

25. Lawn SD, Kerkhoff AD, Vogt M, Wood R. Clinical significance of lipoarabinomannan detection in urine using a low-cost point-of-care diagnostic assay for HIV-associated tuberculosis. AIDS. 2012;26(13):1635-43. Reason for exclusion: Insufficient data

26. Lawn SD, Kerkhoff AD, Vogt M, Wood R. HIV-associated tuberculosis: relationship between disease severity and the sensitivity of new sputum-based and urine-based diagnostic assays. BMC Med. 2013;11(1):231. Reason for exclusion: Insufficient data

27. Lawn SD, Wood R. Point-of-care urine antigen screening tests for tuberculosis and cryptococcosis: Potential for mortality reduction in antiretroviral treatment programs in Africa. Clinical Infectious Diseases. 2012;54 (5):739-40. Reason for exclusion: Review/editorial/commentary

28. Manabe YC, Nonyane BA, Nakiyingi L, Mbabazi O, Lubega G, Shah M, et al. Point-of-care lateral flow assays for tuberculosis and cryptococcal antigenuria predict death in HIV infected adults in Uganda. PLoS One. 2014;9(7):e101459. Reason for exclusion: Duplicate study

29. Mukundan H, Kumar S, Price DN, Ray SM, Lee YJ, Min S, et al. Rapid detection of Mycobacterium tuberculosis biomarkers in a sandwich immunoassay format using a waveguide-based optical biosensor. Tuberculosis (Edinb). 2012;92(5):407-16. Reason for exclusion: Not LF-LAM

30. 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;13(10):1253-9. Reason for exclusion: Not LF-LAM

31. Patel VB, Bhigjee AI, Paruk HF, Singh R, Meldau R, Connolly C, et al. Utility of a novel lipoarabinomannan assay for the diagnosis of tuberculous meningitis in a resource-poor high-HIV prevalence setting. Cerebrospinal Fluid Res. 2009;6:13. Reason for exclusion: Not LF-LAM

32. Patel VB, Singh R, Connolly C, Kasprowicz V, Zumla A, Ndungu T, et al. Comparison of a clinical prediction rule and a LAM antigen-detection assay for the rapid diagnosis of TBM in a high HIV prevalence setting. PLoS One. 2010;5(12):e15664. Reason for exclusion: Not LF-LAM

33. Peter JG, Cashmore TJ, Meldau R, Theron G, van Zyl-Smit R, Dheda K. Diagnostic accuracy of induced sputum LAM ELISA for tuberculosis diagnosis in sputum-scarce patients. Int J Tuberc Lung Dis. 2012;16(8):1108-12. Reason for exclusion: Duplicate study

34. Peter JG, Haripesad A, Mottay L, Kraus S, Meldau R, Dheda K. The clinical utility of urine lipoarabinomannan and the novel point-of-care lateral flow strip Test (Determine TB) for the diagnosis of tuberculosis in hospitalised patients with HIV-related advanced immunosuppression. American Journal of Respiratory and Critical Care Medicine Conference: American Thoracic Society International Conference, ATS. 2011;183(1 MeetingAbstracts). Reason for exclusion: Duplicate study

35. Peter JG, Theron G, Dheda K. Can point-of-care urine LAM strip testing for tuberculosis add value to clinical decision making in hospitalised HIV-infected persons? PLoS One. 2013;8(2):e54875. Reason for exclusion: Duplicate study

36. Peter JG, Theron G, Muchinga TE, Govender U, Dheda K. The diagnostic accuracy of urine-based Xpert MTB/RIF in HIV-infected hospitalized patients who are smear-negative or sputum scarce. PLoS One. 2012;7(7):e39966. Reason for exclusion: Duplicate study

37. 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. Reason for exclusion: Not LF-LAM

38. Sada E, Aguilar D, Torres M, Herrera T. Detection of lipoarabinomannan as a diagnostic test for tuberculosis. J Clin Microbiol. 1992;30(9):2415-8. Reason for exclusion: Not LF-LAM

39. Savolainen L, Kantele A, Sandboge B, Siren M, Valleala H, Tuompo R, et al. Modification of clearview tuberculosis (TB) enzyme-linked immunosorbent assay for TB patients not infected with HIV. Clin Vaccine Immunol. 2013;20(9):1479-82. Reason for exclusion: Not LF-LAM

40. Schmidt R, Jacak J, Schirwitz C, Stadler V, Michel G, Marme N, et al. Single-molecule detection on a protein-array assay platform for the exposure of a tuberculosis antigen. J Proteome Res. 2011;10(3):1316-22. Reason for exclusion: Not LF-LAM

41. Shah M, Dowdy D, Joloba M, Ssengooba W, Manabe YC, Ellner J, et al. Cost-effectiveness of novel algorithms for rapid diagnosis of tuberculosis in HIV-infected individuals in Uganda. AIDS. 2013;27 (18):2883-92. Reason for exclusion: Duplicate study

42. 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;48(8):2972-4. Reason for exclusion: Not LF-LAM

43. Shah M, Ssengooba W, Armstrong D, Nakiyingi L, Holshouser M, Ellner JJ, et al. Comparative performance of urinary lipoarabinomannan assays and Xpert MTB/RIF in HIV-infected individuals. AIDS. 2014;28(9):1307-14. Reason for exclusion: Duplicate study

44. 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;52(2):145-51. Reason for exclusion: Not LF-LAM

45. Singh L, Grover N. Detection of TB antigen by rapid test kit. Medical Journal Armed Forces India. 2011;67 (2):196-7. Reason for exclusion: Not LF-LAM

46. Sun D, Dorman S, Shah M, Manabe YC, Moodley VM, Nicol MP, et al. Cost utility of lateral-flow urine lipoarabinomannan for tuberculosis diagnosis in HIV-infected African adults. Int J Tuberc Lung Dis.2013;17(4):552-8. Reason for exclusion: Insufficient data

47. Swaminathan S, Rekha VV. Antigen detection as a point-of-care test for TB: the case of lipoarabinomannan. Future Microbiol. 2012;7(5):559-64. Reason for exclusion: Review/editorial/commentary

48. Talbot E, Munseri P, Teixeira P, Matee M, Bakari M, Lahey T, et al. Test characteristics of urinary lipoarabinomannan and predictors of mortality among hospitalized HIV-infected tuberculosis suspects in Tanzania. PLoS One. 2012;7(3):e32876. Reason for exclusion: Review/editorial/commentary

49. Tessema TA, Bjune G, Assefa G, Svenson S, Hamasur B, Bjorvatn B. Clinical and radiological features in relation to urinary excretion of lipoarabinomannan in Ethiopian tuberculosis patients. Scand J Infect Dis. 2002;34(3):167-71. Reason for exclusion: Not LF-LAM

50. Tessema TA, Bjune G, Hamasur B, Svenson S, Syre H, Bjorvatn B. Circulating antibodies to lipoarabinomannan in relation to sputum microscopy, clinical features and urinary anti-lipoarabinomannan detection in pulmonary tuberculosis. Scand J Infect Dis. 2002;34(2):97-103. Reason for exclusion: Not LF-LAM

51. 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. Reason for exclusion: Not LF-LAM

52. Wood R, Racow K, Bekker LG, Middelkoop K, Vogt M, Kreiswirth BN, et al. Lipoarabinomannan in urine during tuberculosis treatment: association with host and pathogen factors and mycobacteriuria. BMC Infect Dis. 2012;12:47. Reason for exclusion: Not LF-LAM

Annex 6. References to studies for the review of economic evaluations of LF-LAM

Included studies

1. Sun D, Dorman S, Shah M, Manabe YC, Moodley VM, Nicol MP, et al. Cost utility of lateral-flow urine lipoarabinomannan for tuberculosis diagnosis in HIV-infected African adults. Int J Tuberc Lung Dis. 2013;17(4):552-8. Epub 2013/03/15.

2. Shah M, Dowdy D, Joloba M, Ssengooba W, Manabe YC, Ellner J, et al. Cost-effectiveness of novel algorithms for rapid diagnosis of tuberculosis in HIV-infected individuals in Uganda. AIDS. 2013;27(18):2883-92. Epub 2014/08/15.

Excluded studies

1. Bjerrum S, Kenu E, Lartey M, Newman M, Addo KK, Bengaard Anderson A, et al. Diagnostic accuracy of rapid urine LAM test for diagnosing active tuberculosis among in HIV infected adults in Ghana. In: 45th Union World Conference on Lung Health. Barcelona, Spain; 2014. Reason for exclusion: Not economic evaluation

2. Deng S, Yuan T, Xia J, Huang H, Cheng X, Chen M. Clinical utility of a combination of lipoarabinomannan, 38-kDa, and 16-kDa antigens as a diagnosis tool for tuberculosis. Diagn Microbiol Infect Dis 2011,71:46-50. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

3. 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:e9848. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

4. Dheda K, Van-Zyl Smit RN, Sechi LA, Badri M, Meldau R, Symons G, et al. Clinical diagnostic utility of IP-10 and LAM antigen levels for the diagnosis of tuberculous pleural effusions in a high burden setting. PLoS One 2009,4:e4689. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

5. Drain PK, Grobler A, Gounder L, Sahid F, Wilson D, Bassett IV, et al. Rapid urine lipoarabinomannan testing after two months of tuberculosis treatment independently predicts mortality in a resource-limited setting. In: 44th Union World Conference on Lung Health. Paris, France; 2013. Reason for exclusion: Not economic evaluation

6. Drain PK, Grobler A, Gounder L, Sahid F, Wilson D, Bassett IV, et al. Point-of-care urine lipoarabinomannan for diagnosis and treatment response of pulmonary tuberculosis in sputum smear-negative suspects. In: 44th Union World conference on Lungh Health. Paris, France; 2013. Reason for exclusion: Not economic evaluation

7. Drain PK, Losina E, Coleman SM, Giddy J, Ross D, Katz JN, et al. Value of urine lipoarabinomannan grade and second test for optimizing clinic-based screening for HIV-associated pulmonary tuberculosis. Journal of Acquired Immune Deficiency Syndromes 2015,68:274-280. Reason for exclusion: Not economic evaluation

8. Drain PK, Losina E, Coleman SM, Giddy J, Ross D, Katz JN, et al. Diagnostic accuracy of a point-of-care urine test for tuberculosis screening among newly-diagnosed hiv-infected adults: a prospective, clinic-based study. BMC Infect Dis 2014,14. Reason for exclusion: Not economic evaluation

9. Drain PK, Losina E, Coleman SM, Ross D, Walensky RP, Freedberg KA, et al. Value of urine LAM score and second LAM test for optimising clinic-based screening for active pulmonary tuberculosis among HIV-infected adults. In: 45th Union World Conference on Lung Health. Barcelona, Spain; 2014. Reason for exclusion: Not economic evaluation

10. Kerkhoff AD, Wood R, Vogt M, Lawn SD. Prognostic value of a quantitative analysis of lipoarabinomannan in urine from patients with HIV-associated tuberculosis. PLoS One 2014,9:e103285. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

11. Kerkhoff AD, Wood R, Vogt M, Lawn SD. Predictive Value of Anemia for Tuberculosis in HIV-Infected Patients in Sub-Saharan Africa: An Indication for Routine Microbiological Investigation Using New Rapid Assays. Jaids-Journal of Acquired Immune Deficiency Syndromes 2014,66:33-40. Reason for exclusion: Not economic evaluation

12. Kordy F, Richardson SE, Stephens D, Lam R, Jamieson F, Kitai I. Utility of Gastric Aspirates for Diagnosing Tuberculosis in Children in a Low Prevalence Area: Predictors of Positive Cultures and Significance of Non-tuberculous Mycobacteria. Pediatr Infect Dis J 2015,34:91-93. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

13. 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 2009,23:1875-1880. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

14. Lawn SD, Kerkhoff AD, Burton R, Meintjes G. Underestimation of the incremental diagnostic yield of HIV-associated tuberculosis in studies of the Determine TB-LAM Ag urine assay. AIDS 2014,28:1846-1848. Reason for exclusion: Not economic evaluation

15. Lawn SD, Kerkhoff AD, Burton R, Schutz C, van Wyk G, Vogt M, et al. Massive Diagnostic Yield of HIV-Associated Tuberculosis Using Rapid Urine Assays in South Africa. In: Conference on Retrovirus and Opportunistic Infection. Boston, Massachusetts; 2014. Reason for exclusion: Not economic evaluation

16. Lawn SD, Kerkhoff AD, Vogt M, Wood R. Clinical significance of lipoarabinomannan detection in urine using a low-cost point-of-care diagnostic assay for HIV-associated tuberculosis. AIDS 2012,26:1635-1643. Reason for exclusion: Not economic evaluation

17. Lawn SD, Kerkhoff AD, Vogt M, Wood R. Diagnostic accuracy of a low-cost, urine antigen, point-of-care screening assay for HIV-associated pulmonary tuberculosis before antiretroviral therapy: a descriptive study. Lancet Infect Dis 2012,12:201-209. Reason for exclusion: Not economic evaluation

18. Lawn SD, Wood R. Tuberculosis in antiretroviral treatment services in resource-limited settings: addressing the challenges of screening and diagnosis. J Infect Dis 2011,204 Suppl 4:S1159-1167. Reason for exclusion: Not economic evaluation

19. Manabe YC, Nonyane BAS, Nakiyingi L, Mbabazi O, Lubega G, Shah M, et al. Point-of-Care Lateral Flow Assays for Tuberculosis and Cryptococcal Antigenuria Predict Death in HIV Infected Adults in Uganda. PLoS One 2014,9. Reason for exclusion: Not economic evaluation

20. Murray CJ, Ortblad KF, Guinovart C, Lim SS, Wolock TM, Roberts DA, et al. Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014,384:1005-1070. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

21. Nicol MP, Allen V, Workman L, Isaacs W, Munro J, Pienaar S, et al. Urine lipoarabinomannan testing for diagnosis of pulmonary tuberculosis in children: a prospective study. Lancet Global Health 2014,2:E278-E284. Reason for exclusion: Not economic evaluation

22. Patel VB, Bhigjee AI, Paruk HF, Singh R, Meldau R, Connolly C, et al. Utility of a novel lipoarabinomannan assay for the diagnosis of tuberculous meningitis in a resource-poor high-HIV prevalence setting. Cerebrospinal Fluid Res 2009,6:13. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

23. Peter JG, Cashmore TJ, Meldau R, Theron G, van Zyl-Smit R, Dheda K. Diagnostic accuracy of induced sputum LAM ELISA for tuberculosis diagnosis in sputum-scarce patients. Int J Tuberc Lung Dis 2012,16:1108-1112. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

24. Peter JG, Haripesad A, Mottay L, Kraus S, Meldau R, Dheda K. The clinical utility of urine lipoarabinomannan and the novel point-of-care lateral flow strip Test (Determine(registered trademark) TB) for the diagnosis of tuberculosis in hospitalised patients with HIV-related advanced immunosuppression. American Journal of Respiratory and Critical Care Medicine 2011,183. Reason for exclusion: Not economic evaluation

25. Peter JG, Theron G, van Zyl-Smit R, Haripersad A, Mottay L, Kraus S, et al. Diagnostic accuracy of a urine lipoarabinomannan strip-test for TB detection in HIV-infected hospitalised patients. Eur Respir J 2012,40:1211-1220. Reason for exclusion: Not economic evaluation

26. Sada E, Aguilar D, Torres M, Herrera T. Detection of lipoarabinomannan as a diagnostic test for tuberculosis. J Clin Microbiol 1992,30:2415-2418. Reason for exclusion: Not economic evaluation

27. Sun D, Dorman S, Shah M, Manabe Y, Dowdy D. Cost-effectiveness of lateral-flow urine LAM for TB diagnosis in HIV-positive South African adults. Journal of the International AIDS Society 2012,15:253-254. Reason for exclusion: Duplicate data

28. Talbot E, Munseri P, Teixeira P, Matee M, Bakari M, Lahey T, et al. Test characteristics of urinary lipoarabinomannan and predictors of mortality among hospitalized HIV-infected tuberculosis suspects in Tanzania. PLoS One 2012,7:e32876. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

29. Walters E, Gie R, Hesseling A. Urinary lipoarabinomannan for the diagnosis of paediatric pulmonary tuberculosis: a pilot study. In: 44th Union World Conference on Lung Health. Paris, France; 2013. Reason for exclusion: Not economic evaluation

30. Yokoyama T, Rikimaru T, Kinoshita T, Kamimura T, Oshita Y, Aizawa H. Clinical utility of lipoarabinomannan antibody in pleural fluid for the diagnosis of tuberculous pleurisy. J Infect Chemother 2005,11:81-83. Reason for exclusion: Not Alere Determine™ TB LF-LAM test

Annex 7. Summaries of unpublished studies included in the review

Study 1: Evaluation of the urine Lipoarabinomannan (LAM) test for tuberculosis screening amongst people taking antiretroviral therapy (ART) in South Africa

Yasmeen Hanifa, Violet Chihota, Nontobeko Ndlovu, Alan Karstaedt, Faieza Sahid, Lungiswa Adonis, Crawford Maesela, Sena Jawad, Hans Kinkel, Wendy Stevens, Linda Erasmus, Mark Nicol, Kerrigan McCarthy, Salome Charalambous, Gavin Churchyard, Katherine Fielding, Alison Grant

Background: A point-of-care formulation urine LAM test is available, but its place in TB diagnosis is undefined. As part of the XPHACTOR study, we evaluated Determine TB-LAM Ag to screen adults attending for ART.

Methods: In a systematic sample of adults on ART, Xpert was requested if high priority for TB according to XPHACTOR algorithm (any of: cough, BMI < 18.5, CD4 < 100, weight loss ≥ 10%). Urine was stored at enrolment if CD4 < 200, for LAM at 3 months, with positive defined as +/- or greater. All were reviewed monthly, with reinvestigation if indicated, to 3 months when sputum and blood were taken for TB culture. We defined TB cases as: Xpert+ or culture+ for *M. tuberculosis* at any point.

Results: Amongst 122 participants, (57% female, median age 40 yrs, median CD4 120 cells/mm^3, median duration on ART 26 months), 8/122 (6.6%) participants had TB (4 Xpert-positive, 4 culture-positive).

4/122 (3.3%) were LAM positive (3+/-; 1 1+). Sensitivity and specificity of Determine LAM were 0/8 (0%) and 110/114 (96.5%, 95% CI 91.3%, 99.0%) respectively.

Conclusions: The sensitivity of urine LAM is too low to be useful as a component of TB screening among people on ART with CD4<200.

Study 2: Diagnostic Accuracy of Urine Lipoarabinomannan (LAM) Point-of-Care Assay, Xpert MTB/RIF (GXP) and Smear Microscopy for HIV-Associated Tuberculosis (TB) in Outpatients Enrolling in HIV Care in Lusaka, Zambia

Reid SE, Harris JB, Kaunda K, Chitambi R, Siyambango M, Henostroza G, Kruuner A

Background: Diagnosis of TB in severely immunocompromised HIV-infected patients is challenging due to atypical presentation. Accurate and simple point-of-care tests are urgently needed in this population. We compared accuracy of three tests (LAM, Xpert MTB/Rif (GXP), smear), both alone and in combination, for the diagnosis of TB in HIV clinic enrollees.

Design/Methods: Between July 2011 and April 2012 we enrolled 399 ART-naïve, adult outpatients presenting for HIV care in a primary care setting in Lusaka. All patients were screened for TB, regardless of symptoms, via 3 sputum specimens examined by LED fluorescence and light microscopy. Two sputa and one urine sample were cultured on liquid and solid media; one blood sample was cultured with liquid media. For patients who consented to specimen archiving, one sputa for GXP testing and one urine sample for LAM (Determine TB-LAM Ag) testing were frozen and batched for retrospective testing. Positive LAM results included a cutoff of grade 2 and above. Sensitivity, specificity and associated exact binomial confidence intervals were calculated using culture-confirmed TB as the reference standard.

Results: 249 patients who consented to specimen archiving and had smear, valid culture, GXP (single test) and LAM results were included in the analysis. 37 patients (14.9%) had culture confirmed TB (median CD4 count 144 cells/μL; IQR 95-291). Sensitivities of diagnostic tools are shown in the table. Both LAM and GXP had higher sensitivity in patients with CD4 <200 cells/μL. Combined use of LAM and sputum smear increased sensitivity over either test alone for people with CD4<100, but was still substantially lower than a single GXP. LAM combined with GXP did not improve sensitivity compared to GXP alone. The specificity of all tests was ≥ 98%.

THE
END TB
STRATEGY

World Health Organization

For further information, contact:
Global TB Programme
World Health Organization
Avenue Appia 20, CH-1211
Geneva-27, Switzerland
Information Resource Centre HTM/GTB:
Email: tbdocs@who.int
Website: www.who.int/tb

9 789241 509633