JCI The Journal of Clinical Investigation

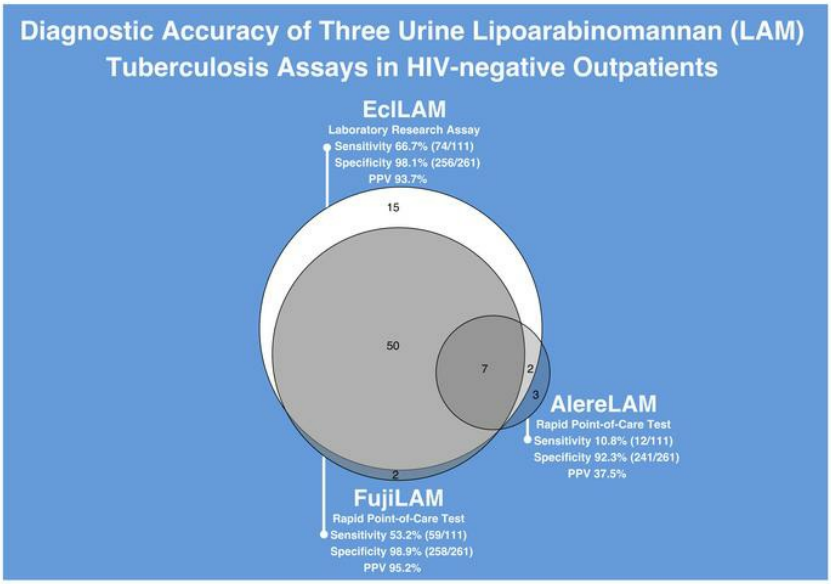
# Diagnostic accuracy of three urine lipoarabinomannan tuberculosis assays in HIV-negative outpatients

Tobias Broger, … , Samuel G Schumacher, Claudia M. Denkinger

*J Clin Invest.* 2020. https://doi.org/10.1172/JCI140461.

Clinical Medicine In-Press Preview Infectious disease

## Graphical abstract



**Title: Diagnostic accuracy of three urine lipoarabinomannan tuberculosis assays in HIV-negative outpatients**

**Authors:** Tobias Broger[1,2#°], Mark P. Nicol.[3,4#], George B. Sigal[5], Eduardo Gotuzzo[6], Alexandra J. Zimmer[1], Shireen Surtie[4], Tatiana Caceres-Nakiche[6], Anna Mantsoki[1], Elena Ivanova Reipold[1], Rita Székely[1], Michael Tsionsky[5], Judith van Heerden[4], Tatiana Plisova[5], Kinuyo Chikamatsu[7], Todd L. Lowary[8,9], Abraham Pinter[10], Satoshi Mitarai[7], Emmanuel Moreau[1&*], Samuel G. Schumacher[1&], Claudia M. Denkinger[1,2&]

**Affiliations:**

[1]FIND, Geneva, Switzerland.

[2]Division of Tropical Medicine, Center for Infectious Diseases, Heidelberg University Hospital, Heidelberg, Germany

[3]Division of Infection and Immunity, School of Biomedical Sciences, University of Western Australia

[4]Division of Medical Microbiology, University of Cape Town, South Africa

[5]Meso Scale Diagnostics, LLC, Rockville, Maryland, USA

[6]Cayetano Heredia University, Lima, Peru

[7]Department of Mycobacterium Reference and Research, Research Institute of Tuberculosis (RIT), Japan Anti-Tuberculosis Association (JATA), Tokyo, Japan

[8]Department of Chemistry, University of Alberta, Edmonton, AB, Canada

[9]Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

[10]Public Health Research Institute Center, New Jersey Medical School, Rutgers University, Newark, New Jersey, USA

[#] Contributed equally, [&] Contributed equally

*To whom correspondence should be addressed: FIND, Emmanuel Moreau, Chemin des Mines 9, 1202 Geneva, Switzerland, Phone +41227100590, e-mail: emmanuel.moreau@finddx.org

°Alternative corresponding author: Tobias Broger, Mettlenweg 4, 9050 Appenzell, Switzerland, Phone +41788432112, e-mail: tobias.broger@gmail.com



Role of funding source: The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Abstract:**

**Background:** Inadequate tuberculosis (TB) diagnostics are a major hurdle in the reduction of disease burden and accurate point-of-care test (POCT) are urgently needed. We assessed the diagnostic accuracy of Fujifilm SILVAMP TB LAM (FujiLAM) for TB diagnosis in HIV-negative outpatients compared to Alere Determine TB LAM Ag (AlereLAM) and a laboratory-based ultrasensitive electrochemiluminescence LAM research assay (EclLAM).

**Methods:** In this multicentre diagnostic test accuracy study, we recruited HIV-negative adults with symptoms suggestive of pulmonary TB presenting to outpatient healthcare centres in Peru and South Africa. Urine samples were tested using FujiLAM, AlereLAM and EclLAM and the diagnostic accuracy was assessed against microbiological (MRS) and composite reference standards.

**Results:** 372 HIV-negative participants were included and the prevalence of microbiologically confirmed TB was 30%. Compared to the MRS, the sensitivities of AlereLAM, FujiLAM and EclLAM were 10.8% (95% CI 6.3to18.0), 53.2% (43.9to62.1), and 66.7% (57.5to74.7) respectively. The specificities of AlereLAM, FujiLAM and EclLAM were 92.3% (88.5to95.0), 98.9% (96.7to99.6), and 98.1% (95.6to99.2) respectively. Positive Likelihood Ratio of AlereLAM, FujiLAM and EclLAM were 1.4, 46.2, and 34.8 and positive predictive values 37.5%, 95.2%, and 93.7% respectively.

**Conclusion:** Compared to AlereLAM, FujiLAM detected five times more TB patients in HIV-negative participants, has a high positive predictive value and has the potential to improve rapid diagnosis of TB at the point-of-care. EclLAM demonstrated that additional sensitivity gains are possible, which highlights LAMs potential as a biomarker. Additional research is required to assess FujiLAMs performance in prospective cohorts, its cost-effectiveness, and its impact in real-world clinical settings.

**[Main Text:]**

## Introduction

Tuberculosis (TB) is the leading single infectious cause of death worldwide with over 1.5 million deaths in 2018.(1) The high rate of unreported TB (estimated at 3.0 million cases) indicates that inadequate diagnostics are a major hurdle in the reduction of disease burden.(1) To address this gap, the World Health Organization (WHO) put forth a set of target product profiles (TPPs)(2, 3) to encourage the development of point-of-care tools to enhance TB case detection. One such TPP is a non-sputum biomarker test for the purpose of initiating TB treatment during the same clinical encounter.(2) An interesting biomarker for this application is the lipoarabinomannan (LAM) antigen found in mycobacterial cell walls.(4, 5)

The Alere Determine TB LAM antigen assay (AlereLAM; Abbott, Chicago, IL, USA) is a TB point-of-care test (POCT) that detects LAM in urine using a simple disposable lateral flow assay. Currently, AlereLAM is the only instrument-free POCT recommended by the WHO for TB, however, its recommended use is limited to assisting in the diagnosis of active TB in people living with HIV (PLHIV) in advanced stages due its limited sensitivity. (6–8) Despite the limited sensitivity, AlereLAM-guided initiation of anti-TB treatment reduced mortality in immunocompromised, hospitalized PLHIV.(9, 10) AlereLAM is not recommended for diagnosis of TB in PLHIV with CD4 greater than 200 cells/µL due to a suboptimal sensitivity of 16% in this population.(6, 7) Performance in HIV-negative patients is very poor, with reported estimated sensitivities ranging from 4 to 31%. (11–15)

Fujifilm (Tokyo, Japan) recently developed a next generation POCT, the Fujifilm SILVAMP TB LAM test (FujiLAM). To improve sensitivity while maintaining high specificity, FujiLAM uses a pair of high-affinity monoclonal antibodies selected to detect LAM presenting the *Mycobacterium tuberculosis (Mtb)*-specific 5-Methylthio-D-xylofuranose epitope (MTX-LAM), and employs a silver-amplification step.(16–18) A recent meta-analysis of 1,595 HIV-positive inpatients and outpatients confirmed FujiLAM's superiority, demonstrating a sensitivity of 71%, twice that of AlereLAM.(19) Further, FujiLAM showed good sensitivity

for the detection of extrapulmonary TB (EPTB) ranging from 47 to 94% across different forms of ETB(20) and could have rapidly diagnosed TB in up to 89% of HIV-positive inpatients who died within 12 weeks.(21)

A non-sputum-based biomarker test would also benefit HIV-negative patients, particularly those with extrapulmonary TB or those unable to produce sputum. This study to assesses FujiLAM's performance in HIV-negative adults with presumptive pulmonary TB. To better understand the relationship between analytical detection limits and clinical sensitivity, the results from FujiLAM are compared to the results from a research assay (EclLAM) employing the same antibodies, but using a more sensitive laboratory immunoassay platform employing electrochemiluminescence (ECL).(14)

## Results

Between February 9, and October 4, 2017, 603 potentially eligible participants were screened. A total of 408 HIV-negative participants met inclusion criteria and 372 were included in the analyses (Figure 1). Of these, 30% (111/372) were classified as definite TB, 3% (10/372) as possible TB, and 67% (251/372) as not TB (Table 1). Prevalence of definite TB was higher in Peru (43%) compared to South Africa (17%). Most participants were young adults (median age 32 years) and 14% had a history of prior TB disease. In participants with definite TB, 68% (76/111) had at least one positive fluorescence sputum smear microscopy (SSM) result. Peruvian participants with TB had shorter mycobacterial growth indicator tube (MGIT) liquid culture time to detection and a larger proportion of patients with positive SSM and Xpert result compared to those from South Africa, suggesting higher mycobacterial load in sputum (Table 1). None of the definite TB patients had a positive blood culture and only 4% (4/111) had a positive urine Xpert result with all of the latter also having a positive sputum Xpert and culture result.

Index test failure rate number was very low (1 repeat FujiLAM and no repeats for AlereLAM). Results of the diagnostic accuracy of urine LAM-based assays (AlereLAM, FujiLAM and EclLAM), sputum-based assays (Xpert and SSM), and combinations of these assays are shown in Figure 2 and in the appendix (p7–9). Overall, compared to the microbiological reference standard (MRS), the sensitivities of AlereLAM, FujiLAM and EclLAM were 10.8% (12/111, [95%Confidence Interval 6.3–18.0]), 53.2% (59/111 [43.9–62.2]), and 66.7% (74/111, [57.5–74.7]), respectively (Figure 1A). The sensitivity of urine FujiLAM and sputum Xpert in combination was 82.0% (91/111, [73.8–88.0]) and the sensitivity of urine FujiLAM in combination with a single SSM was 70.3% (78/111, [61.2–78.0]) and higher than Xpert alone (Figure 2B-C). Using the composite reference standard (CRS), the sensitivities of all assays were not substantially changed (FujiLAM 48.8%, EclLAM 62.0%, AlereLAM 12.4%) (appendix p8).

All tests, except AlereLAM reached specificities of 98% or higher in the MRS-based analysis (AlereLAM 92.3% (241/261, [88.5–95.0]), FujiLAM 98.9% (258/261, [96.7–99.6]), EclLAM 98.1% (256/261, [95.6–

99.2]). When comparing results from the CRS-based analysis to the MRS-based analysis, specificity remained largely unchanged for all assays (AlereLAM 93.2 (234/251), FujiLAM 98.8% (248/251), EclLAM 98.4% (247/251), appendix p8). Against the MRS at study prevalence of 30%, the positive predictive values (PPV) of AlereLAM, FujiLAM and EclLAM were 37.5%, 95.2%, and 93.7% respectively. When assuming a lower pre-test probability of 20% the PPV of AlereLAM, FujiLAM and EclLAM were 26.1%, 92.0%, and 89.7% respectively (Table 2). Positive Likelihood Ratios (LR+) were 46.2 for FujiLAM and 1.4 for AlereLAM. When assuming 20% pre-test probability the NPV of AlereLAM, FujiLAM and EclLAM were, 80.5%, 89.4%, and 92.2% respectively (Table 2). Negative Likelihood Rations (LR−) were 0.47 for FujiLAM and 0.97 for AlereLAM. Fagan nomograms illustrating pretest and posttest probabilities are available in the appendix (p9).

Figure 3 shows the Receiver Operating Characteristic (ROC) curve of the quantitative EclLAM assay and highlights the point estimates of AlereLAM, FujiLAM and EclLAM (at a cut-off of 5.2 pg/ml) in comparison to the TPP performance target. Using the conversion scale based on the EclLAM calibration curve (appendix p6) in Figure 3, we estimate that the LAM threshold of FujiLAM is 10–20 pg/mL and at least 10 times below the threshold of AlereLAM. Data suggests that a threshold around 5 pg/mL or below is required to meet the TPP sensitivity target.

Subgroup analysis revealed that FujiLAM sensitivity was higher in Peru (64.6%) compared to South Africa (25.0%) and this trend was also observed for AlereLAM and EclLAM (Figure 4A and appendix p10-11). Subgroup analyses per MGIT TTD (Figure 4B), SSM status (Figure 4C) and semiquantitative GeneXpert MTB/RIF (Xpert) result (Figure 4D) indicate that FujiLAM sensitivity increases as mycobacterial load in sputum increases and again this trend was confirmed with AlereLAM and EclLAM (appendix p10-11).

Importantly FujiLAM fails to detect 31.6% (24/76) SSM positive patients when using three SSM results as the basis, Figure 4C) or 27.9% (19/68) when using one SSM result as the basis (Figure 2D). On the other hand, FujiLAM detects 23.3% (10/43) of single SSM negative patients with definite TB (Figure 2D). This

increases the sensitivity from 61.3% for a single SSM to 70.3% when FujiLAM and a single SSM are used in combination (Figure 2C and Figure 2D). Even the more sensitive EclLAM fails to detect 16.2% (11/68) single SSM positive patients but would have detected 39.5% (17/43) single SSM negative patients with TB (Figure 2D).

FujiLAM also fails to detect 39% (32/82) sputum Xpert positive patients when using one Xpert as the basis but at the same time it detects 31.0% (9/29) of Xpert negative patients with definite TB (Figure 2D).

**Discussion**

In this multi-centre cohort study of 372 HIV negative outpatients with respiratory symptoms suggestive of pulmonary TB from high-burden TB settings in Peru and South Africa, the FujiLAM point-of-care test was 98.9% specific and identified 53.2% of positive TB cases, representing a five-fold increase in sensitivity among HIV-negative patients compared to AlereLAM. FujiLAM was designed as a rule-in TB diagnostic test to allow rapid treatment initiation and reached a PPV of 95.2%. Together with its high sensitivity for TB diagnosis in PLHIV(19) (sensitivity of 70.7% across CD4 strata) the FujiLAM might have considerable impact on the TB epidemic when scaled-up widely for use in near-patient settings. This is supported by a recent impact modelling analysis. The analysis focusing on LAM-based assays concluded that, relative to the status quo, using a urine-based LAM assay (with 70% sensitivity in PLHIV and 30% in HIV-negatives) in all people presenting to care with TB symptoms would avert 30% of TB deaths and 18% of incident TB cases between 2020 and 2035 in South Africa.(22) The FujiLAM does meet these targets in the study analysed here. While the study might overestimate sensitivity because of the high burden of disease at the study sites, it might at the same time underestimate sensitivity as the study did not consider patients with extra-pulmonary TB or patients that have a hard time producing a sputum (e.g.

children). FujiLAM's NPV is 83.2% and a negative FujiLAM result alone should not be used to rule-out TB and additional microbiological testing is required.

When considering LAM assays for real-world clinical use it is important to evaluate the diagnostic yield, PPV and NPV of algorithms that combine LAM assays and sputum-based assays such as Xpert or SSM.(23) In this study the combination of FujiLAM and Xpert reached 82.0% sensitivity at 98.9% specificity (Figure 2), a PPV of 96.8% and NPV of 92.8% (Table 2). Importantly the combination of FujiLAM and SSM, which is still widely used in clinical practice if Xpert or other molecular tests are not available, reached a similarly high PPV of 96.3% and NPV of 88.7%. The use of these combinations has the potential to rapidly inform TB treatment within a day or less in decentralized settings and treatment in FujiLAM positive patients can immediately be started due to the tests high PPV. The characteristics of FujiLAM and Xpert are complementary: FujiLAM cannot detect drug resistance but Xpert can; Xpert is instrument-based but FujiLAM is a fully disposable POCT; and Xpert uses sputum that is often hard to obtain while FujiLAM uses urine. These findings, as well as the outcomes from the modelling studies(22, 24) suggest that there may be value in integrating LAM-based assays such as FujiLAM into diagnostic algorithms in general populations. A recent assessment further concludes that the Xpert/FujiLAM combination can be cost-effective.(25)

Also an algorithm that starts with X-ray in combination with symptom-based screening to rule-out TB and increase pre-test probability followed by FujiLAM-based diagnosis warrants further investigation. In sum, future studies should carefully assess FujiLAM's added value in real-world scenarios in combination with different tests available at different levels of care and report PPV's and NPV's of such algorithms.

When comparing the two study sites, the performance of all LAM-tests was lower in South Africa compared to Peru. FujiLAMs PPV in South Africa was 72.7% compared with 100% in Peru which was partially due to the lower prevalence in South Africa. Assuming a similar pre-test probability (prevalence) like in Peru, which could be achieved with optimised TB screening, for example with X-ray,

the PPV of FujiLAM would increase to >90%. This is sufficiently high to initiate treatment and substantially higher than clinical diagnosis that is often used in today's clinical practice to initiate empiric treatment. Various indicators suggest that late medical consultation resulting in more advanced disease in Peru or patient selection bias are possible reasons that explain the large differences between sites. This is further supported by a relatively high TB prevalence, high SSM positivity rate and more patients with higher mycobacterial loads in sputum as indicated by 62% of patients with short MGIT time to detection (TTD) in Peru compared to patients in South Africa (Table 1). Subgroup analyses showed that LAM positivity is associated with surrogate markers of body mycobacterial load, such as shorter MGIT TTD, SSM positivity and Xpert semiquantitative result (Figure 4). This finding is in line with an earlier study showing that urine LAM likely reflects total mycobacterial body burden.(26) On the other hand, FujiLAM was negative in a subset of patients with smear-positive disease, suggesting that mycobacterial burden in the sputum is likely not fully reflective of total mycobacterial body burden. Another factor that could impact diagnostic accuracy as a function of geography are structural differences of LAM in different TB strains but there is no scientific evidence of such differences, and further research is needed.

A high specificity (≥98%) of a POCT is necessary to avoid overtreatment. Specificity of FujiLAM in this study with HIV-negatives was 98.9% and higher than AlereLAM's specificity at 92.3% (Figure 2). Earlier studies reported a lower specificity for FujiLAM(16, 19, 27) and the result from this study underlines the importance of a very comprehensive reference standard for a proper specificity assessment of urine biomarker tests.(28)

Our study further shows the potential of LAM as a TB diagnostic biomarker. Using pre-concentration of urine samples and the ultrasensitive EclLAM assay which exploits high-affinity monoclonal antibodies directed towards *Mycobacterium tuberculosis (Mtb)*-specific lipoarabinomannan epitopes(14), we demonstrated that sensitivity increments compared to FujiLAM are feasible. However,

198 this currently requires specialist laboratory equipment allowing electrochemiluminescent-based
199 detection and sophisticated assay protocols. The EclLAM assay reached 66.7% sensitivity at 98.1%
200 specificity showing that a threshold around 5 pg/ml LAM or below is required to meet the TPP sensitivity
201 target. Other recent research studies(26, 29–31) indicated that lower detection limits will translate into
202 higher diagnostic sensitivity. We also showed this in our earlier small case-control study that uses an
203 earlier version of the EclLAM and reached 80% sensitivity in HIV-negative SSM-positive patients at a
204 threshold of 11 pg/ml.(14) In this study, despite the lower threshold due to urine pre-concentration, the
205 sensitivity of EclLAM is lower which is likely a result of the case-control design of the earlier study while
206 this study used a rigorous cohort design with a low risk of bias.

207 It is important to mention that our threshold is an estimate based on the EclLAM assay and non-
208 standardized LAM calibration material and might not be generalizable to other LAM assays with
209 different antibodies, detection technologies or LAM calibration material. Establishing biological
210 reference materials, as has been done by the WHO for other diseases(32), is an urgent priority to
211 support the development, validation and comparison of current and future LAM assays.

212 EclLAM is a research assay employing laboratory equipment and is not designed for use at the POC. In
213 addition, pre-concentration was required to increase sensitivity. Therefore, key challenges in the
214 development of next generation LAM POCT's are to reach a high analytical sensitivity with thresholds in
215 the low pg/ml range while keeping the test simple, affordable and highly specific. Today, the most
216 sensitive POC lateral flow immunoassays detect antigens like the Malaria histidine-rich protein II or LAM
217 in case of FujiLAM, in the low pg/ml range(16, 33) and sample concentration, signal amplification and/or
218 reagent optimisation will likely be needed for POCT's to reach sensitivities like the EclLAM.

219 Taken together, these results suggest that LAM is present in the urine of most HIV-negative patients and
220 that improved assay methods and reagents for LAM detection will lead to increased diagnostic accuracy.
221 The results also suggest that as the detection limits for high sensitivity laboratory-based tests for LAM

continue to improve, centralized urine or blood-based (34) TB antigen detection could also provide a high-throughput complement to nucleic acid tests for TB.

The strengths of this study were the consecutive enrolment of a cohort of HIV-negative patients from two epidemiologically diverse TB endemic settings in Africa and South America, the comparison of three independent LAM assays, the rigorous study design and the comprehensive reference standard. A limitation of the study is that patients unable to provide sputum and patients in whom the disease was thought to be only extrapulmonary and who might benefit from non-sputum based testing(20) were excluded which could have decreased the sensitivity of the LAM assays. Also, the SSM proportion, Xpert and MGIT TTD results suggest more advanced disease in the Peruvian cohort but relatively low burden in South Africa, which could have artificially influenced the sensitivity of the assays. FujiLAM was designed as a POCT and can be used with fresh, unprocessed urine. The use of frozen urine samples for LAM testing in this study could have lowered LAM concentrations as a recent study showed that the use of fresh samples leads to minor sensitivity increases in FujiLAM.(35) Centrifugation of urine is not necessary prior to FujiLAM testing but was a standard procedure in this study.

In conclusion, FujiLAM hast the potential to improve rapid diagnosis of tuberculosis at the point-of-care among all people with presumptive TB presenting to outpatient healthcare centres and could have a high impact on patient outcomes if implemented as a rule-in test in combination with rapid treatment initiation. Further prospective studies are needed to confirm these findings and assess the effect on patient impact to inform policy. Furthermore, the findings highlight the clinical potential of LAM-based diagnosis and research towards an even more sensitive generation of LAM tests should be prioritized.

## Methods

### Study design and participants

In this multi-centre diagnostic accuracy study, we consecutively enrolled adults aged 18 or older with symptoms of pulmonary TB (at least two weeks of persistent cough and at least one additional finding such as haemoptysis, weight loss, fever, night sweats, malaise, contact with an active case, chest pain or loss of appetite) able to produce sputum. In South Africa, outpatients were enrolled at the Town Two and Nolungile primary healthcare facilities in the Khayelitsha township between February 9 and August 31, 2017. In Peru, outpatients were enrolled in 28 primary health care DOTS (Directly Observed Treatment, Short-course) treatment centres in high tuberculosis prevalence areas in the suburbs of Lima and referred to the Universidad Peruana Cayetano Heredia between March 22 and October 4, 2017. Patients in whom the disease was thought to be only extrapulmonary or who received anti-TB treatment in the 60 days prior to enrolment were excluded (appendix p2).

### Procedures

Three sputum (two on day one and a third sputum on day two), one blood, and two urine specimens were collected at enrolment and one sputum specimen at two months follow-up for reference standard testing. Details on the specimen collection and testing flow are provided in the appendix (p3). Reference standard testing was performed in the reference laboratories of the University of Cape Town and the Universidad Peruana Cayetano Heredia on all sputum specimens and included Xpert (Cepheid, Sunnyvale, USA, Xpert testing predated rollout of Xpert Ultra), smear fluorescence microscopy after Auramine O staining, MGIT liquid culture (Becton Dickinson, Franklin Lakes, USA) and solid culture on Löwenstein-Jensen (LJ) medium. Blood cultures from all participants were done in BACTEC™ Myco/F Lytic culture vials (Becton Dickinson, Franklin Lakes, USA). On average 9.4 valid liquid or solid sputum cultures and 2.7 sputum Xpert results were available per patient, 74% of patients had a valid blood culture and 97% had a valid urine Xpert results. The presence of *Mtb* complex in solid and liquid culture

was confirmed with MPT64 antigen (Tauns, Shizuoka, Japan) detection and/or the MTBDRplus line probe assays (Hain Lifesciences, Nehren, Germany). WHO prequalified IVD's were used for HIV testing (rapid diagnostic tests) and CD4 cell counting (flow cytometry). Urine was immediately put on ice after collection and processed within 4 hours. Urine was centrifugated (2000 g at 4°C for 10 minutes), aliquoted on the day of collection and stored at -80°C until batch testing of the liquid fraction with LAM assays. For urinary Xpert testing, 30-40 ml urine was centrifuged (3000 g at 4°C for 15 minutes) and following removal of supernatant the pellet re-suspended in 1 ml PBS and tested using Xpert on the day of collection. Clinical information, index test results and comparator test results were not available to the assessors of the reference standard.

Upon completion of the enrolment, frozen urine aliquots of the complete cohort were shipped to the Research Institute of Tuberculosis of the Japan Anti-Tuberculosis Association (RIT-JATA, Tokyo, Japan) for AlereLAM and FujiLAM testing between January 29 and February 14, 2019 and to Meso Scale Diagnostics LLC (MSD, Rockville, USA) for EclLAM testing between September 18 and 28, 2018.

For AlereLAM and FujiLAM testing at RIT-JATA, frozen urine aliquots were thawed to ambient temperature and mixed by inversion. Samples that were not immediately used for testing were stored at 4 °C for a maximum of 4 h. Testing with FujiLAM was performed according to the manufacturers instructions using urine from the same aliquot as that used for AlereLAM. The five-step FujiLAM test procedure is illustrated in an online video(36) and takes 50–60 minutes from start to end result. The FujiLAM assay does not use a reference scale card and any visible test-line is considered positive. The AlereLAM test was used according to the tests package insert and the four-grade Reference Scale Card, with the Grade 1 cut-off point as the positivity threshold. FujiLAM and AlereLAM were independently read by two readers, each blinded to all other clinical, demographic and test data associated with the samples. After the initial test interpretation, the two readers compared results and, in the event of

discordance, established a final consensus result by mutual agreement. In case of test failure, the test was repeated, and the first valid result was used for the analysis.

Blinded EclLAM testing at MSD followed a previously established assay protocol except for the addition of a pre-concentration step to improve the limit of detection.(14) To pre-concentrate, frozen urine aliquots were thawed to ambient temperature, mixed by inversion, and 490 µL sample was added to Amicon Ultra-0.5 mL centrifugal filters (MilliporeSigma, Burlington, USA) with a 3 KDa cut-off. After ultrafiltration for 20 minutes at 14000 g, approx. 25 µL deionized water was added to the retentate (approx. 45 µL) to get a total of 70 µL and an estimated concentration factor of 7. Prior to analysis, samples were heat treated at 85 °C for 10 min. The immunoassays for LAM used the same antibodies as FujiLAM in a sandwich immunoassay format employing ECL detection as described in the appendix (p5) and elsewhere (14, 17, 37). The research team performing EclLAM had no access to all other clinical, demographic and test data associated with the samples.

**Statistical analysis**

Before data analysis, clinical investigators, who were masked to index test results, categorised patients as having definite TB, possible TB, not TB, and unclassifiable using a combination of clinical and laboratory findings (appendix p4). Patients with definite TB had microbiologically confirmed *Mtb* (any culture or any Xpert positive for *Mtb* during admission). Patients defined as not TB had all microscopy, culture, and Xpert test results negative for *Mtb*, had not started TB treatment, and were alive or had improvement in clinical tuberculosis symptoms at two months' follow-up. Patients defined as possible TB did not satisfy the criteria for definite TB but had clinical features suggestive of TB and were started on TB treatment. Patients that did not fall into any of these categories were defined as unclassifiable and were removed from the main analyses. Definite TB and not TB categories were used to allocate patients into positive and negative, respectively, for both the microbiological reference standard (MRS)

and composite reference standard (CRS). The possible TB group was considered negative by MRS, but positive by CRS as previously proposed in a study guidance publication.(28)

Simple descriptive statistics were used to characterize cohorts. We calculated the point estimates and 95% Wilson Confidence Intervals (CI) for the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), and negative likelihood ratio (LR−) for FujiLAM, AlereLAM, EclLAM, Sputum Smear Microscopy (SSM), Xpert, and combinations of FujiLAM+SSM and FujiLAM+Xpert by comparison with the MRS and CRS. The R package eulerr(38) was used to generate area-proportional Euler diagrams to illustrate the number of positive test results by test in the definite TB group. Fagan nomograms were used to illustrate pretest and posttest probabilities of FujiLAM and AlereLAM.

The cut-off of the EclLAM assay was set at 5.2 pg/ml after review of the data to achieve a specificity of at least 98%. Diagnostic accuracy for LAM assays was determined separately for each cohort and post-hoc was analysed by MGIT time to detection (MGIT TTD), SSM status and semiquantitative Xpert result. EclLAM concentration-based ROC curves were used to show the relationship between LAM concentration, sensitivity and specificity.

**Study approval**

The study was approved by the Human Research Ethics Committee of the University of Cape Town (Cape Town, South Africa), the City of Cape Town (Cape Town, South Africa, Ref. 10364a), the Universidad Peruana Cayetano Heredia (Lima, Peru), and the Peruvian Ministry of Health (Lima, Peru, Ref. 18829-2016). Written informed consent was obtained from patients, as per study protocols. Study participation did not affect standard of care. This study is reported in accordance with the Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines. The study protocol is available on request and relevant data are available online.

**Authors Contributions:** TB, EIR, and CMD designed and oversaw the study. MPN, EG, SS, JH and TCN, coordinated the individual study sites in South Africa and Peru. TB, GBS, MT, TP, TLL, AP, and EM contributed to EclLAM assay and reagent development. KC, RS and SM coordinated measurement of AlereLAM and FujiLAM. AM and RS coordinated data collection and management. TB did the statistical analysis and TB and AJZ wrote the first manuscript draft. All authors contributed to interpretation of data and editing of the article and approved the final version of the manuscript. Authorship, including the order of co-first authors was based on ICMJE criteria.

**Acknowedgments**

The authors thank Mark D. Perkins, and Ranald Sutherland for helping with the conceptualisation of this work, Michelle A. Bulterys, Amanda Jackson, Celeste Worship, Meagan McMaster, Aurélien Macé, Stefano Ongarello and André Trollip for data management, and the clinical and laboratory teams including Jimena Collantes, Cesar Ugarte, Nchimunya Hapeemla, and Widaad Zemanay at the partner sites for their efforts in the implementation, conduct, and timely completion of the study.

**Funding**

This work was funded by the Global Health Innovative Technology (GHIT) Fund (grant number G2015-201), the UK Department for International Development (DFID) (grant number 300341-102), the Dutch Ministry of Foreign Affairs (grant number PDP15CH14), the Bill and Melinda Gates Foundation (grant number OPP1151258), the Australian Department of Foreign Affairs and Trade (grant number 70957), the German Federal Ministry of Education and Research through Kreditanstalt für Wiederaufbau (grant number 2020 60 457), and the National Institutes of Health and National Institute of Allergy and Infectious Diseases (NIH/NIAID; grant number 1R01AI104589).

## Competing Interests

TB, AM, EIP, RS, EM, SGS and CMD were previously or are currently employed by FIND. FIND operates on a charitable non-profit basis and has no financial interest in any of the commercial partners and/or their products. In the case on hand, all funding for this project originates from donors who have, in turn, no financial interest in any of the products which are the subject of these studies. Furthermore, contractual agreements between FIND and its partners, are compliant with the principals of Global Access (as set out under the FIND Policy - https://www.finddx.org/policies/), which includes making all study data freely and readily available, and ensuring the products under study, becomes available to the public sector in resource poor countries at an affordable price and that potential intellectual property issues are not an impediment to product availability. GBS, MT and TP are employed by Meso Scale Diagnostics LLC (Rockville, USA) and received funding from FIND. TB and AP report patents in the field of lipoarabinomannan detection. All other authors declare no competing interests.

**References:**


1. World Health Organization. *Global Tuberculosis Report [Internet]*. 2019:

2. World Health Organization. *High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting [Internet]*. Geneva, Switzerland: 2014:

3. Denkinger CM et al. Defining the Needs for Next Generation Assays for Tuberculosis [Internet]. *J. Infect. Dis.* 2015;211(suppl_2):S29–S38.

4. MacLean E et al. A systematic review of biomarkers to detect active tuberculosis [Internet]. *Nat. Microbiol.* 2019;4(5):748–758.

5. Bulterys MA et al. Point-Of-Care Urine LAM Tests for Tuberculosis Diagnosis: A Status Update [Internet]. *J. Clin. Med.* 2019;9(1):111.

6. Bjerrum S et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in people living with HIV [Internet]. *Cochrane Database Syst. Rev.* [published online ahead of print: October 21, 2019]; doi:10.1002/14651858.CD011420.pub3

7. World Health Organization. *Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV - Policy update 2019 [Internet]*. 2019:

8. World Health Organization. *Consolidated Guidelines on Tuberculosis - Module 3: Diagnosis [Internet]*. Geneva, Switzerland: 2020:

9. Peter JG et al. Effect on mortality of point-of-care, urine-based lipoarabinomannan testing to guide tuberculosis treatment initiation in HIV-positive hospital inpatients: a pragmatic, parallel-group, multicountry, open-label, randomised controlled trial [Internet]. *Lancet* 2016;387(10024):1187–1197.

10. Gupta-Wright A et al. Rapid urine-based screening for tuberculosis in HIV-positive patients admitted to hospital in Africa (STAMP): a pragmatic, multicentre, parallel-group, double-blind, randomised controlled trial [Internet]. *Lancet* 2018;392(10144):292–301.

11. Sahle SN et al. Performance of point-of-care urine test in diagnosing tuberculosis suspects with and without HIV infection in selected peripheral health settings of Addis Ababa, Ethiopia [Internet]. *BMC Res. Notes* 2017;10(1):74.

12. Nicol MP et al. Urine lipoarabinomannan testing for diagnosis of pulmonary tuberculosis in children: a prospective study [Internet]. *Lancet Glob. Heal.* 2014;2(5):e278–e284.

13. Kroidl I et al. Performance of urine lipoarabinomannan assays for paediatric tuberculosis in Tanzania [Internet]. *Eur. Respir. J.* 2015;46(3):761–770.

14. Sigal GB et al. A Novel Sensitive Immunoassay Targeting the 5-Methylthio-D-Xylofuranose–Lipoarabinomannan Epitope Meets the WHO's Performance Target for Tuberculosis Diagnosis [Internet]. *J. Clin. Microbiol.* 2018;56(12):1–17.

15. Suwanpimolkul G et al. Utility of urine lipoarabinomannan (LAM) in diagnosing tuberculosis and predicting mortality with and without HIV: prospective TB cohort from the Thailand Big City TB Research Network [Internet]. *Int. J. Infect. Dis.* 2017;59:96–102.

16. Broger T et al. Novel lipoarabinomannan point-of-care tuberculosis test for people with HIV: a diagnostic accuracy study [Internet]. *Lancet Infect. Dis.* 2019;19(8):852–861.

17. Choudhary A et al. Characterization of the Antigenic Heterogeneity of Lipoarabinomannan, the Major Surface Glycolipid of Mycobacterium tuberculosis , and Complexity of Antibody Specificities toward This Antigen [Internet]. *J. Immunol.* 2018;200(9):3053–3066.

18. Joe M et al. The 5-deoxy-5-methylthio-xylofuranose residue in mycobacterial lipoarabinomannan. Absolute stereochemistry, linkage position, conformation, and immunomodulatory activity. *J. Am. Chem. Soc.* 2006;128(15):5059–5072.

19. Broger T et al. Diagnostic accuracy of a novel tuberculosis point-of-care urine lipoarabinomannan assay for people living with HIV: a meta-analysis of individual in- and outpatient data [Internet]. *PLoS Med.* 2020;17(5):e1003113.

20. Kerkhoff AD et al. Diagnostic sensitivity of SILVAMP TB-LAM (FujiLAM) point-of-care urine assay for extra-

pulmonary tuberculosis in people living with HIV [Internet]. *Eur. Respir. J.* 2020;55(2):1901259.

21. Sossen B et al. "SILVAMP TB LAM" Rapid Urine Tuberculosis Test Predicts Mortality in Patients Hospitalized With Human Immunodeficiency Virus in South Africa [Internet]. *Clin. Infect. Dis.* [published online ahead of print: January 9, 2020]; doi:10.1093/cid/ciaa024

22. Ricks S, Denkinger CM, Schumacher SG, Hallett TB, Arinaminpathy N. The potential impact of urine-LAM diagnostics for tuberculosis in South Africa: a modelling analysis. *PLoS Med.* :(submitted).

23. Huerga H et al. Incremental Yield of Including Determine-TB LAM Assay in Diagnostic Algorithms for Hospitalized and Ambulatory HIV-Positive Patients in Kenya [Internet]. *PLoS One* 2017;12(1):e0170976.

24. Sun AY et al. Modeling the impact of alternative strategies for rapid molecular diagnosis of tuberculosis in Southeast Asia [Internet]. *Am. J. Epidemiol.* 2013;178(12):1740–1749.

25. Krishna R et al. Cost-effectiveness of a novel lipoarabinomannan test for tuberculosis in patients with HIV. *Eur. Respir. J.* :(submitted).

26. Paris L et al. Urine lipoarabinomannan glycan in HIV-negative patients with pulmonary tuberculosis correlates with disease severity [Internet]. *Sci. Transl. Med.* 2017;9(420):eaal2807.

27. Bjerrum S et al. Diagnostic Accuracy of a Novel and Rapid Lipoarabinomannan Test for Diagnosing Tuberculosis Among People With Human Immunodeficiency Virus [Internet]. *Open Forum Infect. Dis.* 2020;7(1):ofz530.

28. Drain PK et al. Guidance for Studies Evaluating the Accuracy of Biomarker-Based Nonsputum Tests to Diagnose Tuberculosis [Internet]. *J. Infect. Dis.* 2019;220(Supplement_3):S108–S115.

29. Shapiro AE et al. A point-of-care test to concentrate and detect urine LAM for TB diagnosis: results from the first-in-human study of FLOW-TB. In: *50th World Conference on Lung Health of the International Union Against Tuberculosis and Lung Disease, Hyderabad, India*. 2019:

30. Wood A et al. Ultrasensitive detection of lipoarabinomannan with plasmonic grating biosensors in clinical samples of HIV negative patients with tuberculosis [Internet]. *PLoS One* 2019;14(3):e0214161.

31. Hamasur B, Ignatowicz L, Ramachandraiah H. SpeClean: urine sample treatment method for ultra-sensitive

LAM diagnostics. In: *50th World Conference on Lung Health of the International Union Against Tuberculosis and Lung Disease, Hyderabad, India*. 2019:

32. WHO International biological reference preparations [Internet]https://www.who.int/biologicals/reference_preparations/en/. cited March 31, 2020

33. Das S et al. Performance of a High-Sensitivity Rapid Diagnostic Test for Plasmodium falciparum Malaria in Asymptomatic Individuals from Uganda and Myanmar and Naive Human Challenge Infections. [Internet]. *Am. J. Trop. Med. Hyg.* 2017;97(5):1540–1550.

34. Broger T et al. Sensitive electrochemiluminescence (ECL) immunoassays for detecting lipoarabinomannan (LAM) and ESAT-6 in urine and serum from tuberculosis patients [Internet]. *PLoS One* 2019;14(4):e0215443.

35. Broger T, Muyoyeta M, Kerkhoff AD, Denkinger CM, Moreau E. Tuberculosis test results using fresh versus biobanked urine samples with FujiLAM [Internet]. *Lancet Infect. Dis.* 2020;20(1):22–23.

36. Fujifilm SILVAMP TB LAM test procedure [Internet]. *YouTube* https://youtu.be/aK-QtzkLBug. cited

37. Kawasaki M et al. Lipoarabinomannan in sputum to detect bacterial load and treatment response in patients with pulmonary tuberculosis: Analytic validation and evaluation in two cohorts [Internet]. *PLOS Med.* 2019;16(4):e1002780.

38. Larsson J. eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses. *R Packag. version 6.1.0* 2020;

**Figure legends**

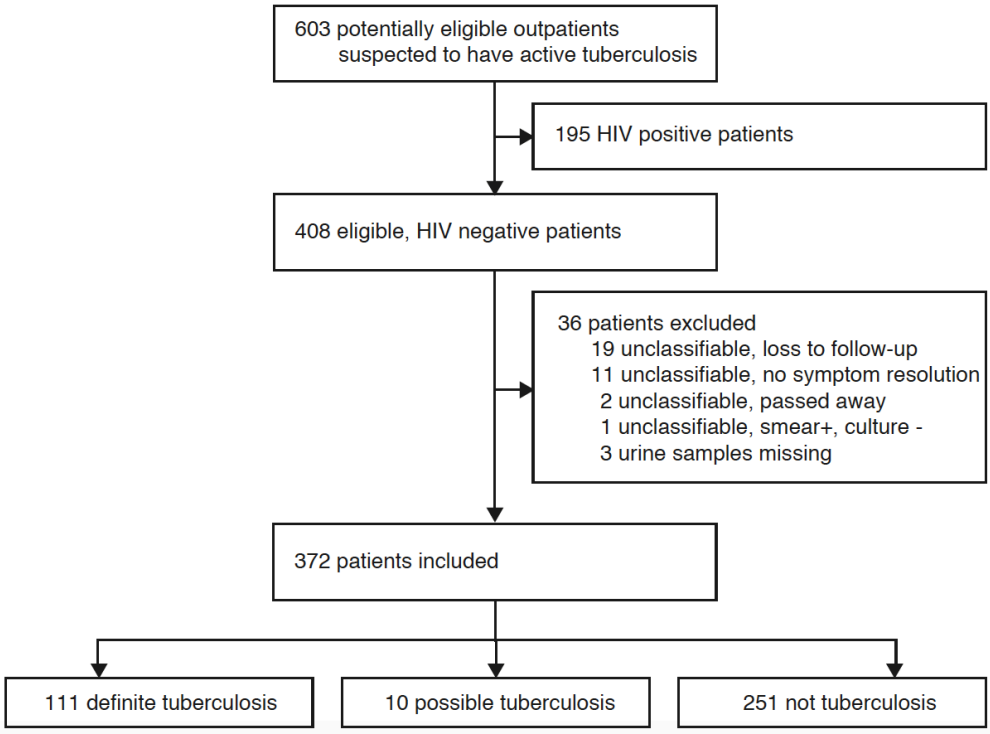



**Figure 1:** Study flow diagram

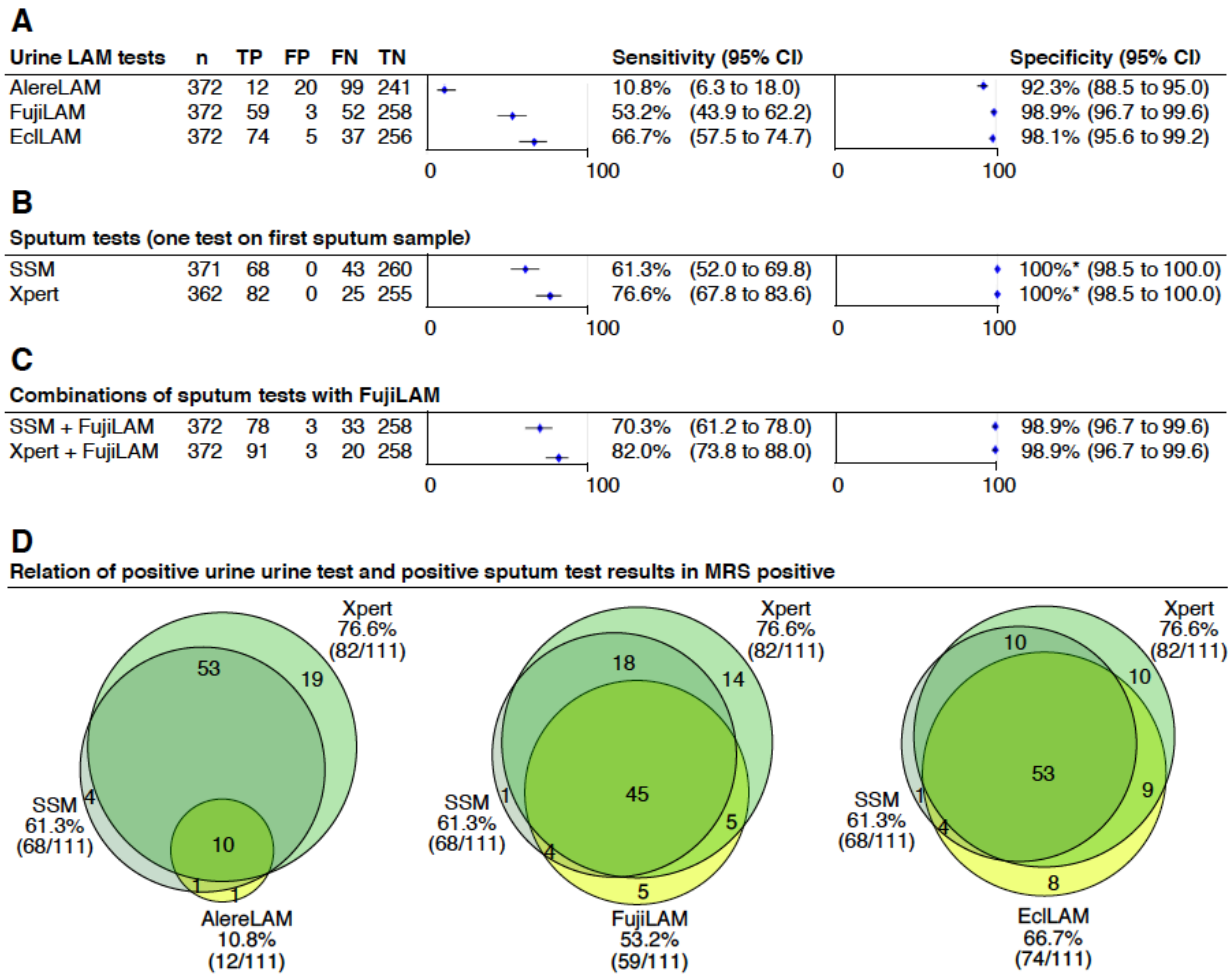



**Figure 2:** Diagnostic accuracy against the microbiological reference standard of (A) Urine LAM tests, (B) sputum tests, (C) combinations of sputum tests with FujiLAM, and (D) positive urine LAM tests in relation with positive sputum tests. *SSM and Xpert were part of the microbiological reference standard and therefore specificity is 100%. TP=true positive. FP=false positive. FN=false negative. TN=true negative. AlereLAM=Alere Determine TB LAM Ag assay. FujiLAM=Fujifilm SILVAMP TB LAM assay. EclLAM=electrochemiluminescence-based LAM detection assay. SSM=Sputum smear microscopy. Xpert=Gene Xpert MTB/RIF assay. MRS=microbiological reference standard.

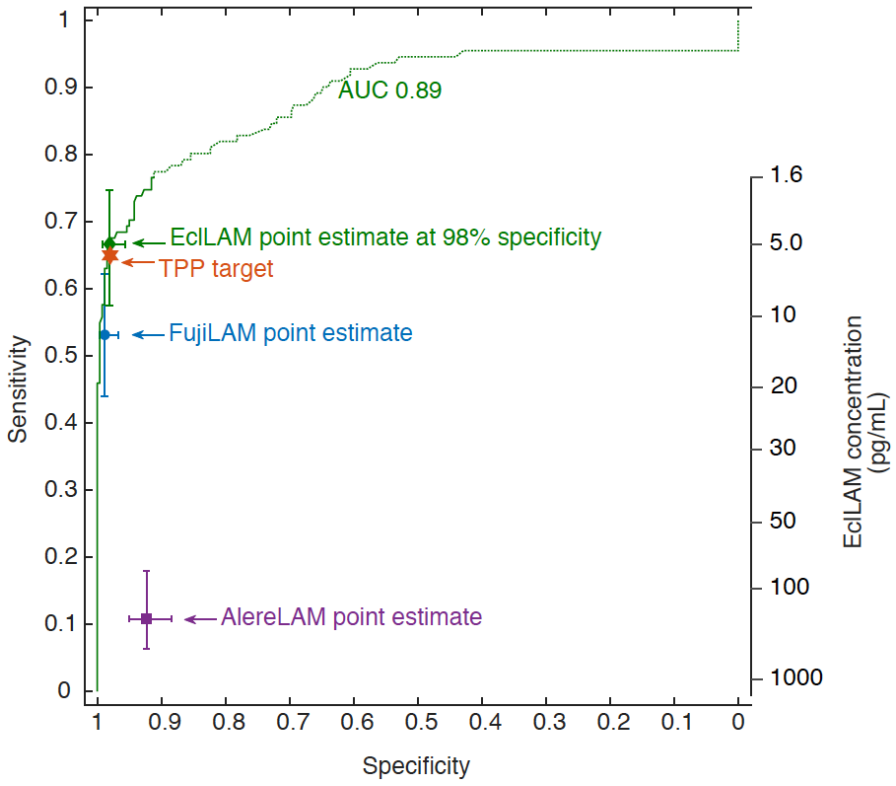



**Figure 3:** Receiver operating characteristic (ROC) analysis of the EclLAM concentration data compared with FujiLAM, AlereLAM and EclLAM performance (n=372 patients). The EclLAM concentration for the ROC curve is indicated on the secondary y-axis. The ROC curve was restricted by the LOD (1.6 pg/ml) of the EclLAM assay, meaning that lower concentrations could not be measured so that the upper part of the ROC curve (green dotted line) should be treated with caution. AUC= area under the curve. AlereLAM=Alere Determine TB LAM Ag assay. FujiLAM=Fujifilm SILVAMP TB LAM assay. EclLAM=electrochemiluminescence-based LAM detection assay.

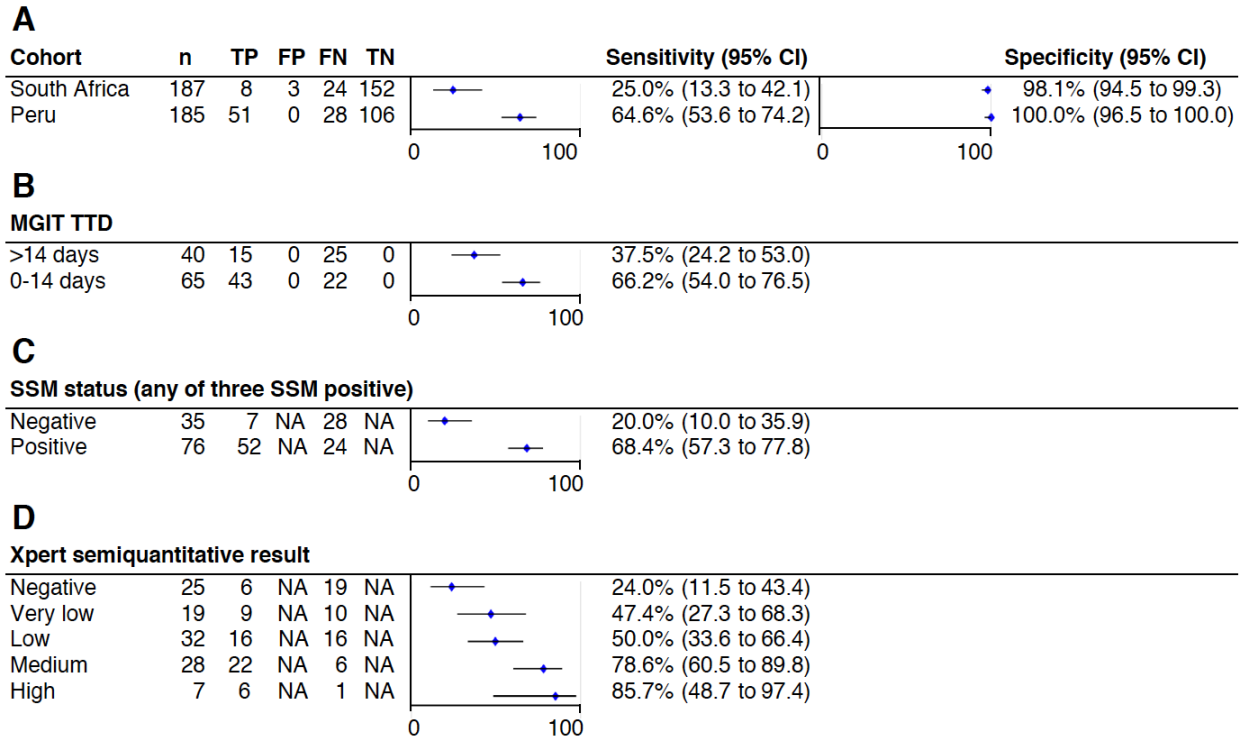



**Figure 4:** Subgroup analysis of FujiLAM. Sensitivity and specificity against microbiological reference standard of (A) FujiLAM by study site, (B) FujiLAM by MGIT TTD, and (C) FujiLAM by SSM status (D) FujiLAM by semiquantitative Xpert result. TP=true positive. FP=false positive. FN=false negative. TN=true negative. SSM=Sputum smear microscopy. Xpert=Gene Xpert MTB/RIF assay. MGIT TTD=Mycobacteria growth indicator tube time to detection. NA=Not applicable.

| | All participants (N=372) | | South Africa, Cape Town (N=187) | | Peru, Lima (N=185) | |
|---|---|---|---|---|---|---|
| **Demographic or clinical characteristic** | | | | | | |
| Median age [IQR] - years | 32 | (25–47) | 33 | (26–48) | 31 | (24–42) |
| Sex | | | | | | |
| Female - no. (%) | 158/372 | (42%) | 96/187 | (51%) | 62/185 | (34%) |
| Male- no. (%) | 214/372 | (58%) | 91/187 | (49%) | 123/185 | (66%) |
| History of TB - no. (%) | 52/372 | (14%) | 22/187 | (12%) | 30/185 | (16%) |
| | | | | | | |
| **Distribution in diagnostic categories** | | | | | | |
| Definite TB - no. (%) | 111/372 | (30%) | 32/187 | (17%) | 79/185 | (43%) |
| Possible TB - no. (%) | 10/372 | (3%) | 1/187 | (1%) | 9/185 | (5%) |
| Not TB - no. (%) | 251/372 | (67%) | 154/187 | (82%) | 97/185 | (52%) |
| | | | | | | |
| **Reference standard** | | | | | | |
| MRS positive - no. (%) | 111/372 | (30%) | 32/187 | (17%) | 79/185 | (43%) |
| CRS positive - no. (%) | 121/372 | (33%) | 33/187 | (18%) | 88/185 | (48%) |
| | | | | | | |
| **Distribution in MRS positive patients** | | | | | | |
| Sputum smear microscopy positive (any of three tests on three sputum samples positive) - no. (%) | 76/111 | (68%) | 15/32 | (47%) | 61/79 | (77%) |
| Sputum smear microscopy positive (one test on first sputum sample) - no. (%) | 68/111 | (61%) | 12/32 | (38%) | 56/79 | (71%) |
| Blood culture positive - no. (%) | 0/111 | (0%) | 0/32 | (0%) | 0/79 | (0%) |
| Urine Xpert positive - no. (%) | 4/111 | (4%) | 0/32 | (0%) | 4/79 | (5%) |
| Sputum Xpert positive (any of three tests on three sputum samples positive) - no. (%) | 102/111 | (92%) | 27/32 | (84%) | 75/79 | (95%) |
| Sputum Xpert positive (one test on first sputum sample) - no. (%) | 82/111 | (74%) | 22/32 | (69%) | 60/79 | (76%) |
| Sputum Xpert result (from testing on first sputum sample, Xpert was repeated once using the same sample in case of an indeterminate result) | | | | | | |
| Negative - no. (%) | 25/111 | (23%) | 10/32 | (31%) | 15/79 | (19%) |
| Very Low - no. (%) | 19/111 | (17%) | 4/32 | (13%) | 15/79 | (19%) |
| Low - no. (%) | 32/111 | (29%) | 6/32 | (19%) | 26/79 | (33%) |
| Medium - no. (%) | 28/111 | (25%) | 7/32 | (22%) | 21/79 | (27%) |
| High - no. (%) | 7/111 | (6%) | 5/32 | (16%) | 2/79 | (3%) |
| Average sputum MGIT time to detection | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| MGIT negative (only Xpert positive) - no. (%) | 6/111 | (5%) | 3/32 | (9%) | 3/79 | (4%) |
| >14 days - no. (%) | 40/111 | (36%) | 13/32 | (41%) | 27/79 | (34%) |
| 0–14 days - no. (%) | 65/111 | (59%) | 16/32 | (50%) | 49/79 | (62%) |

**Table 1:** Demographic and clinical characteristics of the study participants

| Test | n | PPV (95% CI) | | NPV (95% CI) | | LR+ | LR − | PPV at 20%Prev. | NPV at 20%Prev. |
|---|---|---|---|---|---|---|---|---|---|
| AlereLAM | 372 | 37.5% | (22.9 to 54.7) | 70.9% | (65.8 to 75.5) | 1.4 | 0.97 | 26.1% | 80.5% |
| FujiLAM | 372 | 95.2% | (86.7 to 98.3) | 83.2% | (78.7 to 87.0) | 46.2 | 0.47 | 92.0% | 89.4% |
| EclLAM | 372 | 93.7% | (86.0 to 97.3) | 87.4% | (83.1 to 90.7) | 34.8 | 0.34 | 89.7% | 92.2% |
| SSM | 371 | 100%* | (94.7 to 100.0) | 85.8% | (81.4 to 89.3) | * | 0.39 | 100.0%* | 91.2% |
| Xpert | 362 | 100%* | (95.5 to 100.0) | 91.1% | (87.2 to 93.9) | * | 0.23 | 100.0%* | 94.5% |
| FujiLAM + SSM | 372 | 96.3% | (89.7 to 98.7) | 88.7% | (84.5 to 91.8) | 61.1 | 0.30 | 93.9% | 93.0% |
| FujiLAM + Xpert | 372 | 96.8% | (91.0 to 98.9) | 92.8% | (89.1 to 95.3) | 71.3 | 0.18 | 94.7% | 95.6% |

**Table 2:** Predictive values and likelihood ratios of urine LAM tests, sputum tests, and combinations of sputum tests with FujiLAM against the microbiological reference standard. PPV and NPV were re-calculated at an assumed TB prevalence of 20%. *SSM and Xpert were part of the microbiological reference standard and therefore PPV is 100% and LR+ not defined. AlereLAM=Alere Determine TB LAM Ag assay. FujiLAM=Fujifilm SILVAMP TB LAM assay. EclLAM=electrochemiluminescence-based LAM detection assay. SSM=Sputum smear microscopy. Xpert=Gene Xpert MTB/RIF assay. PPV=Positive Predictive Value. NPV=Negative Predictive Value. LR+=Positive likelihood ratio. LR −=Negative likelihood ratio. Prev.=Prevalence.