Diagnostic yield of urine lipoarabinomannan and sputum tuberculosis tests in people living with HIV: a systematic review and meta-analysis of individual participant data

Tobias Broger, Lisa Koeppel, Helena Huerga, Poppy Miller, Ankur Gupta-Wright, François-Xavier Blanc, Aliasgar Esmail, Byron W P Reeve, Marco Floridia, Andrew D Kerkhoff, Fausto Ciccacci, Margaret P Kasaro, Swe Swe Thit, Mathieu Bastard, Gabriella Ferlazzo, Christina Yoon, Daniël J Van Hoving, Bianca Sossen, Juan Ignacio García, Matthew J Cummings, Rachel M Wake, Josh Hanson, Adithya Cattamanchi, Graeme Meintjes, Gary Maartens, Robin Wood, Grant Theron, Keertan Dheda, Ioana Diana Olaru, Claudia M Denkinger* on behalf of the TBYield Study Consortium†*

Summary

Background Sputum is the most widely used sample to diagnose active tuberculosis, but many people living with HIV are unable to produce sputum. Urine, in contrast, is readily available. We hypothesised that sample availability influences the diagnostic yield of various tuberculosis tests.

Methods In this systematic review and meta-analysis of individual participant data, we compared the diagnostic yield of point-of-care urine-based lipoarabinomannan tests with that of sputum-based nucleic acid amplification tests (NAATs) and sputum smear microscopy (SSM). We used microbiologically confirmed tuberculosis based on positive culture or NAAT from any body site as the denominator and accounted for sample provision. We searched PubMed, Web of Science, Embase, African Journals Online, and clinicaltrials.gov from database inception to Feb 24, 2022 for randomised controlled trials, cross-sectional studies, and cohort studies that assessed urine lipoarabinomannan point-of-care tests and sputum NAATs for active tuberculosis detection in participants irrespective of tuberculosis symptoms, HIV status, CD4 cell count, or study setting. We excluded studies in which recruitment was not consecutive, systematic, or random; provision of sputum or urine was an inclusion criterion; less than 30 participants were diagnosed with tuberculosis; early research assays without clearly defined cutoffs were tested; and humans were not studied. We extracted study-level data, and authors of eligible studies were invited to contribute deidentified individual participant data. The main outcomes were the tuberculosis diagnostic yields of urine lipoarabinomannan tests, sputum NAATs, and SSM. Diagnostic yields were predicted using Bayesian random-effects and mixed-effects meta-analyses. This study is registered with PROSPERO, CRD42021230337.

Findings We identified 844 records, from which 20 datasets and 10 202 participants (4561 [45%] male participants and 5641 [55%] female participants) were included in the meta-analysis. All studies assessed sputum Xpert (MTB/RIF or Ultra, Cepheid, Sunnyvale, CA, USA) and urine Alere Determine TB LAM (AlereLAM, Abbott, Chicago, IL, USA) in people living with HIV aged 15 years or older. Nearly all (9957 [98%] of 10 202) participants provided urine, and 82% (8360 of 10 202) provided sputum within 2 days. In studies that enrolled unselected inpatients irrespective of tuberculosis symptoms, only 54% (1084 of 1993) of participants provided sputum, whereas 99% (1966 of 1993) provided urine. Diagnostic yield was 41% (95% credible interval [CrI] 15–66) for AlereLAM, 61% (95% CrI 25–88) for Xpert, and 32% (95% CrI 10–55) for SSM. Heterogeneity existed across studies in the diagnostic yield, influenced by CD4 cell count, tuberculosis symptoms, and clinical setting. In predefined subgroup analyses, all tests had higher yields in symptomatic participants, and AlereLAM yield was higher in those with low CD4 counts and inpatients. AlereLAM and Xpert yields were similar among inpatients in studies enrolling unselected participants who were not assessed for tuberculosis symptoms (51% *vs* 47%). AlereLAM and Xpert together had a yield of 71% in unselected inpatients, supporting the implementation of combined testing strategies.

Interpretation AlereLAM, with its rapid turnaround time and simplicity, should be prioritised to inform tuberculosis therapy among inpatients who are HIV-positive, regardless of symptoms or CD4 cell count. The yield of sputum-based tuberculosis tests is undermined by people living with HIV who cannot produce sputum, whereas nearly all participants are able to provide urine. The strengths of this meta-analysis are its large size, the carefully harmonised denominator, and the use of Bayesian random-effects and mixed-effects models to predict yields; however, data were geographically restricted, clinically diagnosed tuberculosis was not considered in the denominator, and little information exists on strategies for obtaining sputum samples.

Funding FIND, the Global Alliance for Diagnostics.

Lancet Glob Health 2023; 11: e903–16

See **Comment** page e809

*Contributed equally

†Members of the TBYield Study Consortium are listed in the appendix

Division of Infectious Disease and Tropical Medicine (T Broger MSc, L Koeppel PhD, I D Olaru PhD, C M Denkinger MD) **and German Center for Infection Research, partner site** (C M Denkinger), **Heidelberg University Hospital, Heidelberg, Germany; Field Epidemiology Department, Epicentre, Paris, France** (H Huerga PhD, M Bastard MSc); **New Zealand Institute for Plant and Food Research, Auckland, New Zealand** (P Miller PhD); **Institute for Global Health, University College London, London, UK** (A Gupta-Wright PhD); **Clinical Research Department** (A Gupta-Wright, I D Olaru) **and Faculty of Infectious and Tropical Diseases, Department of Immunology and Infection** (K Dheda PhD), **London School of Hygiene & Tropical Medicine, London, UK; Service de Pneumologie, l'institut du thorax, Nantes Université, CHU Nantes, Nantes, France** (F-X Blanc MD); **Centre for Lung Infection and Immunity, Division of Pulmonology, Department of Medicine and UCT Lung Institute** (A Esmail PhD, K Dheda), **South African MRC Centre for the Study of Antimicrobial Resistance** (A Esmail, K Dheda), **Division of Emergency Medicine** (D J Van Hoving PhD), **Department of Medicine** (G Meintjes PhD, G Maartens MMed, B Sossen MBChB), **Wellcome Centre for Infectious Diseases**

Research in Africa, Institute of Infectious Disease and Molecular Medicine (G Meintjes, G Maartens), **and Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences** (R Wood DSc, B Sossen), **University of Cape Town, Cape Town, South Africa; DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, Faculty of Medicine and Health Sciences** (B W P Reeve PhD, G Theron PhD), **South African Medical Research Council Centre for Tuberculosis Research, Faculty of Medicine and Health Sciences** (B W P Reeve, G Theron), **Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences** (B W P Reeve, G Theron), **and Division of Emergency Medicine** (D J Van Hoving), **Stellenbosch University, Cape Town, South Africa; National Center for Global Health, Istituto Superiore di Sanità, Rome, Italy** (M Floridia PhD); **Division of HIV, Infectious Diseases and Global Medicine** (A D Kerkhoff MD) **and Department of Medicine, Division of Pulmonary and Critical Care Medicine** (C Yoon MD), **Zuckerberg San Francisco General Hospital, San Francisco, CA, USA; Trauma Center** (A D Kerkhoff) **and Center for Tuberculosis** (A D Kerkhoff, C Yoon, A Cattamanchi MD), **University of California San Francisco, San Francisco, CA, USA; UniCamillus, International University of Health and Medical Science, Rome, Italy** (F Ciccacci PhD); **Community of Sant'Egidio, DREAM programme, Rome, Italy** (F Ciccacci); **Centre for Infectious Disease Research in Zambia, Lusaka, Zambia** (M P Kasaro MScPH); **UNC Global Projects, LLC Zambia, Lusaka, Zambia** (M P Kasaro); **Department of Medicine, University of Medicine 2, Yangon, Myanmar** (S S Thit FRCP); **Department of Medicine, Médecins Sans Frontières, Paris, France** (G Ferlazzo MD); **Population Health Program, Tuberculosis Group, Texas Biomedical Research Institute, San Antonio, TX, USA** (J I García PhD); **Division of**

Research in context

Evidence before this study

Despite advances in tuberculosis diagnostics and their global roll-out, most methods still require sputum for testing; however, people living with HIV might have difficulty producing sputum samples. Furthermore, the diagnosis of tuberculosis in people living with HIV can be challenging, as they are more likely to produce paucibacillary sputum samples and to have extrapulmonary disease. We searched PubMed and the Cochrane Infectious Diseases Group Specialized Register for meta-analyses published from inception until Feb 24, 2022 using search terms "tuberculosis", "lipoarabinomannan", "LAM", "Xpert", and terms related to these concepts without restrictions. We included only publications in English. The identified meta-analyses reported 42% sensitivity and 91% specificity for urine lipoarabinomannan, 77% sensitivity and 98% specificity for sputum Xpert MTB/RIF, 88% sensitivity and 93% specificity for Xpert Ultra, and 53% sensitivity and 96% specificity for sputum smear fluorescent microscopy in people living with HIV (appendix p 24). We found no meta-analyses on diagnostic yield and sample provision but identified only meta-analyses reporting on diagnostic accuracy.

Added value of this study

To our knowledge, this is the first individual participant data meta-analysis assessing diagnostic yield of urine lipoarabinomannan, sputum nucleic acid amplification tests, and sputum smear microscopy. The diagnostic yield of a test better reflects its clinical usefulness as it examines the number of individuals testing positive among all individuals eligible for testing, irrespective of their ability to actually provide a sample. Nearly all people living with HIV were able to produce a urine sample for testing, but one in five participants was unable to produce sputum. Thus, the diagnostic yields of sputum Xpert and sputum smear microscopy were lower than their sensitivities. By contrast, the diagnostic yield of urine AlereLAM was unaffected by the ability to provide a sample, because urine samples were readily obtained from almost all people living with HIV. This study emphasises the challenges of diagnosing tuberculosis in people living with HIV using only sputum-based tests, especially in those requiring hospitalisation.

Implications of all the available evidence

Our study has policy implications for the diagnosis of tuberculosis in people living with HIV. Urine lipoarabinomannan testing in hospitalised people living with HIV should be prioritised in addition to sputum-based diagnostics to maximise yield. In outpatients, urine lipoarabinomannan testing should be used to aid in the diagnosis of tuberculosis in people living with HIV with tuberculosis symptoms. Participants unable to produce a diagnostic sample should not be excluded from future tuberculosis diagnostic studies. Both the number of participants unable to produce a sample and the diagnostic yield of tests should be reported alongside sensitivity and specificity. Reporting these values is of particular importance when considering novel tuberculosis diagnostics based on non-sputum samples (eg, urine, swab, breath, and blood-based assays), which might have lower sensitivity than existing sputum-based assays but similar diagnostic yields when the ability to collect a sample for testing is considered.

Introduction

In 2021, it was estimated that 10·6 million people developed active tuberculosis, 1·6 million of whom died. Of these 10·6 million cases, only 6·4 million were reported, and inadequate tuberculosis diagnostics are still a major challenge in reducing disease burden.[1,2] Sputum has been used to diagnose tuberculosis for more than a century and is the most used sample type. However, sputum can be difficult to obtain, particularly in people living with HIV, and it cannot be used to diagnose extrapulmonary tuberculosis. Furthermore, results for diagnostic tests that rely on sputum are usually not available during the same clinical encounter.[3] To address these gaps, WHO developed a target product profile in 2014 to encourage the development of non-sputum biomarker tests, with the ultimate aim of enabling appropriate tuberculosis treatment initiation during the same clinical encounter.[4] Detection of lipoarabinomannan antigen in urine has the greatest potential to fill this diagnostic void.[5,6] In 2019, WHO made a conditional recommendation to use the Alere Determine TB-LAM Ag (AlereLAM, Abbott, Chicago, IL, USA) lateral flow assay for assisting in the diagnosis of active tuberculosis in people living with HIV.[7] Despite WHO's recommendation and evidence that implementation of AlereLAM reduces tuberculosis-related mortality,[8,9] adoption and uptake of the test have been slow.[10,11]

Previous meta-analyses of urine lipoarabinomannan tests, sputum nucleic acid amplification tests (NAATs; eg, Xpert MTB/RIF or Ultra, Cepheid, Sunnyvale, CA, USA), and sputum smear microscopy (SSM) have focused only on diagnostic sensitivity and specificity.[12–16] However, test accuracy does not account for ability to provide a sample. Modelling studies showed that a rapid test with moderate sensitivity on an easily obtainable non-sputum sample could be more useful than a sensitive test that is reliant on sputum, which can be difficult to obtain.[17,18] Thus, diagnostic accuracy alone gives an incomplete picture of the usefulness of a test to diagnose tuberculosis in routine settings. By contrast, the diagnostic yield of a test considers both the patient's ability to produce the sample necessary to conduct the test and the sensitivity of the test. Diagnostic yield measures the proportion of tuberculosis cases that are detected by a diagnostic test among all tuberculosis cases identified as positive (ie, the denominator).

We aimed to do an individual participant data (IPD) meta-analysis to determine the comparative tuberculosis diagnostic yield of urine lipoarabinomannan point-of-care tests on the first available urine sample, sputum NAATs on the first available sputum sample, and sputum smear microscopy on the first available sputum sample within 2 days of enrolment against a harmonised denominator. The advantages of IPD meta-analysis over aggregate meta-analysis are the ability to harmonise the variables and denominator across studies, inclusion of participants that were excluded from the primary studies, and the ability to assess interactions and perform subgroup analyses.

Methods

Search strategy and selection criteria

For this systematic review and IPD meta-analysis, we searched PubMed, Web of Science, Embase, and African Journals Online for papers published between database inception and Feb 24, 2022 without any language restrictions, using search terms that combined the outcome ("tuberculosis") and the intervention (biomarker "lipoarabinomannan" or "LAM") with the test name or principle ("AlereLAM", "antigen", "lateral flow", "urine", "FujiLAM", etc). Search terms used for each database are shown in the appendix (p 3). We also searched the references of identified studies and review articles, contacted tuberculosis researchers, and searched clinicaltrials.gov to identify unpublished studies.

We included randomised controlled trials, cross-sectional studies, and cohort studies without any date restrictions that assessed both urine lipoarabinomannan point-of-care tests and sputum NAATs for active tuberculosis detection in participants irrespective of tuberculosis symptoms, HIV status, CD4 cell count, or study setting. There were no patient age restrictions; however, for this analysis, investigators decided to report on studies with adults and adolescents (ie, aged ≥15 years) and to publish results for children separately due to the different tuberculosis case definitions used for children. There were no restrictions on the types of NAAT used. We excluded studies without consecutive, systematic, or random recruitment; studies where the ability to provide sputum or urine were an inclusion criterion; studies with less than 30 participants diagnosed with tuberculosis; studies evaluating early research assays without clearly defined cutoffs; and animal studies. After removing duplicates, two independent reviewers (TB and IDO) screened the titles and abstracts and subsequently the full texts to confirm eligibility, with any disagreements resolved by discussing with the primary authors of the study in question. Covidence and Excel were used to manage references and for the purpose of screening.

Data extraction, study quality, and processing

Study-level data were extracted independently by the two reviewers (TB and IDO), with disagreements resolved by consensus after discussing with the primary authors of the extracted studies. We extracted prespecified variables (appendix p 4). Risk of bias in primary studies was independently assessed by the two reviewers according to the Quality of Diagnostic Accuracy Studies-2 tool,[19] with disagreements resolved by consensus. We invited authors of eligible studies by email to contribute de-identified IPD (appendix pp 5–6). On receipt of IPD, we checked the number of participants, participants who were lipoarabinomannan-positive, and participants who were NAAT-positive against the original publication to confirm that we had received the full dataset and queried missing or inconsistent data. Duplicates were removed and data were cleaned, standardised, and pooled into one database using an R script. For AlereLAM, we used the manufacturer's threshold for test positivity: either the updated reference card with four bands (grade 1 of 4) or the corresponding previous reference card with five bands (grade 2 of 5). Participants without data for age, tuberculosis symptoms, sex, HIV status, or antiretroviral therapy (ART) were excluded.

Denominator and diagnostic yield

The main study outcomes—tuberculosis diagnostic yields of urine lipoarabinomannan, NAAT, and SSM from the first baseline diagnostic sample—were compared independently against a meta-analysis denominator (MAD). Diagnostic yields were calculated using simple proportions and predicted using random-effects and mixed-effects meta-analyses. We defined diagnostic yield (DY) as the proportion of tuberculosis cases identified by a single diagnostic test on the first diagnostic sample collected within 2 days of enrolment (PT) among those with a positive MAD:

$$DY = \frac{PT}{MAD} \times 100\%.$$

2 days was specifically chosen to allow for a second collection attempt the following day. Participants who were unable to provide samples for index tests (eg, urine or sputum) were still included in the analysis. The MAD was reconstructed and harmonised across all studies and included participants with microbiologically confirmed tuberculosis, defined as any culture (ie, liquid or solid) or any NAAT positive for *Mycobacterium tuberculosis* from any sample (eg, sputum, urine, blood, and other extrapulmonary samples). Culture and NAAT were included in the MAD due to their high specificity, and NAAT was included because culture was often not available in studies that assessed performance of tests in programmatic settings. In a secondary analysis, participants with a positive lipoarabinomannan test were added to the denominator (MAD–LAM). Clinically diagnosed tuberculosis was not considered in the denominator because harmonisation across studies was not feasible. More details on the denominators are shown in the appendix (p 7).

Pulmonary, Allergy, and Critical Care Medicine, Columbia University Irving Medical Center, New York, NY, USA (M J Cummings MD); **Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, NY, USA** (M J Cummings); **Centre for Healthcare-Associated Infections, Antimicrobial Resistance and Mycoses, National Institute for Communicable Diseases, Johannesburg, South Africa** (R M Wake PhD); **Institute for Infection and Immunity, St George's University of London, London, UK** (R M Wake); **The Kirby Institute, University of New South Wales, Sydney, NSW, Australia** (J Hanson FRACP); **Department of Medicine, Division of Pulmonary Diseases and Critical Care Medicine, University of California Irvine, Irvine, CA, USA** (A Cattamanchi)

Correspondence to:
Dr Claudia M Denkinger, Division of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, 69120 Heidelberg, Germany
claudia.denkinger@uni-heidelberg.de

See Online for appendix

Statistical analysis

The R package eulerr[20] was used to generate area-proportional Euler diagrams to show the number of positive test results by test type in MAD-positive participants. Furthermore, diagnostic yields were predicted using one-stage random-effects and mixed-effects meta-analyses. We followed a Bayesian approach to obtain posterior distributions for the diagnostic yields per study with Markov Chain Monte Carlo methods using the brms R package for all models.[21] We ensured model fit through residual analysis and posterior predictive checks. The overall diagnostic yield posterior distribution was summarised to compute the mean diagnostic yield and its 95% credible intervals (95% CrI).

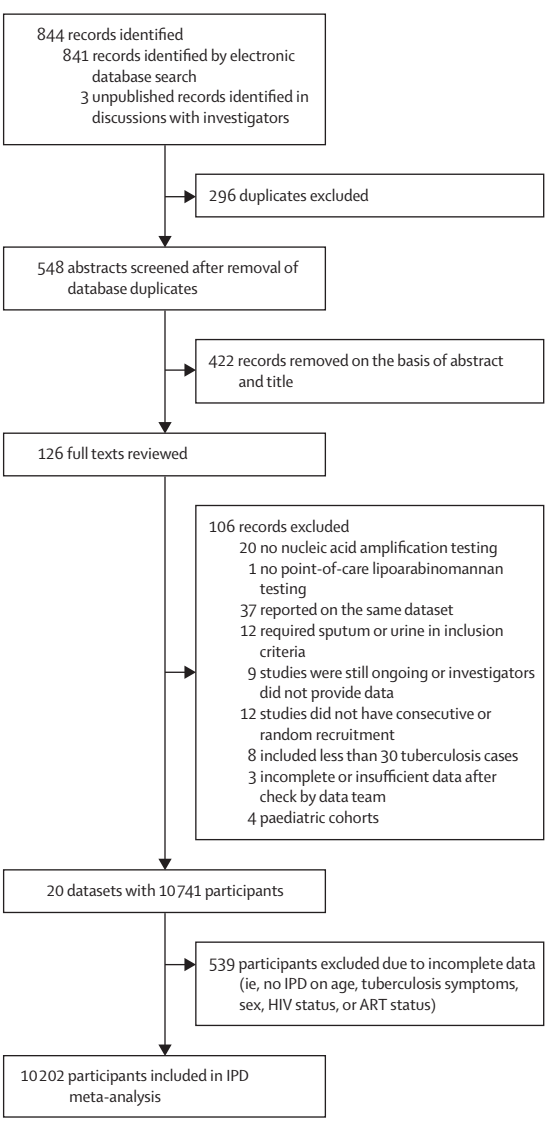

Figure 1: **PRISMA diagram of studies included in this meta-analysis and reasons for exclusion**
ART=antiretroviral therapy. IPD=individual participant data.

For the summary overall diagnostic yield prediction, we included one random effect to account solely for heterogeneity at the study level.

To predict diagnostic yield in subgroups and assess the sources of heterogeneity, we extended the model by adding fixed effects resulting in a multivariable generalised linear mixed model. The generalised linear mixed model included fixed effects that accounted for population effects (ie, age, sex, presence of tuberculosis symptoms [cough, fever, weight loss, or night sweats[22]], CD4 cell count, Xpert cartridge type [MTB/RIF or Ultra], recruitment setting [inpatients or outpatients], number of valid results from sputum Xpert or culture, antiretroviral therapy [ART] status, and random effects [study and country]; appendix p 8). For the generalised linear mixed model, missing CD4 cell counts were imputed per individual on the basis of sex, ART status, and setting through Monte Carlo sampling during the model inference. We inferred adjusted odds ratios (ORs) for all variables included in the models to explore the effects of the variable towards a positive test result. Next, we performed subgroup analyses for relevant variables through estimated marginal means, providing predicted weighted means of the diagnostic yield and 95% prediction intervals (95% PrIs) after accounting for all covariates. Relevant variables included those with a significant effect based on ORs and variables that were predefined in the analysis plan (ie, CD4 stratum, recruitment setting, and tuberculosis symptoms). Prespecified secondary analyses assessed the tuberculosis diagnostic yield of combinations of tests (ie, urine lipoarabinomannan with sputum NAAT and urine lipoarabinomannan with SSM). Remaining heterogeneity between the different studies was assessed by considering statistical relevance of the posterior distributions of the random effect for each study. If the 95% CrI did not intersect with 0, then we considered this study to vary more than only through random sampling error.

We further determined the proportion of participants who were able to provide a baseline urine sample within 2 days of enrolment and, separately, the proportion of participants who were able to provide a sputum sample within the same time period. Prespecified sensitivity analyses for the overall diagnostic yield were first performed using MAD–LAM, second using MAD and excluding studies that did not perform microbiological testing on samples other than sputum and urine, and third using MAD and excluding studies with a high or unclear overall risk of bias, by use of the generalised linear mixed model.

All analyses were done in R, version 4.0.4. The meta-analysis protocol was registered with PROSPERO (CRD42021230337) and the prospectively defined statistical analysis plan is shown in the appendix (pp 31–38). Our findings are reported in accordance with PRISMA–Individual Patient Data and PRISMA–Diagnostic Test Accuracy statements (appendix pp 27–30).[23,24] The primary

	All participants (n=10 202)
Country	
Guatemala	295 (3%)
Kenya	867 (8%)
Malawi	1406 (14%)
Mozambique	1276 (13%)
Myanmar	517 (5%)
South Africa	3472 (34%)
Tanzania	205 (2%)
Uganda	610 (6%)
Zambia	936 (9%)
Zimbabwe	618 (6%)
WHO region	
African region	9390 (92%)
Region of the Americas	295 (3%)
South-East Asia region	517 (5%)
Recruitment setting	
Inpatients	3662 (36%)
Outpatients	6540 (64%)
Age, years	36 (30–44)
Sex	
Male	4561 (45%)
Female	5641 (55%)
HIV-positive	10 202 (100%)
CD4 count, cells per µL	187 (66–365)
CD4 count group	
≤100 cells per µL	3138 (31%)
101–200 cells per µL	1797 (18%)
>200 cells per µL	4502 (44%)
Unknown	765 (7%)
On antiretroviral therapy	4716 (46%)
Previous tuberculosis	1784 (17%)
Tuberculosis symptoms	8525 (84%)
Number of valid sputum Xpert and sputum culture results	
0	1542 (15%)
1	2868 (28%)
2	3431 (34%)
>2	2361 (23%)
Positive tuberculosis results	
MAD*	1615 (16%)
MAD–LAM†	2531 (25%)
Original study reference standard	1791 (18%)
Sample available in the first 2 days	
Urine	9957 (98%)
Sputum	8360 (82%)
Induced	277 (3%)
Spontaneously expectorated	5284 (63%)
Unknown	2799 (33%)
Positive test in the first 2 days	
Urine AlereLAM	1550 (15%)
Sputum Xpert	982 (10%)
Sputum smear microscopy	490 (5%)

Data are n (%) or median (IQR). AlereLAM=Alere Determine TB LAM Ag assay. LAM=lipoarabinomannan. MAD=meta-analysis denominator. Xpert=Xpert MTB/RIF or Xpert Ultra assay. *Number of positive participants as defined by the harmonised MAD based on microbiologically confirmed tuberculosis. †MAD–LAM is the number of positive participants as defined by the MAD based on microbiologically confirmed tuberculosis, including participants with a positive AlereLAM test in the denominator.

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

studies all had ethics approval and this IPD meta-analysis was approved by the ethics committee of the Medical Faculty Heidelberg (S-260/2022).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Of the 844 records identified, 20 datasets of 10 741 people were eligible for inclusion (figure 1; appendix p 9; NCT03187964):[8,9,25–42] 14 cohort studies, three cross-sectional studies, and three randomised controlled trials. All studies enrolled people living with HIV. 11 studies enrolled outpatients, five enrolled inpatients, and four enrolled both inpatients and outpatients. Ten studies enrolled participants with tuberculosis symptoms and the other ten enrolled unselected participants irrespective of tuberculosis symptoms. The risk of bias was generally low, with 15 of 20 studies having a low overall risk of bias (appendix pp 10–11). All studies used AlereLAM as the lipoarabinomannan test and sputum Xpert as the NAAT (17 Xpert MTB/RIF and 3 Xpert Ultra; participants who were trace-positive in Xpert Ultra studies were classified as having tuberculosis). 16 studies also performed SSM and one study also performed Fujifilm Silvamp TB LAM (FujiLAM, Fujifilm, Odawara, Japan).

Records for 10 741 participants were obtained (accounting for all 20 eligible datasets) and 10 202 records (from 4561 [45%] male participants and 5641 [55%] female participants) were included in the analyses after IPD harmonisation (figure 1). The dataset contains IPD from three continents, but 9390 (92%) of 10 202 participants were from sub-Saharan Africa (table 1). CD4 cell count was available for 9437 (93%) participants. The mean number of valid sputum Xpert and sputum culture results per participant was 1·8 (SD 1·2). 50% (ten of 20) of studies performed Xpert or culture testing on non-sputum samples, and microbiological confirmation was exclusively based on non-sputum samples in 124 (8%) of 1615 MAD-positive participants.

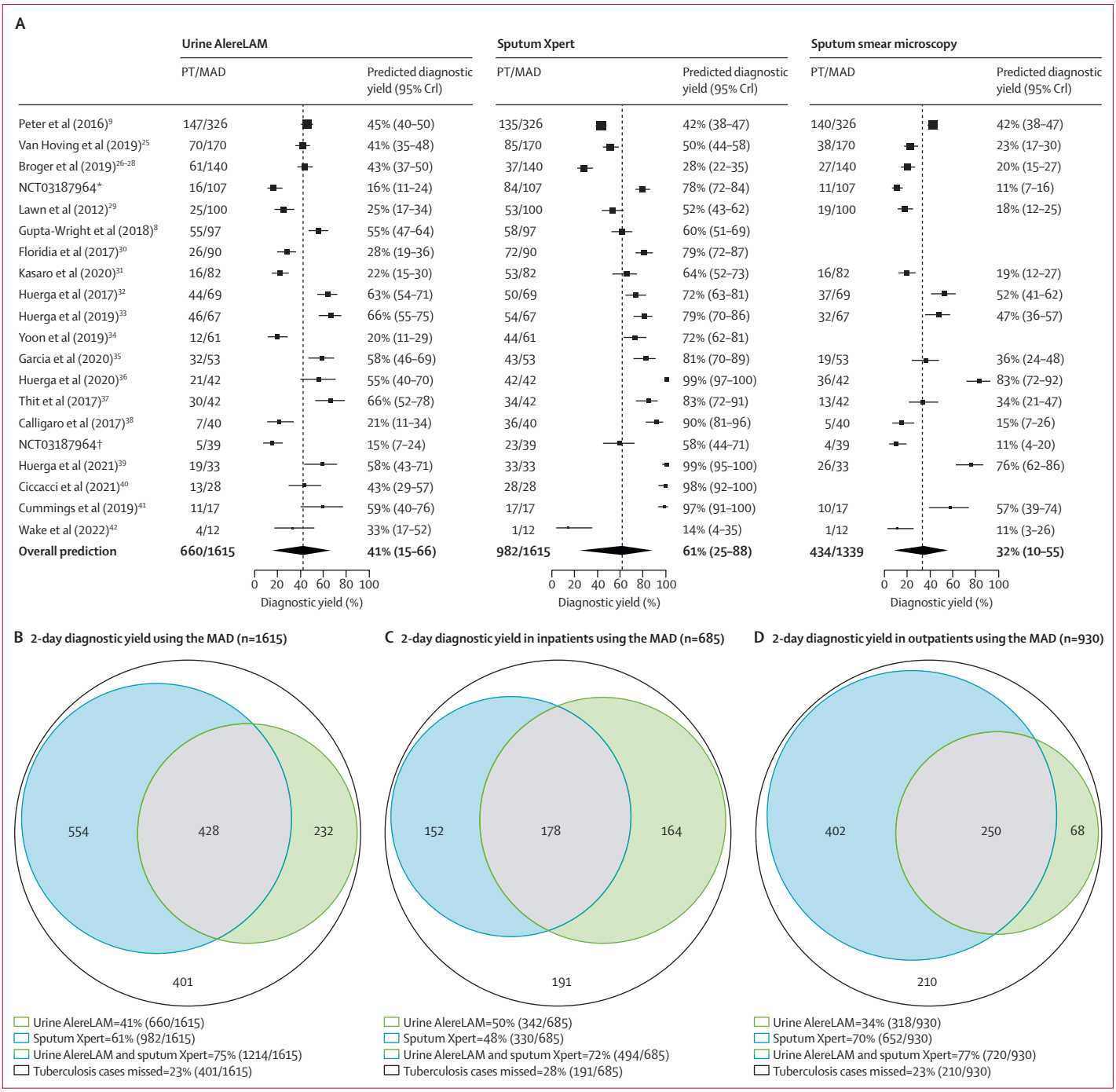

***Figure 2*: Diagnostic yields of urine AlereLAM, sputum Xpert, and sputum smear microscopy**

(A) Forest plots per study and test and overall prediction of diagnostic yield. Squares represent predicted diagnostic yields and whiskers represent 95% CrI. The size of the square is proportional to the number of participants with MAD-positive tuberculosis in each study, and studies are sorted by size. The vertical dashed lines indicate the overall predicted mean by the random effects model and the diamond represents the 95% CrI around that prediction. The PT/MAD represents the number of participants with a positive test from the first sample collected in the initial 2 days after enrolment divided by the number of positive participants as defined by the harmonised MAD based on microbiologically confirmed tuberculosis. Euler diagrams and proportion of positive urine AlereLAM and sputum Xpert test results and their overlap in all participants who were MAD-positive (B), inpatients who were MAD-positive (C), and outpatients who were MAD-positive (D). A Euler diagram for the subset of studies that performed all three tests, including sputum smear microscopy, is shown in the appendix (p 39). 95% CrI=95% credible interval. AlereLAM=Alere Determine TB LAM Ag assay. MAD=meta-analysis denominator. PT=positive tests. Xpert=Xpert MTB/RIF or Xpert Ultra assay. *Data are from the Kraaifontein Tuberculosis substudy. †Data are from the Antiretroviral Therapy Tuberculosis substudy.

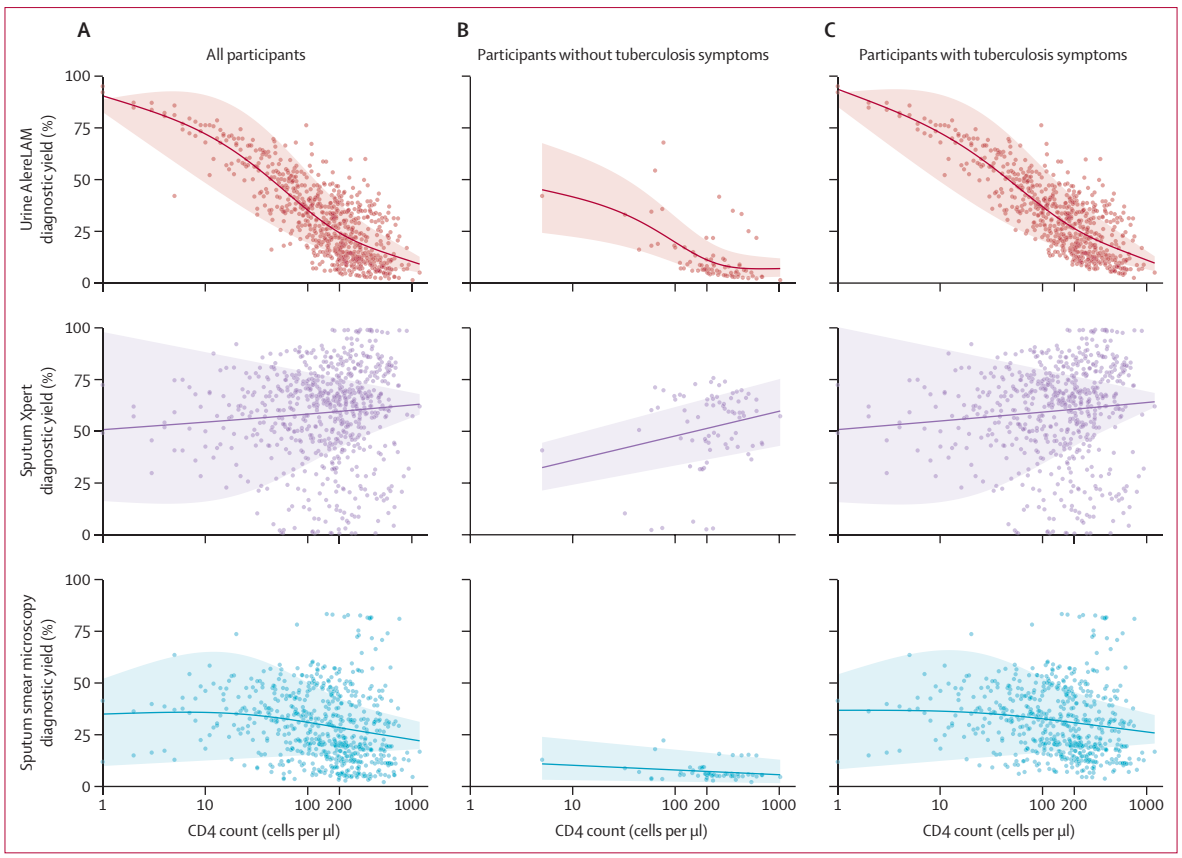

***Figure 3*: Diagnostic yield predictions as a function of CD4 cell count**

Diagnostic yield predictions in participants with tuberculosis based on the MAD in all participants (A), participants without tuberculosis symptoms (B), and participants with tuberculosis symptoms (C). We used the number of positive participants as defined by the harmonised MAD based on microbiologically confirmed tuberculosis as the denominator for diagnostic yield. Solid lines represent mean predictions, dashed lines represent 95% prediction intervals, and dots represent the participant data. AlereLAM=Alere Determine TB LAM Ag assay. MAD=meta-analysis denominator. Xpert=Xpert MTB/RIF or Xpert Ultra assay.

Within the first 2 days, nearly all participants provided a urine sample, with fewer participants providing a sputum sample (table 1). The ability to provide sputum differed by setting (5829 [89%] of 6540 outpatients *vs* 2531 [69%] of 3662 inpatients), and 2452 (78%) of 3138 participants with a CD4 count of 100 cells per µL or less provided sputum samples. By contrast, the ability to provide urine exceeded 96% in all subgroups (appendix p 12). In five studies enrolling inpatients irrespective of tuberculosis signs and symptoms, only 1084 (54%) of 1993 of people living with HIV were able to provide a sputum sample, whereas 1966 (99%) provided a urine sample within the first 2 days.

Pooled overall predicted diagnostic yield in people living with HIV was 41% (95% CrI 15–66) for urine AlereLAM, 61% (25–88) for sputum Xpert, and 32% (10–55) for SSM (figure 2A). AlereLAM and Xpert, in combination, detected 75% (1214 of 1615) of all participants with tuberculosis in the first 2 days (figure 2B). In inpatients, the diagnostic yields were 50% (342 of 685) for AlereLAM, 48% (330) for Xpert, and 72% (494) for AlereLAM plus Xpert (figure 2C). In outpatients, the diagnostic yields were 34% (318 of 930) for AlereLAM, 70% (652) for Xpert, and 77% (720) for AlereLAM plus Xpert (figure 2D).

For all tests, there was a large degree of uncertainty regarding their diagnostic yields across studies (figure 2A), suggesting substantial heterogeneity across studies with respect to relevant subgroups (eg, CD4 cell count, presence of symptoms, and severity of disease as inferred from inpatient or outpatient status). We examined the effect of these variables in subsequent analyses.

The analysis of variable effect on the diagnostic yield identified a significant effect of CD4 cell count and tuberculosis symptoms, causing multimodality in the posterior distribution of the modelled diagnostic yield and thus large 95% CrIs (appendix p 13). Figure 3 shows predicted diagnostic yields for AlereLAM, Xpert, and SSM as a function of CD4 cell count for participants with and without tuberculosis symptoms. For AlereLAM, diagnostic yield increased with lower CD4 cell count in people living with HIV (OR 3·47, 95% CrI 2·77–4·36, per 200 cells per µL CD4 count decrease; table 2; appendix p 14). The predicted diagnostic yield for

	Urine AlereLAM				Sputum Xpert				SSM			AlereLAM + Xpert			AlereLAM + SSM		
	WHO recommendation	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)	WHO recommendation	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)
Tuberculosis symptoms																	
Any setting																	
Any CD4 cell count	Recommended	651	1538	42% (7–82)	Recommended	943	1538	60% (1–100)	431	1283	34% (5–83)	1173	1538	76% (37–99)	721	1283	53% (14–91)
>200	Recommended	61	334	19% (5–51)	Recommended	213	334	63% (6–100)	86	281	30% (6–85)	226	334	68% (31–99)	109	281	36% (10–85)
≤200	Recommended	546	1120	48% (13–89)	Recommended	666	1120	59% (1–100)	325	938	35% (5–84)	872	1120	78% (42–99)	570	938	58% (20–93)
≤100	Recommended	450	794	56% (24–84)	Recommended	468	794	59% (1–100)	235	675	35% (5–82)	646	794	81% (52–99)	443	675	64% (28–92)
Inpatients																	
Any CD4 cell count	Recommended	340	678	49% (12–83)	Recommended	323	678	47% (1–100)	236	581	40% (7–79)	492	678	71% (32–99)	366	581	61% (26–91)
>200	Recommended	20	115	24% (8–49)	Recommended	53	115	46% (1–99)	33	95	36% (6–79)	59	115	58% (26–69)	39	95	43% (20–78)
≤200	Recommended	296	523	56% (19–90)	Recommended	245	523	47% (1–99)	185	446	41% (7–79)	396	523	74% (39–99)	297	446	66% (34–92)
≤100	Recommended	247	393	62% (37–85)	Recommended	180	393	47% (1–100)	145	336	42% (7–80)	309	393	77% (50–99)	239	336	70% (46–92)
Outpatients																	
Any CD4 cell count	Recommended	311	860	36% (6–80)	Recommended	614	860	71% (22–100)	195	702	28% (4–86)	681	860	78% (45–99)	355	702	47% (12–90)
>200	Recommended	41	219	17% (4–52)	Recommended	150	219	72% (30–100)	53	186	28% (6–87)	167	219	72% (42–99)	70	186	33% (10–87)
≤200	Recommended	250	597	42% (11–87)	Recommended	421	597	70% (19–100)	140	492	29% (4–86)	476	597	81% (51–99)	273	492	52% (18–93)
≤100	Recommended	203	401	51% (21–84)	Recommended	288	401	70% (19–100)	90	339	28% (3–85)	337	401	85% (62–99)	204	339	58% (25–93)
Unselected, symptoms not assessed*																	
Any setting																	
Any CD4 cell count	NA	253	721	35% (3–81)	NA	439	721	61% (6–99)	101	473	22% (3–80)	525	721	72% (32–99)	200	473	42% (6–89)
>200	NA	29	212	16% (2–52)	NA	125	212	60% (6–98)	26	151	19% (2–76)	130	212	64% (30–97)	36	151	26% (4–75)
≤200	NA	220	500	42% (8–90)	NA	307	500	61% (6–99)	75	314	24% (3–81)	387	500	76% (32–99)	161	314	48% (12–92)
≤100	Recommended	176	331	52% (20–85)	NA	202	331	61% (5–100)	49	210	25% (4–82)	270	331	81% (40–99)	121	210	56% (23–92)

(Table 2 continues on next page)

	Urine AlereLAM				Sputum Xpert				SSM			AlereLAM + Xpert			AlereLAM + SSM		
	WHO recommendation	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)	WHO recommendation	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)	Positive tests, n	MAD positive tests, n	2-day diagnostic yield (95% PrI)
(Continued from previous page)																	
Inpatients																	
Any CD4 cell count	NA	143	281	51% (15–82)	NA	133	281	47% (4–100)	57	184	31% (4–84)	203	281	71% (31–100)	110	184	59% (27–93)
>200	NA	10	58	26% (11–48)	NA	26	58	43% (4–100)	9	38	30% (4–84)	28	58	60% (25–99)	12	38	43% (22–82)
≤200	Recommended	133	222	57% (20–93)	NA	106	222	48% (4–100)	48	145	31% (4–85)	174	222	75% (35–100)	98	145	63% (33–94)
≤100	Recommended	108	159	65% (40–85)	NA	70	159	49% (4–100)	32	102	32% (4–85)	127	159	79% (39–100)	74	102	69% (38–94)
Outpatients																	
Any CD4 cell count	Not recommended	110	440	24% (3–74)	NA	306	440	70% (26–92)	44	289	16% (3–41)	322	440	73% (34–97)	90	289	30% (5–79)
>200	Not recommended	19	154	12% (2–53)	NA	99	154	67% (33–90)	17	113	15% (2–39)	102	154	65% (34–95)	24	113	19% (4–68)
≤200	Not recommended	87	278	31% (7–77)	NA	201	278	71% (19–92)	27	169	18% (3–43)	213	278	77% (30–97)	63	169	36% (10–81)
≤100	Recommended	68	172	39% (18–85)	NA	132	172	73% (8–93)	17	108	18% (3–44)	149	172	82% (40–98)	47	108	43% (20–87)
No tuberculosis symptoms																	
Any setting																	
Any CD4 cell count	NA	9	77	13% (2–61)	NA	39	77	52% (3–82)	3	56	8% (2–25)	41	77	55% (20–91)	7	56	13% (2–61)
Inpatients																	
Any CD4 cell count	NA	2	7	39% (9–88)	NA	1	7	15% (2–82)	1	7	8% (1–31)	2	7	40% (12–96)	2	7	32% (11–74)
Outpatients																	
Any CD4 cell count	NA	7	70	10% (2–45)	NA	38	70	56% (25–82)	2	49	8% (2–24)	39	70	56% (25–90)	5	49	13% (3–49)

Positive tests are the number of positive tests from the first sample collected in the initial 2 days after enrolment. MAD positive tests are the number of positive tests as defined by the harmonised MAD, based on microbiologically confirmed tuberculosis. NA is shown where there was no existing recommendation. AlereLAM=Alere Determine TB LAM Ag assay. MAD=meta-analysis denominator. NA=not applicable. PrI=prediction interval. Xpert=Xpert MTB/RIF or Xpert Ultra assay. *Includes all participants from studies that enrolled participants irrespective of signs and symptoms of tuberculosis.

Table 2: **Predicted diagnostic yields for tuberculosis tests and test combinations in various clinical scenarios and among subgroups**

AlereLAM was 34% among outpatients and 50% among inpatients (OR 0·80, 95% CrI 0·49–1·24). The predicted diagnostic yield for Xpert was 70% in the outpatient setting and 48% in the inpatient setting (OR 1·51, 95% CrI 0·74–2·61). All three tests had lower predicted diagnostic yields in participants without tuberculosis symptoms than in participants with tuberculosis symptoms (figure 3; table 2), but tuberculosis was detected in only 77 participants who were asymptomatic. CD4 cell count, age, ART status, and the Xpert cartridge type (ie, Xpert MTB/RIF *vs* Xpert Ultra) had no significant effect on Xpert diagnostic yield (appendix p 13). Only 124 participants had tuberculosis detected by Xpert Ultra (76% yield) compared with 858 with Xpert MTB/RIF (59% yield), and the OR for yield in Xpert Ultra versus Xpert MTB/RIF was 1·51 (0·33–4·41).

Among inpatients, AlereLAM and Xpert had similar diagnostic yields (table 2). AlereLAM had higher diagnostic yield than Xpert at low CD4 counts (both ≤200 and ≤100 cells per μL) in both symptomatic and unselected inpatients, whereas Xpert performed better in inpatients with CD4 counts above 200 cells per μL (table 2). When used in combination, Xpert and AlereLAM had higher diagnostic yields than did the individual tests alone.

In outpatients, the diagnostic yield of AlereLAM was low compared with Xpert in symptomatic and unselected participants. The combination of Xpert and AlereLAM diagnosed 78% of symptomatic outpatients, but the incremental increase in diagnostic yield when adding AlereLAM to Xpert was small—7% among people living with HIV who were symptomatic and 3% among unselected people living with HIV.

SSM diagnostic yield was below 42% in all scenarios, settings, and across CD4 strata. SSM diagnostic yield was influenced by test method: of the 16 studies that performed SSM, seven used fluorescence microscopy, two used conventional Ziehl-Neelsen microscopy, and seven did not specify. Fluorescence microscopy showed consistent diagnostic yields in the inpatient (39%) and the outpatient (41%) setting. Studies that used Ziehl-Neelsen microscopy showed lower SSM diagnostic yield than did studies that used fluorescence microscopy (12% in the outpatient setting; appendix p 21).

Only one study[26] evaluated for tuberculosis using urine FujiLAM. FujiLAM diagnostic yield was 65% (91 of 140, 95% CrI 57–72), which was 24% higher than AlereLAM and 4% higher than Xpert overall estimates (appendix p 15). FujiLAM yield increased in participants with tuberculosis symptoms and those with decreasing CD4 cell counts, with approximately 20% higher yields than AlereLAM (appendix p 16).

Eight of 20 studies included AlereLAM in their primary study reference standard definition. In a prespecified sensitivity analysis across all 20 studies using the MAD–LAM, the predicted 2-day diagnostic yield of AlereLAM was 61%, that of Xpert was 39%, and that of SSM was 22% (appendix p 19). An additional post-hoc analysis of diagnostic yield, using MAD–LAM as the denominator and adjusting yield for false positives, resulted in similar diagnostic yields of 59% for AlereLAM, 38% for Xpert, and 21% for SSM (appendix p 20). In the subgroup of participants who were unable to produce sputum in the first 2 days, AlereLAM detected 78% (359 of 462) of people with tuberculosis using MAD–LAM as the denominator. Overall, 28% (99 of 359) of those with an AlereLAM positive test were microbiologically confirmed with a positive Xpert or culture on at least one sample from any body site in the subsequent diagnostic workup, showing the challenges of confirming tuberculosis in people living with HIV who might struggle to provide a sputum sample (appendix p 22). Despite the incomplete tuberculosis microbiology due to the difficulties of getting samples, data from a subset of the studies show that antituberculosis therapy was initiated in 74% (176 of 238) participants with a positive urine AlereLAM result who were unable to produce sputum in the first 2 days.

Two additional prespecified sensitivity analyses, after excluding studies that did not perform microbiological testing on samples other than sputum and urine and studies with high or unclear risk of bias, both showed similar overall diagnostic yields to the primary analysis using MAD (appendix p 23). A comparison of diagnostic yield to calculated diagnostic yields based on data from previous meta-analyses is shown in the appendix (p 24).

Discussion

This IPD meta-analysis in adolescents and adults (ie, aged ≥15 years) showed that urine was obtainable in nearly all people living with HIV, but that nearly a fifth of participants were unable to provide sputum within 2 days of enrolment. Thus, the diagnostic yields of sputum Xpert (61%) and SSM (32%) were lower than their sensitivities. By contrast, the diagnostic yield of urine AlereLAM (41%) was unaffected by sample provision as samples were readily obtained from almost all people living with HIV.

The considerable heterogeneity in diagnostic yield observed within and across studies with respect to CD4 cell count, presence of tuberculosis symptoms, and clinical setting is unsurprising given the known heterogeneity in sensitivity estimates as reported in previous meta-analyses. However, combining our findings on sample provision with sensitivity estimates from previous meta-analyses[12–16,43] would result in diagnostic yields of 62–66% for Xpert, 35–43% for SSM, and 42% for AlereLAM, which are all similar to our reported results.

AlereLAM diagnostic yield was highest in participants with a CD4 count below 100 cells per μL and in inpatients. Our finding of increased diagnostic yield at low CD4 cell count is in line with earlier studies,[12,44] showing that lipoarabinomannan positivity is associated with total body mycobacterial load[45–47] and disseminated

tuberculosis, which are more likely in immunocompromised people living with HIV compared with immunocompetent people without HIV.[48] As previously reported by Dhana and colleagues,[49] we observed that AlereLAM 2-day diagnostic yield was similar to that of sputum Xpert among unselected inpatients. Lipoarabinomannan testing should therefore be considered a priority in the inpatient setting, as disseminated tuberculosis is common[48,50,51] and people who test positive for urine lipoarabinomannan have a higher mortality risk compared with people who test negative.[52–54] The results also support follow-up Xpert testing when AlereLAM results are negative, as this combination increased the 2-day diagnostic yield from 51% to 71%.

In the outpatient setting, Xpert clearly outperformed AlereLAM. However, AlereLAM still had an incremental yield in participants with tuberculosis symptoms, detecting 78% of tuberculosis cases when combined with Xpert testing in outpatients. Rapid point-of-care Xpert testing is often not available, and Xpert results take several days to come back to care providers. Therefore, AlereLAM should still be considered as a urine-based point-of-care diagnostic option to inform rapid treatment until Xpert results are available and in settings with no access to Xpert. SSM is still widely used for tuberculosis diagnosis in outpatients[55] and relevant as a comparator in WHO's prequalification of tuberculosis tests.[56] SSM had an overall diagnostic yield of only 34% in people with tuberculosis symptoms and performed poorly in all subgroups and clinical settings, never exceeding 42%. As a result, and in line with the WHO recommendation,[57] NAATs and urine lipoarabinomannan should be used instead of SSM for tuberculosis diagnosis in people living with HIV.

Taken together, our results suggest that the current WHO lipoarabinomannan guidelines[57] could be simplified and extended to prioritise lipoarabinomannan testing to diagnose tuberculosis in all inpatients who are HIV-positive and to aid in tuberculosis diagnosis in outpatients who are HIV-positive with tuberculosis symptoms (panel). A multicountry survey assessed the reasons for low uptake of urine lipoarabinomannan testing, and a prominent reason was that the test is only for a small perceived population, and thus is not considered a priority to implement.[10] Therefore, a simpler and broader recommendation could result in increased adoption. The same survey also identified that budget limitations; scarcity of country-specific data; administrative hurdles, such as regulatory agency approval; and insufficient coordination between national tuberculosis and HIV programmes are important implementation barriers. Our results make a compelling argument for broad use of urine lipoarabinomannan testing in people living with HIV, and we urge donors, ministries of health, and implementors to address these hurdles.

False positives could result in an overestimation of diagnostic yield, but the specificity of Xpert is higher than 97%, which is sufficiently high to avoid overestimation.[57] Pooled specificities of AlereLAM were reported to be 91% in people living with HIV with tuberculosis symptoms and 95% in unselected people living with HIV who were not assessed for tuberculosis symptoms.[12] If 9% of AlereLAM positive results were false positives, then diagnostic yield would reduce from 41% to 37% in this study. However, AlereLAM specificities might be underestimated due to underdiagnosis by imperfect reference standards, particularly as many studies used sputum-based tests but not non-sputum-based tests to establish the reference standard.[58,59] The MAD based on microbiological confirmation of tuberculosis might have missed participants with tuberculosis, as the number and type of tests performed differed between studies, with half of the studies performing only Xpert and culture on sputum samples. This focus on sputum-based reference standard testing could have led to a greater underestimation of AlereLAM yield than Xpert yield, particularly in participants who were unable to produce sputum. By including AlereLAM in the denominator (MAD–LAM), we showed that AlereLAM yield increased to 61% and Xpert yield dropped to 39%. Yield for AlereLAM was high (59%) even after adjusting for false positives. AlereLAM was positive in 78% of participants who were unable to produce sputum in the first 2 days, 28% of whom had subsequent tuberculosis confirmation

***Panel:* Proposal for a simplified urine lipoarabinomannan guideline for tuberculosis testing in people living with HIV**

In inpatient settings

Use urine lipoarabinomannan testing for all people who are HIV-positive (regardless of tuberculosis symptoms and CD4 cell count).

In outpatient settings

Use urine lipoarabinomannan testing for all people who are HIV-positive with tuberculosis symptoms* (regardless of CD4 cell count) and people who are HIV-positive, irrespective of tuberculosis symptoms with a CD4 count of less than 100 cells per μL or who are seriously ill.†

In all settings

All people who are HIV-positive with a positive urine lipoarabinomannan result should start tuberculosis therapy. Along with urine lipoarabinomannan testing, additional tuberculosis and, as necessary, drug-resistance testing‡ should be conducted. A negative urine lipoarabinomannan test does not rule out tuberculosis.

*Pulmonary or extrapulmonary symptoms of tuberculosis. †Seriously ill is defined on the basis of four danger signs: respiratory rate of more than 30 breaths per min, temperature of more than 39°C, heart rate of more than 120 beats per min, and unable to walk unaided. ‡The need for drug-resistance testing depends on prevalence of drug-resistant tuberculosis.

and 74% initiated tuberculosis treatment, suggesting that many people with positive AlereLAM results who were unable to produce sputum did indeed have tuberculosis. Therefore, AlereLAM-based treatment initiation should be a priority, particularly in inpatients who are HIV-positive, as the reliance on a combination of sputum-based diagnosis and clinically guided empirical treatment leaves people at an unacceptably high risk of death from undiagnosed tuberculosis.[8]

AlereLAM does not meet the requirements of a broadly applicable non-sputum-based diagnostic test,[4] but this important diagnostic void could be filled by next-generation lipoarabinomannan tests. We identified only one cohort study[26,27] of the next-generation FujiLAM assay that satisfied inclusion criteria, and the manufacturer reported product modifications, suggesting that previously evaluated assays might vary from the final commercial product.[60] Nevertheless, the results showed the potential of a next-generation lipoarabinomannan test, with FujiLAM reaching 65% overall diagnostic yield in more than 400 people living with HIV who were admitted to hospital, regardless of CD4 cell count and tuberculosis symptoms—the highest of all tests, including sputum Xpert.[58,61] These results are similar to those from a large prospective study reporting 60% diagnostic yield for FujiLAM among ambulatory outpatients in four African countries.[62] Next-generation lipoarabinomannan tests thus have great potential to avert tuberculosis deaths and incident tuberculosis cases,[17] and their development should be prioritised.[6,63] However, despite their potential for rapid point-of-care diagnosis, lipoarabinomannan-based tuberculosis tests will still need to be done in conjunction with NAATs for drug-resistance testing. Improved tests on easily obtainable samples are needed, and detection of mycobacterial DNA in oral swabs,[64] exhaled breath aerosols,[65,66] blood,[67] and urine[68,69] could potentially be used for investigation of drug resistance.

This study highlights important aspects that should be considered in future evaluations of tuberculosis tests. Many primary studies had to be excluded as they selectively enrolled only participants able to produce sputum. Thus, the findings and conclusions of these studies are restricted to people who can produce sputum, potentially biasing accuracy results in favour of sputum-based tests and making assessment of the diagnostic yield impossible. We propose that future tuberculosis diagnostic studies, particularly those evaluating non-sputum-based diagnostic tests, include participants regardless of ability to produce sputum and that they report on sample provision, diagnostic yield, and the composite yield of test algorithms.[70] We did not include children, but we will report on them in a separate analysis.

The strengths of this meta-analysis are its large size, with individual participant data from more than 10 000 people from three continents, including data from three randomised controlled trials. Furthermore, we carefully harmonised the denominator across 20 studies and used Bayesian random-effects and mixed-effects models to predict diagnostic yields in clinical scenarios and subgroups after accounting for potentially important confounders. Two denominators, one based on a widely accepted microbiological reference standard (MAD) and one combining MAD and lipoarabinomannan (MAD–LAM) were compared, and the influence of test specificity was evaluated in detail. However, our study has important limitations. Clinically diagnosed tuberculosis was not considered in the denominator. Most IPD were from Africa, with only one study from Asia and one from Central America. Little information was provided on the strategies and efforts used for obtaining sputum samples in the different studies, which might have influenced the availability of samples for testing. Information on sputum induction was missing for a third of participants, which might have biased the diagnostic yield assessment of the sputum-based tests and introduced heterogeneity. It is unclear whether AlereLAM yield is different if service-level staff perform the test, who might have little experience of the challenges of interpreting the results on the basis of the reference scale card and have a high workload.

The diagnostic yield of sputum-based tuberculosis tests is limited by people who cannot produce sputum, hampering diagnostic evaluation, whereas nearly all adults can provide urine. Urine AlereLAM had a similar diagnostic yield to sputum Xpert in inpatients who were HIV-positive. Furthermore, among outpatients, combined Xpert and AlereLAM testing can diagnose tuberculosis in more than three quarters of people living with HIV with tuberculosis symptoms. Therefore, guidelines should recommend prioritising lipoarabinomannan testing for tuberculosis diagnosis in all inpatients who are HIV-positive (irrespective of tuberculosis symptoms and regardless of CD4 cell count) and to aid in tuberculosis diagnosis in outpatients who are HIV-positive with tuberculosis symptoms. Next-generation lipoarabinomannan tests and other non-sputum-based assays could have broad usefulness in the fight against tuberculosis, and their development should be prioritised.

Contributors

TB designed the study and protocol and wrote the statistical analysis plan with assistance from LK, IDO, and CMD. CMD supervised the study. TB and IDO did the systematic review. HH, AG-W, AE, BWPR, MF, ADK, FC, MPK, JH, CY, DJVH, BS, JIG, MJC, RMW, and KD contributed data to the meta-analysis. TB and IDO merged and harmonised the IPD. TB, IDO, LK, and PM accessed the IPD and verified the data. LK and PM came up with the statistical method and analysed the data with assistance from TB, IDO, and CMD. TB, IDO, LK, and CMD wrote the first draft of the manuscript. All authors contributed to the interpretation of data and editing of the article and approved the final version of the manuscript. TB, IDO, LK, PM, and CMD had full access to all the data in the study, and all authors had final responsibility for the decision to submit for publication.

Declaration of interests

TB reports patent applications in the field of tuberculosis detection, reports consulting fees from the FINDdx, and is a shareholder of Avelo.

GMe was supported by the Wellcome Trust (214321/Z/18/Z and

203135/Z/16/Z) and the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa (grant number 64787). The opinions, findings, and conclusions expressed in this manuscript reflect those of the authors alone. All other authors declare no competing interests.

Data sharing

The aggregate datasets could not be made available due to data protection regulations. The study investigators of the original studies retain ownership of their data. Any requests for access to IPD should be made directly to study investigators of the original studies.

Acknowledgments

We thank the people that consented to participate in the primary studies. We thank Nandini Dendukuri for advice on Bayesian modelling and Miriam Compton for the manuscript review. We used ChatGPT for proofreading and shortening sections of text.

References

1 Pai M, Kasaeva T, Swaminathan S. Covid-19's devastating effect on tuberculosis care—a path to recovery. *N Engl J Med* 2022; **386:** 1490–93.

2 WHO. Global tuberculosis report 2022. 2022. https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022 (accessed April 1, 2023).

3 Gupta RK, Lucas SB, Fielding KL, Lawn SD. Prevalence of tuberculosis in post-mortem studies of HIV-infected adults and children in resource-limited settings: a systematic review and meta-analysis. *AIDS* 2015; **29:** 1987–2002.

4 Denkinger CM, Kik SV, Cirillo DM, et al. Defining the needs for next generation assays for tuberculosis. *J Infect Dis* 2015; **211** (suppl 2)**:** S29–38.

5 MacLean E, Broger T, Yerlikaya S, Fernandez-Carballo BL, Pai M, Denkinger CM. A systematic review of biomarkers to detect active tuberculosis. *Nat Microbiol* 2019; **4:** 748–58.

6 Bulterys MA, Wagner B, Redard-Jacot M, et al. Point-of-care urine LAM tests for tuberculosis diagnosis: a status update. *J Clin Med* 2019; **9:** 111.

7 WHO. Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV—policy update 2019. 2019. https://apps.who.int/iris/handle/10665/329479 (accessed April 1, 2023).

8 Gupta-Wright A, Corbett EL, van Oosterhout JJ, 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. *Lancet* 2018; **392:** 292–301.

9 Peter JG, Zijenah LS, Chanda D, 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. *Lancet* 2016; **387:** 1187–97.

10 Singhroy DN, MacLean E, Kohli M, et al. Adoption and uptake of the lateral flow urine LAM test in countries with high tuberculosis and HIV/AIDS burden: current landscape and barriers. *Gates Open Res* 2020; **4:** 24.

11 Kraef C, Yedilbayev A, Seguy N, et al. Uptake of the lateral flow urine LAM test in Europe and Central Asia. *Int J Tuberc Lung Dis* 2022; **26:** 835–41.

12 Bjerrum S, Schiller I, Dendukuri N, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in people living with HIV. *Cochrane Database Syst Rev* 2019; **10:** CD011420.

13 Horne DJ, Kohli M, Zifodya JS, et al. Xpert MTB/RIF and Xpert MTB/RIF Ultra for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst Rev* 2019; **6:** CD009593.

14 Steingart KR, Henry M, Ng V, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis* 2006; **6:** 570–81.

15 Zifodya JS, Kreniske JS, Schiller I, et al. Xpert Ultra versus Xpert MTB/RIF for pulmonary tuberculosis and rifampicin resistance in adults with presumptive pulmonary tuberculosis. *Cochrane Database Syst Rev* 2021; **2:** CD009593.

16 Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst Rev* 2014; **1:** CD009593.

17 Ricks S, Denkinger CM, Schumacher SG, Hallett TB, Arinaminpathy N. The potential impact of urine-LAM diagnostics on tuberculosis incidence and mortality: a modelling analysis. *PLoS Med* 2020; **17:** e1003466.

18 Ryckman TS, Dowdy DW, Kendall EA. Infectious and clinical tuberculosis trajectories: Bayesian modeling with case finding implications. *Proc Natl Acad Sci USA* 2022; **119:** e2211045119.

19 Whiting PF. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155:** 529–36.

20 Larsson J. eulerr: area-proportional Euler and Venn diagrams with ellipses. Dec 10, 2020. https://cran.r-project.org/package=eulerr (accessed April 1, 2023).

21 Bürkner P-C. Bayesian item response modeling in R with brms and Stan. *J Stat Softw* 2021; **100:** 1–54.

22 WHO. Guidelines for intensified tuberculosis case-finding and isoniazid preventive therapy for people living with HIV in resource-constrained settings. June 9, 2011. http://whqlibdoc.who.int/publications/2011/9789241500708_eng.pdf (accessed April 1, 2023).

23 McInnes MDF, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018; **319:** 388–96.

24 Stewart LA, Clarke M, Rovers M, et al. Preferred reporting items for a systematic review and meta-analysis of individual participant data. *JAMA* 2015; **313:** 1657.

25 Van Hoving DJDJ, Lahri S, Lategan HJ, Nicol MP, Maartens G, Meintjes G. Brief report: real-world performance and interobserver agreement of urine lipoarabinomannan in diagnosing HIV-associated tuberculosis in an emergency center. *J Acquir Immune Defic Syndr* 2019; **81:** e10–14.

26 Broger T, Sossen B, du Toit E, et al. Novel lipoarabinomannan point-of-care tuberculosis test for people with HIV: a diagnostic accuracy study. *Lancet Infect Dis* 2019; **19:** 852–61.

27 Lawn SD, Kerkhoff AD, Burton R, et al. Rapid microbiological screening for tuberculosis in HIV-positive patients on the first day of acute hospital admission by systematic testing of urine samples using Xpert MTB/RIF: a prospective cohort in South Africa. *BMC Med* 2015; **13:** 192.

28 Lawn SD, Kerkhoff AD, Burton R, et al. Diagnostic accuracy, incremental yield and prognostic value of Determine TB-LAM for routine diagnostic testing for tuberculosis in HIV-infected patients requiring acute hospital admission in South Africa: a prospective cohort. *BMC Med* 2017; **15:** 67.

29 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:** 201–09.

30 Floridia M, Ciccacci F, Andreotti M, et al. Tuberculosis case finding with combined rapid point-of-care assays (Xpert MTB/RIF and Determine TB LAM) in HIV-positive individuals starting antiretroviral therapy in Mozambique. *Clin Infect Dis* 2017; **65:** 1878–83.

31 Kasaro MP, Chilyabanyama ON, Shah NS, et al. Performance of Xpert MTB/RIF and Determine TB-LAM Ag in HIV-infected adults in peri-urban sites in Zambia. *Public Health Action* 2020; **10:** 134–40.

32 Huerga H, Ferlazzo G, Bevilacqua P, et al. Incremental yield of including determine-TB LAM assay in diagnostic algorithms for hospitalized and ambulatory HIV-positive patients in Kenya. *PLoS One* 2017; **12:** e0170976.

33 Huerga H, Rucker SCM, Cossa L, et al. Diagnostic value of the urine lipoarabinomannan assay in HIV-positive, ambulatory patients with CD4 below 200 cells/μl in 2 low-resource settings: a prospective observational study. *PLoS Med* 2019; **16:** e1002792.

34 Yoon C, Semitala FC, Asege L, et al. Yield and efficiency of novel intensified tuberculosis case-finding algorithms for people living with HIV. *Am J Respir Crit Care Med* 2019; **199:** 643–50.

35 García JI, Meléndez J, Álvarez R, et al. Accuracy of the tuberculosis point-of-care Alere determine lipoarabinomannan antigen diagnostic test using α-mannosidase treated and untreated urine in a cohort of people living with HIV in Guatemala. *AIDS Res Ther* 2020; **17:** 62.

36 Huerga H, Rucker SCM, Bastard M, et al. Should urine-LAM tests be used in TB symptomatic HIV-positive patients when no CD4 count is available? A prospective observational cohort study from Malawi. *J Acquir Immune Defic Syndr* 2020; **83:** e24–30.

37 Thit SS, Aung NM, Htet ZW, et al. The clinical utility of the urine-based lateral flow lipoarabinomannan assay in HIV-infected adults in Myanmar: an observational study. *BMC Med* 2017; **15:** 145.

38 Calligaro GL, Zijenah LS, Peter JG, et al. Effect of new tuberculosis diagnostic technologies on community-based intensified case finding: a multicentre randomised controlled trial. *Lancet Infect Dis* 2017; **17:** 441–50.

39 Huerga H, Mathabire Rucker SC, Bastard M, et al. Urine lipoarabinomannan (LAM) testing for all HIV patients hospitalized in the medical wards identifies a large proportion of patients with tuberculosis at risk of death. *Open Forum Infect Dis* 2021; **8:** ofaa639.

40 Ciccacci F, Welu B, Ndoi H, et al. High-sensitivity C-reactive protein in HIV care: tuberculosis diagnosis and short-term mortality in a cohort of Kenyan HIV patients in the DREAM programme. *Int J Infect Dis* 2021; **104:** 329–34.

41 Cummings MJ, Bakamutumaho B, Owor N, et al. Operational feasibility and diagnostic yield of urine TB-LAM testing among HIV-infected patients hospitalized with sepsis and septic shock in Uganda. *Am J Respir Crit Care Med* 2019; **199:** A5169.

42 Wake RM, Govender NP, Omar SV, et al. Rapid urine-based screening tests increase the yield of same-day tuberculosis diagnoses among patients living with advanced HIV disease. *AIDS* 2022; **36:** 839–44.

43 Chang EW, Page A-L, Bonnet M. Light-emitting diode fluorescence microscopy for tuberculosis diagnosis: a meta-analysis. *Eur Respir J* 2016; **47:** 929–37.

44 Shah M, Hanrahan C, Wang ZY, et al. Urine lateral flow lipoarabinomannan assay for diagnosing active tuberculosis in adults living with HIV. *Cochrane Database Syst Rev* 2014; **12:** CD011420.

45 Broger T, Nicol MP, Sigal GB, et al. Diagnostic accuracy of 3 urine lipoarabinomannan tuberculosis assays in HIV-negative outpatients. *J Clin Invest* 2020; **130:** 5756–64.

46 Paris L, Magni R, Zaidi F, et al. Urine lipoarabinomannan glycan in HIV-negative patients with pulmonary tuberculosis correlates with disease severity. *Sci Transl Med* 2017; **9:** eaal2807.

47 Cox JA, Lukande RL, Kalungi S, et al. Is urinary lipoarabinomannan the result of renal tuberculosis? Assessment of the renal histology in an autopsy cohort of Ugandan HIV-infected adults. *PLoS One* 2015; **10:** e0123323.

48 Kerkhoff AD, Sossen B, Schutz C, et al. Diagnostic sensitivity of SILVAMP TB-LAM (FujiLAM) point-of-care urine assay for extra-pulmonary tuberculosis in people living with HIV. *Eur Respir J* 2020; **55:** 1901259.

49 Dhana A, Hamada Y, Kengne AP, et al. Diagnostic accuracy of WHO screening criteria to guide lateral-flow lipoarabinomannan testing among HIV-positive inpatients: a systematic review and individual participant data meta-analysis. *J Infect* 2022; **85:** 40–48.

50 Barr DA, Lewis JM, Feasey N, et al. Mycobacterium tuberculosis bloodstream infection prevalence, diagnosis, and mortality risk in seriously ill adults with HIV: a systematic review and meta-analysis of individual patient data. *Lancet Infect Dis* 2020; **20:** 742–52.

51 Kerkhoff AD, Barr DA, Schutz C, et al. Disseminated tuberculosis among hospitalised HIV patients in South Africa: a common condition that can be rapidly diagnosed using urine-based assays. *Sci Rep* 2017; **7:** 10931.

52 Gupta-Wright A, Peters JA, Flach C, Lawn SD. Detection of lipoarabinomannan (LAM) in urine is an independent predictor of mortality risk in patients receiving treatment for HIV-associated tuberculosis in sub-Saharan Africa: a systematic review and meta-analysis. *BMC Med* 2016; **14:** 53.

53 Nathavitharana RR, Lederer P, Chaplin M, Bjerrum S, Steingart KR, Shah M. Impact of diagnostic strategies for tuberculosis using lateral flow urine lipoarabinomannan assay in people living with HIV. *Cochrane Database Syst Rev* 2021; **8:** CD014641.

54 Sossen B, Broger T, Kerkhoff AD, et al. "SILVAMP TB LAM" rapid urine tuberculosis test predicts mortality in patients hospitalized with human immunodeficiency virus in South Africa. *Clin Infect Dis* 2020; **71:** 1973–76.

55 Cazabon D, Pande T, Kik S, et al. Market penetration of Xpert MTB/RIF in high tuberculosis burden countries: a trend analysis from 2014 – 2016. *Gates Open Res* 2018; **2:** 35.

56 WHO. In vitro diagnostic medical devices used for the qualitative detection of Mycobacterium tuberculosis complex DNA and mutations associated with drug-resistant tuberculosis. 2022. https://apps.who.int/iris/handle/10665/366068 (accessed April 1, 2023).

57 WHO. Consolidated guidelines on tuberculosis—module 3: diagnosis. 2020. https://apps.who.int/iris/rest/bitstreams/1284627/retrieve (accessed April 1, 2023).

58 Broger T, Nicol MP, Székely R, 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. *PLOS Med* 2020; **17:** e1003113.

59 Lawn SD, Kerkhoff AD, Nicol MP, Meintjes G. Underestimation of the true specificity of the urine lipoarabinomannan point-of-care diagnostic assay for HIV-associated tuberculosis. *J Acquir Immune Defic Syndr* 2015; **69:** e144–46.

60 Székely R, Sossen B, Mukoka M, et al. Multicentre accuracy trial of FUJIFILM SILVAMP TB LAM test in people with HIV reveals lot variability. *medRxiv* 2022; published online Sept 9. https://doi.org/10.1101/2022.09.07.22278961 (preprint).

61 Li Z, Tong X, Liu S, Yue J, Fan H. The value of FujiLAM in the diagnosis of tuberculosis: a systematic review and meta-analysis. *Front Public Health* 2021; **9:** 757133.

62 Huerga H, Bastard M, Lubega AV, et al. Novel FujiLAM assay to detect tuberculosis in HIV-positive ambulatory patients in four African countries: a diagnostic accuracy study. *Lancet Glob Health* 2023; **11:** e126–35.

63 Treatment Action Group. Pipeline report 2022—tuberculosis diagnostics. November, 2022. https://www.treatmentactiongroup.org/resources/pipeline-report/2022-pipeline-report (accessed April 1, 2023).

64 Andama AO, Reza T, Jaganath D, et al. Accuracy of oral swab testing using Xpert MTB-RIF Ultra for tuberculosis diagnosis. *Am J Respir Crit Care Med* 2022; **205:** A4928.

65 Williams CM, Abdulwhhab M, Birring SS, et al. Exhaled Mycobacterium tuberculosis output and detection of subclinical disease by face-mask sampling: prospective observational studies. *Lancet Infect Dis* 2020; **20:** 607–17.

66 Dinkele R, Gessner S, McKerry A, et al. Aerosolization of Mycobacterium tuberculosis by tidal breathing. *Am J Respir Crit Care Med* 2022; **206:** 206–216.

67 Huang Z, LaCourse SM, Kay AW, et al. CRISPR detection of circulating cell-free Mycobacterium tuberculosis DNA in adults and children, including children with HIV: a molecular diagnostics study. *Lancet Microbe* 2022; **3:** e482–92.

68 Oreskovic A, Panpradist N, Marangu D, et al. Diagnosing pulmonary tuberculosis by using sequence-specific purification of urine cell-free DNA. *J Clin Microbiol* 2021; **59:** e0007421.

69 Fernández-Carballo BL, Broger T, Wyss R, Banaei N, Denkinger CM. Toward the development of a circulating free DNA-based in vitro diagnostic test for infectious diseases: a review of evidence for tuberculosis. *J Clin Microbiol* 2018; **57:** e01234–18.

70 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:** 1846–48.