THE LANCET
Global Health

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Broger T, Koeppel L, Huerga H, et al. Diagnostic yield of urine lipoarabinomannan and sputum tuberculosis tests in people living with HIV: a systematic review and meta-analysis of individual participant data. *Lancet Glob Health* 2023; **11:** e903–16.

Table of Contents

TBYield Study Consortium authors 2

Search terms 3

Study level variables 4

Patient level variables 5

Denominators 7

Multivariable generalized linear mixed model (GLMM) 8

Detailed primary study characteristics and proportion of missing data 9

Risk of bias assessment questionnaire based on QUADAS-2 and results 10

Risk of bias assessment 11

Two-day sputum and urine sample provision 12

Analysis of variable effect – odds ratios 13

Tuberculosis diagnostic yield predictions as a function of CD4 count and setting 14

FujiLAM Diagnostic yield prediction 15

FujiLAM Diagnostic yield prediction as a function of CD4 count and TB symptoms 16

GLMM model diagnostic yield predictions 17

Quantification of random effect introduced by study (“heterogeneity”) 18

Sensitivity analysis including LAM in the denominator (MAD–LAM) 19

Sensitivity analysis including LAM in the denominator (MAD–LAM) with adjustment for test specificity 20

Sputum smear microscopy (SSM) diagnostic yield for different microscopy methods 21

Urine AlereLAM diagnostic yield in patients unable to provide sputum 22

Overview of sensitivity analyses 23

Comparison of diagnostic yield to calculated diagnostic yields based on previous meta-analyses 24

Supplementary References 25

PRISMA-IPD Checklist 27

Statistical analysis plan 31

Two day diagnostic yields in the subset of studies that assessed SSM 39

TBYield Study Consortium authors

Suzette Oelofse	MBChB	Centre for Lung Infection and Immunity, Division of Pulmonology, Department of Medicine and UCT Lung Institute South African MRC Centre for the Study of Antimicrobial Resistance, University of Cape Town, Cape Town, South Africa.
Didier Laureillard	MD	Department of Infectious and Tropical Diseases, University Hospital, Nîmes, France Research unit "Pathogenesis and Control of Chronical and Emerging Infections", INSERM, French Blood Center, University of Montpellier, Montpellier, France
Mauro Andreotti	BSc	National Center for Global Health, Istituto Superiore di Sanità, Rome, Italy
Obvious Nchimunya Chilyabanyama	PhD	African Centre of Excellence in Data Science (ACEDS), University of Rwanda, Kigali, Rwanda Research Division, Centre for Infectious Disease Research in Zambia, Lusaka, Zambia
Benjamin Welu	PhD	Community of Sant'Egidio, DREAM programme, Meru, Kenya
Lucas Molfino	MSc	Department of Medicine, Médecins Sans Frontières, Maputo, Mozambique
Sekai Chenai Mathabire Rücker	MPH	Field Epidemiology Department, Epicentre, Chiradzulu, Malawi
Elisabeth Szumilin	MD	Department of Medicine, Médecins Sans Frontières, Paris, France
Loide Cossa	MD	Department of Medicine, Médecins Sans Frontières, Maputo, Mozambique
Johanna Meléndez	MD	Unidad de Atención Integral del VIH e Infecciones Crónicas “Dr. Carlos Rodolfo Mejía Villatoro”, Hospital Roosevelt, Guatemala City, Guatemala
Linda Mbuthini	MPH	Centre for Lung Infection and Immunity, Division of Pulmonology, Department of Medicine and UCT Lung Institute South African MRC Centre for the Study of Antimicrobial Resistance, University of Cape Town, Cape Town, South Africa.
Max O'Donnell	MPHMD	Division of Pulmonary, Allergy, and Critical Care Medicine, Columbia University Irving Medical Center, New York, New York Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York
Joseph N Jarvis	PhD	Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK
Gcobisa Ndlangalavu	PhD	DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research; South African Medical Research Council Centre for Tuberculosis Research; Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa.
Katherine Fielding	PhD	TB Centre, London School of Hygiene & Tropical Medicine, London, UK

Search terms

The search generally follows PICO: However, the search focusses on **o**utcome (Tuberculosis) and **i**ntervention (urine LAM tests) without restrictions regarding **p**articipants or **c**omparator to avoid missing relevant papers.

MEDLINE (Pubmed)

#1	Outcome	"Tuberculosis"[MeSH Terms] OR "Mycobacterium tuberculosis"[MeSH Terms] OR "Tuberculosis"[Title/Abstract] OR "TB"[Title/Abstract]
#2	Intervention Biomarker	((lipoarabinomannan[Title/Abstract]) OR (LAM[Title/Abstract]) OR (lipoarabinomannan[Supplementary Concept]))
#3	Intervention Test	("Alere"[Title] OR "AlereLAM"[Title] OR "Fujifilm"[Title] OR "FujiLAM"[Title] OR "test"[Title] OR "assay"[Title] OR "point of care"[Title] OR "point-of-care"[Title] OR "lateral flow"[Title] OR "LFA"[Title] OR "antigen"[Title] OR "Ag"[Title] OR "diagnostic*"[Title] OR "Silvamp"[Title] OR "determine"[Title]) OR (urin*[Title/Abstract])
#4		(#1) AND (#2) AND (#3)

Web of Science

#1	Outcome	TS=(Tuberculosis OR TB OR Mycobacterium tuberculosis)
#2	Intervention Biomarker	TS=(lipoarabinomannan OR LAM)
#3	Intervention Test	TI=(Alere OR AlereLAM OR Fujifilm OR FujiLAM OR test OR assay OR point of care OR point-of-care OR lateral flow OR LFA OR antigen OR Ag OR diagnostic* OR Silvamp OR determine) OR TS=(urin*)
#4		#1 AND #2 AND #3

EMBASE

#1	Outcome	tuberculsis OR 'mycobacterium tuberculosis':ti,ab
#2	Intervention Biomarker	lipoarabinomannan OR lam:ti,ab
#3	Intervention Test	urin*:ti,ab
#4		#1 AND #2 AND #3

ClinicalTrial.gov
Condition or Disease: Tuberculosis
Other terms: Lipoarabinomannan OR LAM

African Journals Online
Search Term: “Lipoarabinomannan or LAM”

Study level variables

- Study identification
    - First Author
    - Corresponding author and email
    - Journal
    - Title
    - DOI
    - Publication year
- Study details
    - Population age group (children <10 years, adolescent 10 to 18 years, adults ≥18 years, children & adolescent, adolescent & adults, children & adolescent & adults, other (specify))
    - Population details
    - CD4 range (≤100, ≤200, irrespective of CD4, other (specify)
    - Study design (RCT, cross-sectional, cohort, other (specify))
    - NCT (national clinical trial number)
    - Country(ies)
    - Number of sites
    - Clinical study setting (outpatient, inpatient, both (specify), unclear)
    - Inclusion criteria
    - Clinical status (symptomatic TB, asymptomatic, irrespective of TB symptoms)
    - Participant selection (consecutive, random, convenience, other (specify), unknown)
    - Timing (prospective, retrospective, unknown)
    - Sample size
        - Enrolled
        - For diagnostic accuracy
        - For diagnostic yield
    - Sputum sample type
    - Proportion of patients with non-induced sputum >50% (yes, no)
    - Sputum tests (SmearFluor; SmearZn; SmearUnspecified; Xpert; XpertUltra; MGIT; LJ; TLA; Other)
    - Xpert Version (G4, Ultra)
    - Other micro TB tests (bCulture, uCulture, uXpert, etc.)
    - Microbio on non-sputum samples (yes, no)
    - Chest X-ray (yes, no)
    - Urine LAM tests (AlereLAM; FujiLAM)
    - AlereLAM reference card and cut-off used (4LG1, 5LG2, Other, Unclear)
    - Specimens used LAM (fresh, frozen, unclear)
    - Index tests (availability and descriptions of methods for sputum smear microscopy, sputum NAATs, AlereLAM incl. cut-off, and FujiLAM, and/or other LAM tests)
    - Target condition (PTB, ETB, PTB&ETB, other (specify))
    - Reference standard definition used in study
    - Study results
        - AlereLAM (#Positive, Sensitivity, Specificity, diagnostic yield)
        - FujiLAM (#Positive, Sensitivity, Specificity)
        - Xpert (#Positive, Sensitivity, Specificity, diagnostic yield)
        - Oritinal study reference standard (#Positive, #Negative)
    - Denominator for dx yield definition
- Answers to QUADAS-2 signalling questions for risk of bias assessment (see below) for the four domains
    - Patient selection
    - Index test
    - Reference standard
    - Flow and timing

Patient level variables

Key variables [Exemplary coding is provided in brackets]

- Patient study ID
- Sputum results (for every available sputum specimen in case multiple specimen were collected)
    - Sputum specimen available [yes, no]
    - Sputum specimen collected in the first 48 hours? [yes / no / unknown]
    - Sputum induction [yes / no / unknown / no sample collected]
    - Smear microscopy result(s) [positive / negative / Unknown:TestfailedOrIndet / Unknown:NotDone / Unknown: OtherReason]
    - NAAT result(s) [positive / negative / Unknown:TestfailedOrIndet / Unknown:NotDone / Unknown:OtherReason]
    - Culture result(s) with confirmed Mtb [positive / negative / NTM / Unknown:Contaminated Unknown:TestfailedOrIndet / Unknown:NotDone / Unknown:OtherReason]
- Urine LAM results (for every available urine specimen in case multiple specimen were collected)
    - Urine specimen available [yes, no]
    - Urine specimen collected in the first 48 hours? [yes / no /unknown]
    - AlereLAM [positive / negative / Unknown:TestfailedOrIndet / Unknown:NotDone / Unknown:OtherReason] Important: Please mention the cut-off and AlereLAM reference card used in the codebook. AlereLAM resultsin this column should either use a
        - (a) grade 1 cut-off based on the 4-grade reference card (tests produced by Alere after January 2014) or
        - (b) grade 2 cut-off based on the 5-grade reference card. (tests produced by Alere before January 2014)
        - Please add grading in a separate, additional column if available
    - FujiLAM [positive / negative / Unknown:TestfailedOrIndet / Unknown:NotDone / Unknown:OtherReason]
    - Other LAM tests [positive / negative / no result]
- HIV status [positive/ negative / indeterminate / unknown]
- Recruitment setting [inpatient /outpatient /unknown]
- CD4 count [result in cells per ul / unknown]
- Clinical status [asymptomatic / symptomatic / unknown] (based on WHO symptom screen (2 weeks of cough or haemoptysis, weight loss, fever or night swats) or as per study definition. Please add the study definition to the codebook)
- Age [in years / unknown / if years can't be provided Age group in line with age groups from this study is acceptable, age groups are: children below 10 years, adolescent 10 to 18 years, adults greater or equal to 18 years)]
- Sex [female/ male/ unknown]
- Reference standard group results of primary study: TB diagnosis used in your primary study [TB / Non-TB / Probably TB / or as per study definition / unknown] (Please use the groups as they are described in the published paper or provide us with definitions in the codebook)
- Patient included in main analysis group?: We request data for the full cohort (including patients unable to provide sputum or urine) for this diagnostic yield meta-analysis. Typically, the main analysis of your study was performed in an analysis subgroup. Please indicate in this column those included in your main analysis subgroup (e.g. for diagnostic accuracy analysis). This will allow us to perform data checks [yes/no]

Additional variables based on availability [Exemplary coding is provided in brackets]

- On ART [yes / no / unknown]
- TB History [yes /no / unknown]
- Date of admission or study recruitment date [dd.mm.yy / unknown]
- Date of collection per sample [dd.mm.yy / not collected / unknown / missing sample]
- Additional test results
    - If available results from other reference standard definitions (microbiological reference standard (MRS), composite reference standard (CRS), etc.) [positive / negative / unknown]
    - If available other relevant results that were used for the definition of the TB diagnosis/denominator for diagnostic yield calculation:
        - Urine NAATs (e.g. Xpert MTB/RIF and/or GeneXpert MTB/RIF Ultra result(s)) [positive / negative / unknown]
        - Blood culture with confirmed Mtb [positive / negative / unknown / NTM / contaminated]

- Xpert or culture results from samples other than sputum [positive / negative / unknown]
- Patient follow-up results [positive / negative / unknown]
- Timepoint of follow-up visit
- X-ray
- TB treatment started [yes/no/unknown]
- Mortality outcome
- Timepoint of mortality outcome

Denominators

The meta-analysis denominator (MAD) is defined in the table below. LAM was not included in the denominator of the main analysis (MAD) but a sensitivity analysis was conducted with LAM in the denominator (MAD+LAM).

Meta-analysis denominator (MAD) (used for primary analysis)	**Meta-analysis denominator including LAM (MAD+LAM) (used for sensitivity analysis)**	**Original non-harmonized study reference standard (OSR) (used for descriptive overview only)**

Microbiologically confirmed TB was used as a denominator and defined as follows:
- Any culture (liquid or solid) positive for *Mtb* from any sample type or
- Any Xpert positive for *Mtb* from any sample type
 	
Samples include:
- Sputum
- Urine
- Blood
- Other extrapulmonary samples
  o Ascitic fluid
  o Bone marrow
  o Cerebrospinal fluid
  o Fine Needle aspirate
  o Lymph node aspirate
  o Gastric lavage
  o Bronchoalveolar lavage
  o Pus
  o Pleural fluid
  o Stool
 	
Only microbiologically confirmed TB was deemed TB positive. Possible TB or clinical definitions of TB were deemed negative because harmonisation across studies was not feasible due to methodological differences in the studies. For example some studies included X-ray, others follow-up data in the definition of clinical TB.
	Same definition as MAD (on the left) but in addition including participants with LAM positive urine sample(s).	The main reference standard that was used in the original study and provided by the primary study authors. Typically primary studies used a culture-based microbiological reference standard. Some studies included LAM in their reference standard definition.

Multivariable generalized linear mixed model (GLMM)

Let X_{ti} a set of Bernoulli distributed random variables, whereby *i =1,…n* represents the individual in the population of size *n* and *t* in {lama_48, sp_micro1_48, lamaSSM, bf_lamfSSM, sp_xpert1_48, lamaSpxpert, lamfSpxpert, lamf_48} the diagnostic test with success probability p_{ti}.

For *t* in { lama_48, sp_micro1_48, lamaSSM, bf_lamfSSM} we define the linear predictor

$$\text{logit}(p_{ti}) = \beta_{t0} + \beta_{t1} age_{ti} + \beta_{t2} \log(cd4)_{ti} + \beta_{t3} inout_{ti} + \beta_{t4} sx_{ti} + \beta_{t5} art_{ti} + \beta_{t6} sex_{ti} + \beta_{t7} nusp_{ti} + r\ tagctry_{ti},$$

for *t* in {sp_xpert1_48, lamaSpxpert, lamfSpxpert}

$$\text{logit}(p_{ti}) = \beta_{t0} + \beta_{t1} age_{ti} + \beta_{t2} \log(cd4)_{ti} + \beta_{t3} inout_{ti} + \beta_{t4} sx_{ti} + \beta_{t5} art_{ti} + \beta_{t6} sex_{ti} + \beta_{t7} nusp_{ti} + \beta_{t7} xpver_{ti} + r\ tagctry_{ti},$$

and for *t* = lamf_48

$$\text{logit}(p_{ti}) = \beta_{t0} + \beta_{t1} age_{ti} + \beta_{t2} \log(cd4)_{ti} + \beta_{t3} inout_{ti} + \beta_{t4} sx_{ti} + \beta_{t5} art_{ti} + \beta_{t6} sex_{ti} + \beta_{t7} nusp_{ti},$$

with coefficients β_{tj} and r_t, *j*=1,…7.

In the case of t in {sp1, u1} we further define

$$\text{logit}(p_{ti}) = \beta_{t0} + \beta_{t1} age_{ti} + \beta_{t2} \log(cd4)_{ti} + \beta_{t3} inout_{ti} + \beta_{t4} sx_{ti} + \beta_{t5} art_{ti} + \beta_{t6} sex_{ti} + r\ tagctry_{ti}.$$

We imputed missing cd4 counts (denoted by $cd4^*_{ti}$) via sampling from a normal distribution

$$cd4^*_{ti} \sim N(\propto_0 + \propto_1 art_{ti} + \propto_2 sex_{ti} + \propto_3 inout_{ti}, 1)$$

For all coefficients we chose standard Normal priors N(0,1).

Variable	Description
age	Age in years
art	On ART
cd4	CD4 count
inout	Recruitment Setting (inpatients, outpatients)
lama_48	Positive Alere LAM result for first urine available within 48 hours
lamf_48	Positive Fuji LAM result for first urine available within 48 hours
sex	sex
sp_micro1_48	Positive Smear microscopy for first sputum available within 48 hours
sp_xpert1_48	Positive Xpert result for first sputum available within 48 hours
sp1	First sputum available
sx	TB symptoms
u1	Urine sample available
xpver	Was Xpert MTB/RIF or Xpert MTB/RIF Ultra used?
nusp	Number of sputum Xpert and Cultures
Tagctry	Variable combining tag and country. Tag is the study and country the country
lamaSSM	lama_48 positive or sp_micro1_48 positive
lamfSSM	lamf_48 positive or sp_micro1_48 positive
lamaSpxpert	lama_48 positive or sp_xpert1_48 positive
lamfSpxpert	lamf_48 positive or sp_xpert1_48 positive

Detailed primary study characteristics and proportion of missing data

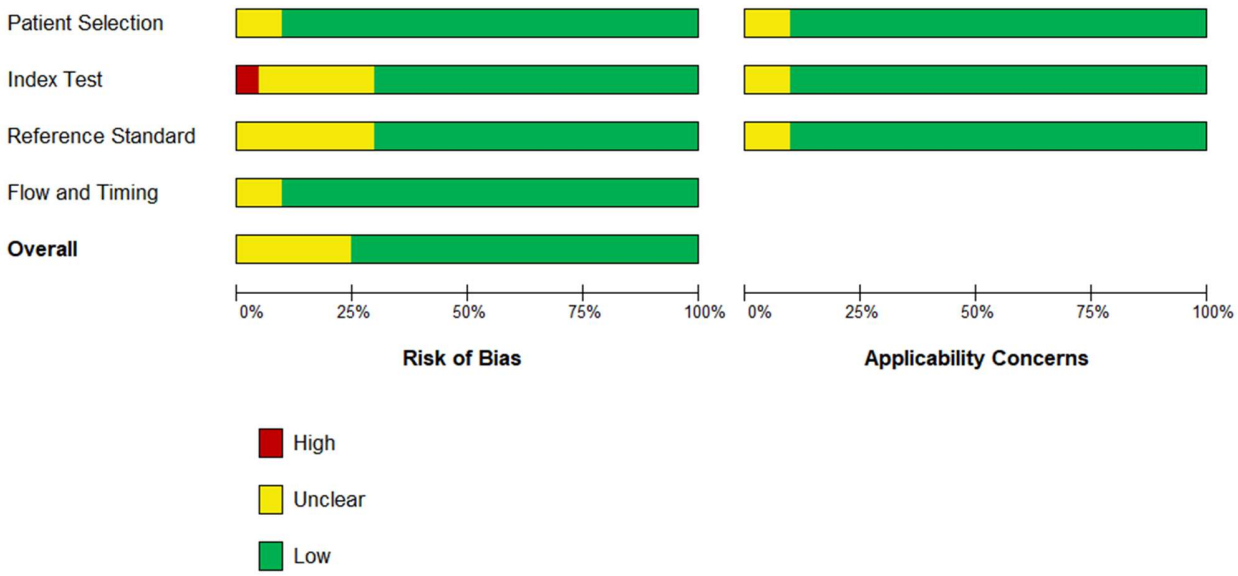
Study and acronym	Number of IPD	Countries	Design	Study population & inclusion			CD4 count		ART		TB symptom screen		TB reference standard						Specimen available in the first 48 hours				
				Recruitment setting	TB symptom	CD4 count	Median (IQR)	no data	On ART	no data	Sympto-matic	no data	MRS		MRS including LAM		Original study reference standard		Urine	Sputum			
													positive	unknown	positive	unknown	positive	unknown	Urine	all	induced	expecto-rated	no data
Broger et al (2019)[1]; Lawn et al (2015)[2]; Lawn et al. (2017)[3]	418	South Africa	cohort	inpatients	irrespective of symptoms	irrespective of CD4	149 (56, 313)	0%	42%	0%	90%	0%	33%	21%	41%	16%	33%	21%	98%	36%	1%	35%	64%
Calligaro et al (2017)[4], XACT	252	South Africa; Zimbabwe	RCT	outpatients	symptomatic TB	irrespective of CD4	234 (88, 422)	4%	45%	0%	100%	0%	16%	0%	20%	0%	12%	2%	99%	100%	12%	88%	0%
Ciccacci et al (2021)[5]	589	Kenya	cohort	outpatients	symptomatic TB	irrespective of CD4	477 (290, 710)	64%	100%	0%	100%	0%	5%	3%	15%	2%	15%	0%	100%	97%	0%	97%	3%
Cummings et al (2019)[6]	118	Uganda	cohort	inpatients	symptomatic TB	irrespective of CD4	NA (NA, NA)	100%	64%	0%	100%	0%	14%	41%	38%	29%	38%	2%	96%	62%	0%	62%	38%
Floridia et al (2017)[7], DREAM	997	Mozambique	cross-sectional	outpatients	irrespective of symptoms	irrespective of CD4	277 (142, 395)	0%	0%	0%	35%	0%	9%	0%	10%	0%	10%	0%	95%	95%	0%	0%	100%
Garcia et al (2020)[8]	295	Guatemala	cohort	inpatients and outpatients	symptomatic TB	irrespective of CD4	130 (48, 290)	1%	100%	0%	100%	0%	18%	20%	35%	11%	18%	20%	100%	79%	0%	0%	100%
Gupta-Wright et al (2018)[9], STAMP	1,093	Malawi; South Africa	RCT	inpatients	irrespective of symptoms	irrespective of CD4	246 (83, 471)	0%	85%	0%	90%	0%	9%	41%	16%	37%	16%	0%	99%	57%	0%	57%	43%
Huerga et al (2017)[10]	278	Kenya	cohort	inpatients and outpatients	symptomatic TB	≤200 cells/µl	109 (48, 216)	4%	96%	0%	100%	0%	25%	25%	47%	14%	29%	45%	100%	78%	4%	73%	23%
Huerga et al (2019)[11]	279	Malawi; Mozambique	cohort	outpatients	symptomatic TB	≤200 cells/µl	40 (16, 82)	0%	39%	0%	100%	0%	24%	16%	52%	10%	24%	16%	100%	79%	0%	79%	21%
Huerga et al (2020)[12]	481	Malawi	cohort	outpatients	symptomatic TB	irrespective of CD4	341 (131, 546)	0%	89%	0%	100%	0%	9%	36%	21%	31%	9%	0%	100%	65%	0%	65%	35%
Huerga et al (2021)[13]	372	Malawi	cohort	inpatients	irrespective of symptoms	irrespective of CD4	159 (49, 367)	2%	82%	0%	90%	0%	9%	38%	29%	28%	9%	38%	99%	62%	0%	62%	38%
Kasaro et al (2020)[14]	598	Zambia	cohort	outpatients	symptomatic TB	irrespective of CD4	298 (134, 453)	19%	40%	0%	100%	0%	14%	6%	16%	6%	12%	18%	86%	87%	0%	87%	13%
Lawn et al (2012)[15]	600	South Africa	cohort	outpatients	irrespective of symptoms	irrespective of CD4	168 (95, 231)	1%	0%	0%	87%	0%	17%	10%	19%	9%	16%	0%	99%	85%	10%	75%	16%
Peter et al (2016), LAMRCT[16]	1,257	South Africa; Zimbabwe; Zambia; Tanzania	RCT	inpatients	symptomatic TB	irrespective of CD4	81 (26, 198)	8%	48%	0%	100%	0%	26%	7%	36%	5%	27%	7%	96%	94%	3%	83%	14%
Theron et al (2021a, NCT03187964)[17]	740	South Africa	cohort	outpatients	irrespective of symptoms	irrespective of CD4	291 (166, 478)	2%	0%	0%	62%	0%	14%	1%	16%	1%	11%	7%	99%	99%	11%	20%	69%
Theron et al (2021b, NCT03187964)[17]	228	South Africa	cohort	outpatients	irrespective of symptoms	irrespective of CD4	325 (200, 489)	1%	0%	0%	31%	0%	17%	2%	18%	2%	15%	3%	100%	98%	0%	0%	100%
Thit et al (2017)[18]	517	Myanmar	cohort	inpatients and outpatients	irrespective of symptoms	irrespective of CD4	270 (129, 442)	0%	70%	0%	53%	0%	8%	0%	41%	0%	10%	0%	100%	100%	0%	0%	100%
Van Hoving et al (2019)[19]	417	South Africa	cross-sectional	outpatients	symptomatic TB	irrespective of CD4	84 (29, 209)	2%	48%	0%	100%	0%	41%	2%	47%	1%	41%	2%	99%	65%	0%	0%	100%
Wake et al (2022)[20]	181	South Africa	cross-sectional	inpatients and outpatients	irrespective of symptoms	≤100 cells/µl	35 (13, 61)	0%	18%	0%	83%	0%	7%	4%	16%	4%	17%	0%	88%	46%	33%	12%	55%
Yoon et al (2019)[21]	492	Uganda	cohort	outpatients	irrespective of symptoms	≤350 cells/µl	149 (60, 246)	0%	0%	0%	90%	0%	12%	0%	14%	0%	12%	21%	99%	100%	0%	100%	0%

Risk of bias assessment questionnaire based on QUADAS-2 and results

Domain	Overall*	Patient selection					Index test			Reference standard				Flow and timing			
Signalling Question	High or unclear overall Risk of Bias	Was a consecutive or random sample of patients enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?	Summary Risk of Bias: Could the selection of patients have introduced bias?	Applicability concerns Are there concerns that the included patients do not match the review question?	Were the index test results (uLAM and spNAAT) interpreted without knowledge of the results of the reference standard?	Summary Risk of Bias:Could the conduct or interpretation of the index test have introduced bias?	Applicability concerns: Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Is the denominator likely to correctly classify the target condition? (Xpert only=no; Xpert + Culture= yes	Were the reference standard results interpreted without knowledge of the results of the index test?	Summary Risk of Bias: Could the reference standard, its conduct, or its interpretation have introduced bias?	Applicability concerns: Are there concerns that the target condition as defined by the reference standard does not match the review question?	Did all patients receive a reference standard?	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Summary Risk of Bias: Could the patient flow have introduced bias?
Broger et al (2019); Lawn et al (2015)	low	yes	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	no	no	low
Calligaro et al (2017)	low	yes	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	yes	yes	low
Ciccacci et al (2021)	high/unclear	yes	yes	yes	low	low	unclear	unclear	unclear	no	unclear	unclear	unclear	yes	yes	yes	low
Cummings et al (2019)	high/unclear	yes	yes	yes	low	low	unclear	unclear	unclear	no	unclear	unclear	unclear	unclear	yes	yes	low
Floridia et al (2017)	high/unclear	yes	yes	yes	low	low	unclear	unclear	low	no	unclear	unclear	unclear	yes	yes	yes	low
Gupta-Wright et al (2018), STAMP	low	yes	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	yes	yes	low
Huerga et al (2017)	low	yes	yes	yes	low	low	yes	low	low	yes	unclear	low	low	yes	yes	yes	low
Huerga et al (2019)	low	yes	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	no	yes	low
Huerga et al (2020)	low	yes	yes	yes	low	low	yes	low	low	no	yes	unclear	low	yes	yes	yes	low
Huerga et al (2021)	low	yes	yes	yes	low	low	yes	low	low	no	yes	unclear	low	yes	yes	yes	low
Kasaro et al (2020)	high/unclear	unclear	yes	yes	unclear	unclear	unclear	unclear	low	yes	unclear	unclear	low	yes	yes	no	unclear
Lawn et al (2012)	low	yes	yes	yes	low	low	unclear	unclear	low	yes	unclear	low	low	yes	yes	yes	low
Peter et al (2016)	low	yes	yes	unclear	unclear	unclear	yes	low	low	yes	yes	low	low	yes	yes	yes	low
Theron et al (2021a, NCT03187964)	low	unclear	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	yes	yes	low
Theron et al (2021b, NCT03187964)	low	unclear	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	yes	yes	low
Van Hoving et al (2019)	low	yes	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	yes	yes	low
Yoon et al (2019)	low	yes	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	yes	unclear	unclear
Thit et al (2017)	low	yes	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	yes	yes	low
Garcia et al (2020)	low	yes	yes	yes	low	low	yes	low	low	yes	yes	low	low	yes	yes	yes	low
Wake et al (2022)	high/unclear	Yes	Yes	Yes	low	low	unclear	high	low	yes	unclear	low	low	yes	yes	yes	low

*High or unclear overall risk of bias was pre-defined using the following logic: ≥2 summary risk of bias (sRoB) questions answered as “high” OR 1 sRoB questions answered as “high” and ≥1 sRoB questions answered as “unclear” OR ≥2 sRoB questions answered as “unclear”

Risk of bias assessment



Two-day sputum and urine sample provision

TB Symptoms	Setting	CD4	Total # participants	Urine # participants	Urine %	Sputum # participants	Sputum %
Any	any	any	10202	9957	98%	8360	82%
	inpatients	any	3662	3585	98%	2531	69%
	outpatients	any	6540	6372	97%	5829	89%
Any	any	>200	4502	4393	98%	3744	83%
	any	101-200	1797	1769	98%	1508	84%
	any	<=100	3138	3065	98%	2452	78%
TB symptoms	any	any	8525	8321	98%	6874	81%
	inpatients	any	3461	3388	98%	2457	71%
	outpatients	any	5064	4933	97%	4417	87%
No TB symptoms	any	any	1677	1636	98%	1486	89%
	inpatients	any	201	197	98%	74	37%
	outpatients	any	1476	1439	97%	1412	96%
Unselected, symptoms not assessed	any	any	5638	5534	98%	4508	80%
	inpatients	any	1993	1966	99%	1084	54%
		<200	1005	989	98%	545	54%
	outpatients	any	3645	3568	98%	3424	94%
		<=200	1613	1573	98%	1467	91%
		<=100	864	832	96%	761	88%

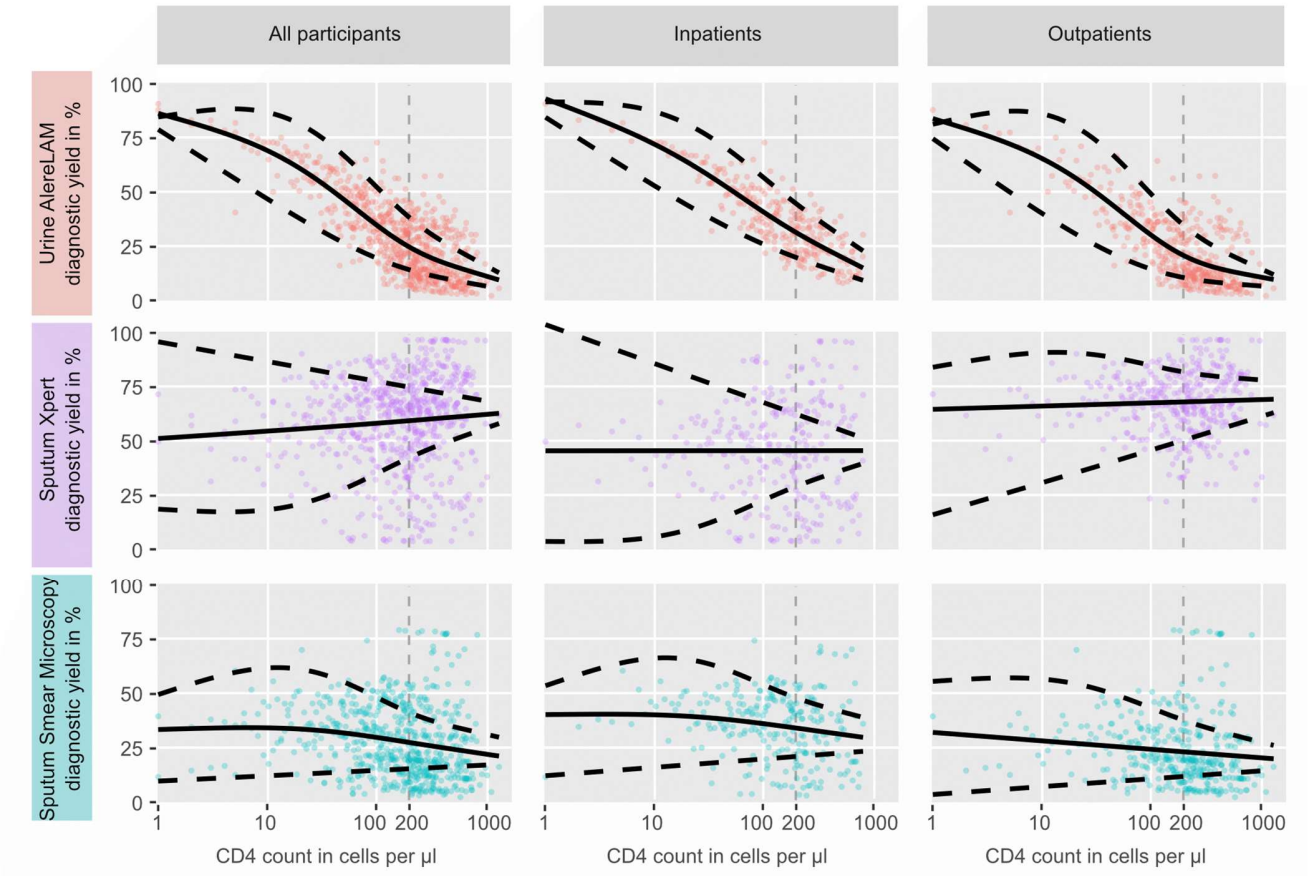
Two-day sample provision per TB symptoms, setting and CD4 count groups. CD4 counts are in cells per µL.

Analysis of variable effect – odds ratios

	Total n	Urine AlereLAM OR (95% CrI)	Sputum Xpert OR (95% CrI)	Sputum smear microscopy OR (95% CrI)
CD4 count				
per 200 cells/μL decrease	1531	3.47 (2.77, 4.36)*	1.04 (0.86, 1.25)	1.17 (0.97, 1.38)
TB symptom screen				
symptomatic	1538	Reference condition	Reference condition	Reference condition
asymptomatic	77	0.49 (0.21, 0.94) *	0.38 (0.21, 0.64) *	0.40 (0.15, 0.87)*
Setting				
inpatient	685	Reference condition	Reference condition	Reference condition
outpatient	930	0.80 (0.49, 1.24)	1.51 (0.74, 2.61)	0.77 (0.45, 1.26)
Age				
per 10 years' increase	1615	0.93 (0.81, 1.06)	0.88 (0.77, 1.00)	1.00 (0.86, 1.15)
On ART				
not on ART	984	Reference condition	Reference condition	Reference condition
on ART	631	1.12 (0.84, 1.43)	1.21 (0.89, 1.59)	0.89 (0.65, 1.16)
Sex				
female	757	Reference condition	Reference condition	Reference condition
male	858	0.77 (0.60, 0.98)*	1.01 (0.78, 1.29)	1.06 (0.79, 1.41)
Xpert cartridge				
Xpert MTB/RIF	858	Reference condition	Reference condition	Reference condition
Xpert Ultra	124	1.00 (1.00, 1.00)	1.51 (0.33, 4.41)	1.00 (1.00, 1.00)

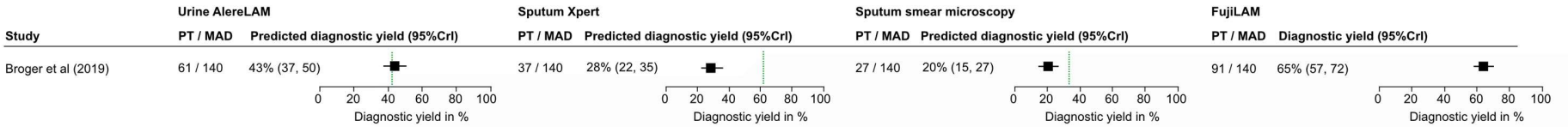
Adjusted odds ratios demonstrating the effect of potential confounding variables on diagnostic yield for each tuberculosis test among participants with tuberculosis based on the MAD. Odds ratios above one indicate higher diagnostic yield and odds ratios below one lower diagnostic yield. Significant odds rations with credibility intervals not including zero are highlighted with*. The effects of the key confounders are further analyzed in Figure 3 and Table 2 of the main paper. Crl=credible interval. ART=HIV antiretroviral therapy. Xpert=Xpert MTB/RIF or Xpert Ultra assay. AlereLAM=Alere Determine TB LAM Ag assay.

Tuberculosis diagnostic yield predictions as a function of CD4 count and setting



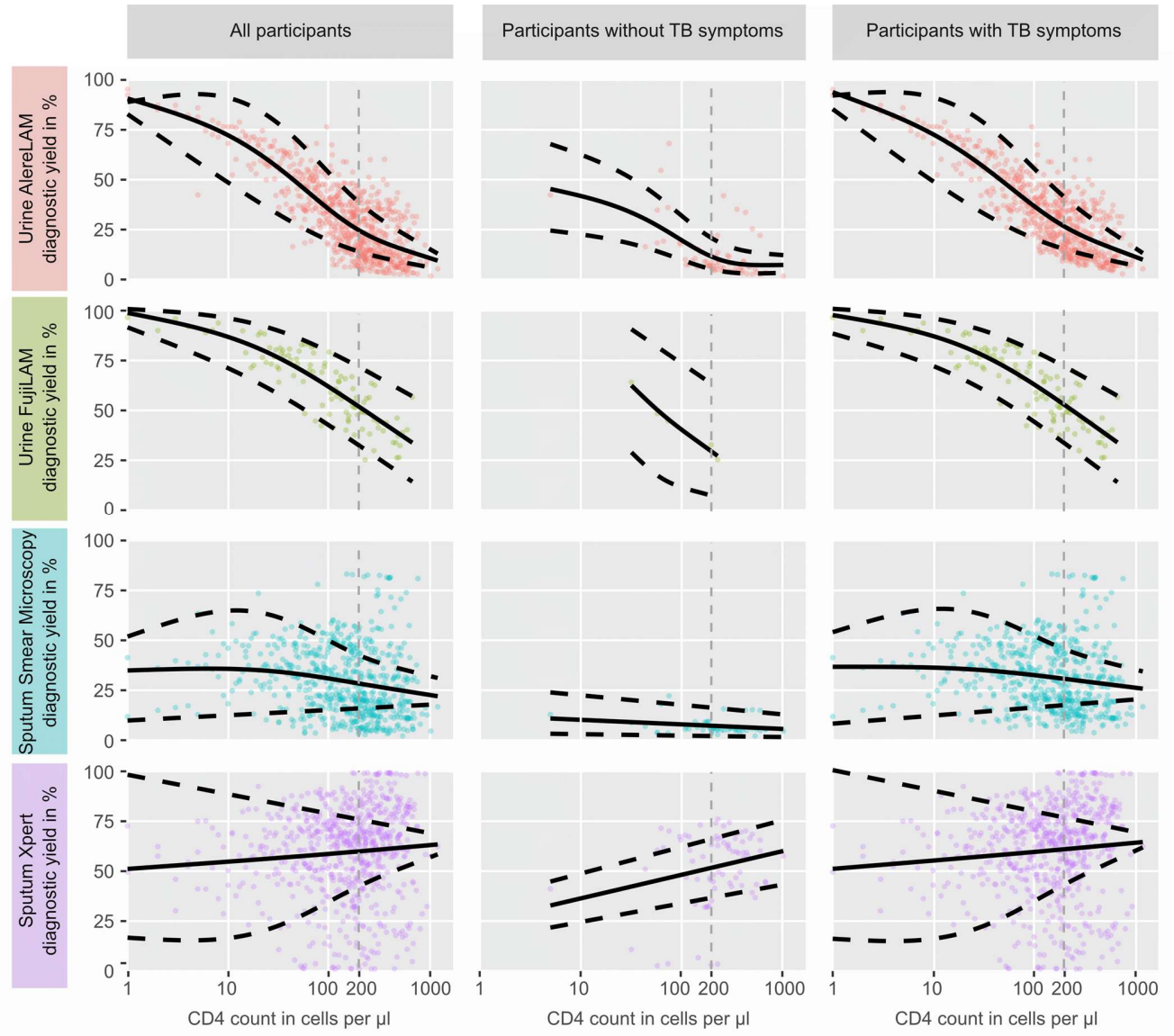
Tuberculosis diagnostic yield predictions as a function of CD4 count and setting. The MAD based on microbiologically confirmed tuberculosis was used as a denominator. Solid lines represent mean predictions, dashed lines 95% prediction intervals, and dots the participant data. AlereLAM=Alere Determine TB LAM Ag assay. Xpert=Xpert MTB/RIF or Xpert Ultra assay.

FujiLAM Diagnostic yield prediction



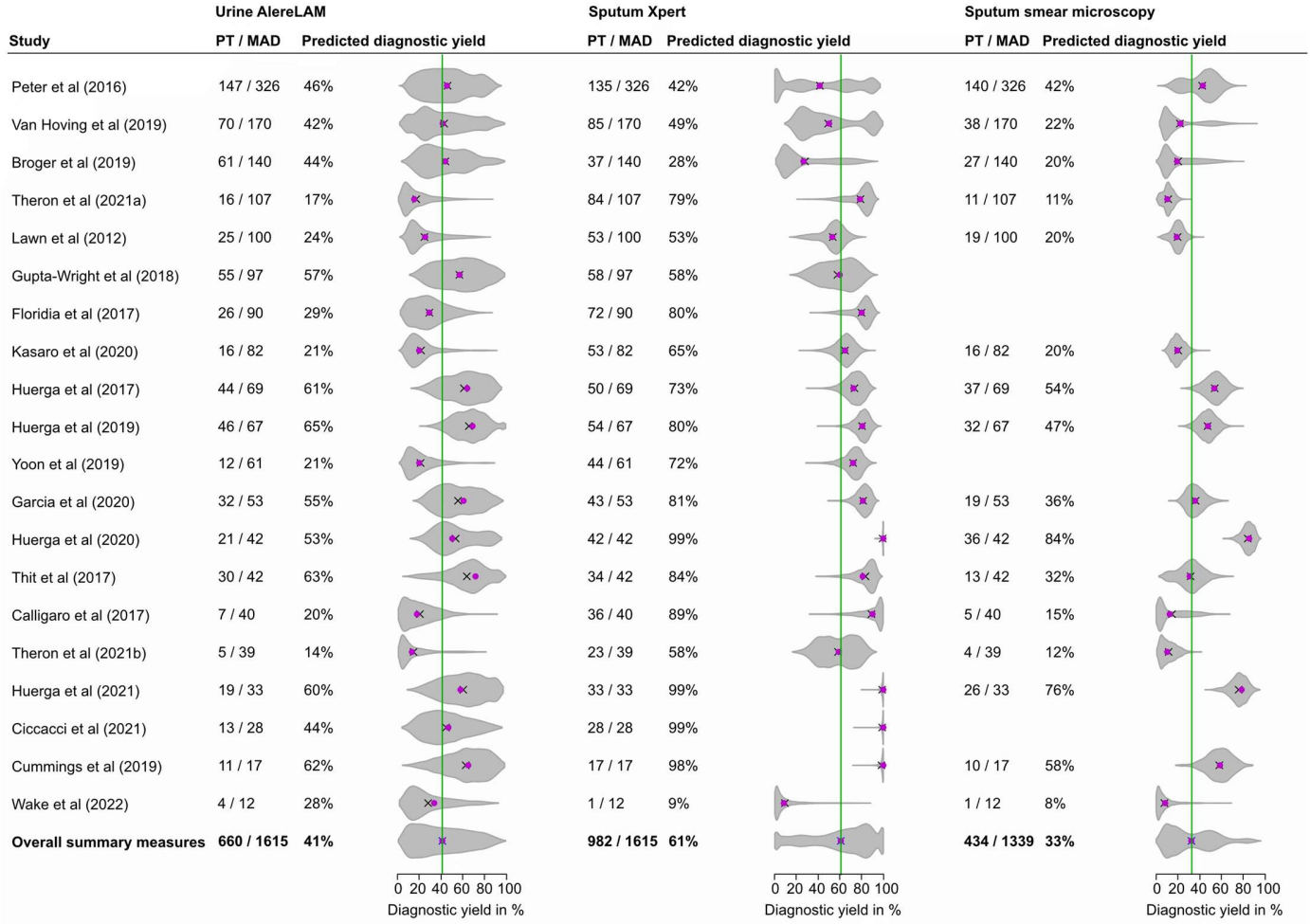
Diagnostic yields of urine AlereLAM, sputum Xpert, sputum smear microscopy, and FujiLAM for one study that allowed analysis of FujiLAM diagnostic yield. Solid squares represent mean predicted diagnostic yields. Horizontal lines indicate 95% credible intervals. The vertical dashed green lines indicate the predicted population mean (all datasets combined, see Figure 2A in the main paper). PT= number of positive tests from the first sample collected in the initial two days after enrolment. MAD=number of positive patients as defined by the harmonized meta-analysis denominator based on microbiologically confirmed tuberculosis. AlereLAM=Alere Determine TB LAM Ag assay. Xpert=Xpert MTB/RIF or Xpert Ultra assay. FujiLAM=Fujifilm Silvamp TB LAM assay. CrI=credible intervals

FujiLAM Diagnostic yield prediction as a function of CD4 count and TB symptoms



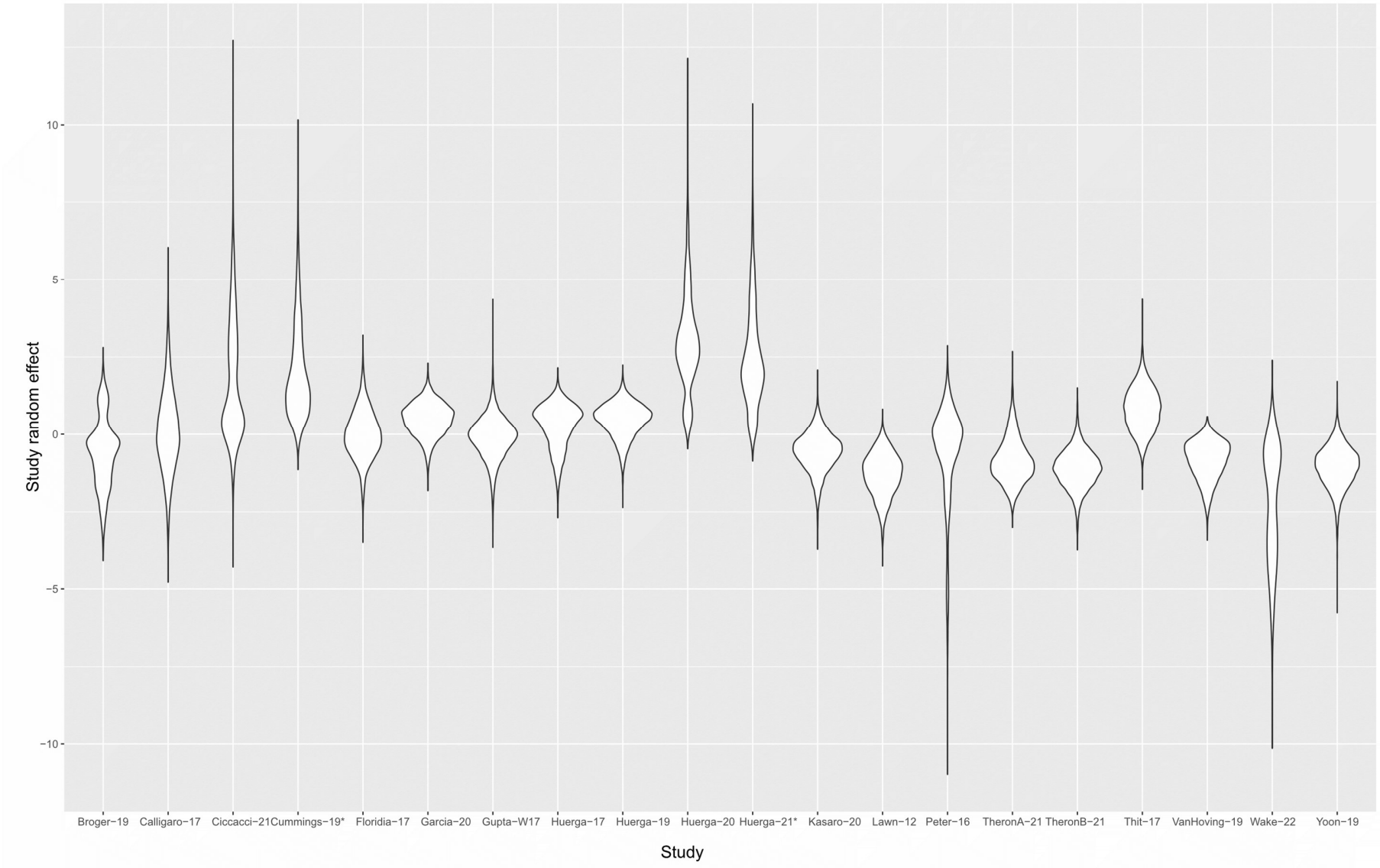
Diagnostic yield predictions as a function of CD4 count and TB symptoms. The MAD based on microbiologically confirmed tuberculosis was used as a denominator. Solid lines represent predictions, dashed lines 95% prediction intervals, and dots predictions for individual participants. AlereLAM=Alere Determine TB LAM Ag assay. FujiLAM=Fujifilm Silvamp TB LAM assay. Xpert=Xpert MTB/RIF or Xpert Ultra assay.

GLMM model diagnostic yield predictions



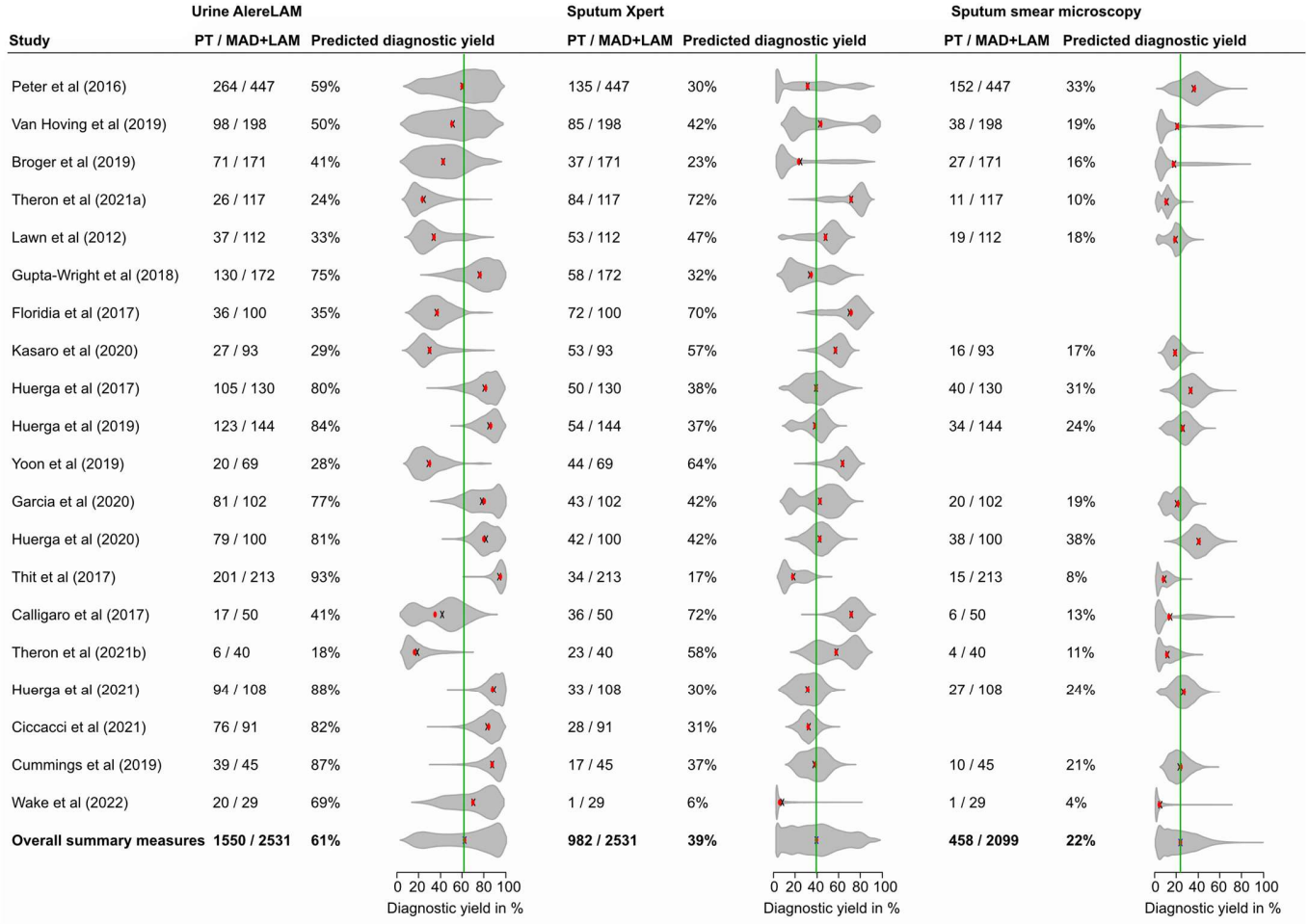
GLMM model diagnostic yield predictions for urine AlereLAM, sputum Xpert and sputum smear microscopy. Crosses represent mean predicted diagnostic yields and purple dots diagnostic yields from proportions (PT/MAD). Grey violin plots indicate the distribution of the data. Studies are sorted by size. PT= number of positive tests from the first sample collected in the initial two days after enrolment. MAD=number of positive patients as defined by the harmonized meta-analysis denominator based on microbiologically confirmed tuberculosis. AlereLAM=Alere Determine TB LAM Ag assay. Xpert=Xpert MTB/RIF or Xpert Ultra assay.

Quantification of random effect introduced by study (“heterogeneity”)



Study random effect. Studies with a significant random effect are marked with *.

Sensitivity analysis including LAM in the denominator (MAD–LAM)



Diagnostic yields of urine AlereLAM, sputum Xpert and sputum smear microscopy using the meta-analysis denominator including LAM (MAD–LAM). Crosses represent mean predicted diagnostic yields and red dots diagnostic yields from ratios. Grey violin plots indicate the distribution of the data. Studies are sorted by size. PT= number of positive tests from the first sample collected in the initial two days after enrolment. MAD–LAM=number of positive patients as defined by the harmonized meta-analysis denominator based on microbiologically confirmed tuberculosis or a positive LAM test. AlereLAM=Alere Determine TB LAM Ag assay. Xpert=Xpert MTB/RIF or Xpert Ultra assay.

Sensitivity analysis including LAM in the denominator (MAD–LAM) with adjustment for test specificity

Imperfect test specificity leads to an overestimation of diagnostic yield. In this sensitivity analysis we adjusted the diagnostic yield for test specificity by subtracting estimated number of false positives in the numerator and denominator. This was done for all tests (AlereLAM; Xpert, and sputum smear microscopy). As an example we show the approach for the AlereLAM adjusted yield:

$$DY_{adjusted,LAM} = \frac{PT_{LAM} - FP_{LAM}}{MAD - FP_{LAM}} \times 100\%$$

Whereas:

$$FP_{LAM} = PT_{LAM}\ x\ (1 - Specificity_{LAM})$$

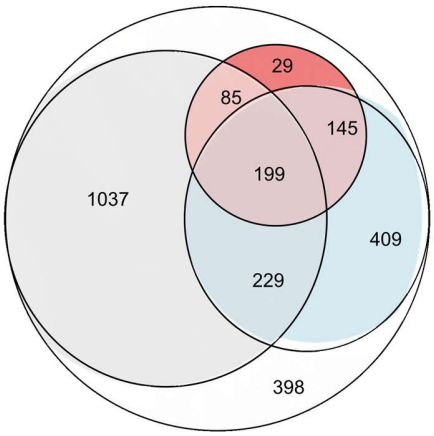
$DY_{adjusted,LAM}$: AlereLAM diagnostic yield adjusted for test specificity

PT_{LAM}: Number of AlereLAM positive tests on the first sample collected in the initial 2 days after enrolment

FP_{LAM}: Number of expected false positive AlereLAM results

MAD: Number of reference-standard positive patients as defined by the harmonized meta-analysis denominator

$Specificity_{LAM}$: Specificity of AlereLAM based on published meta-analysis



Test	DY Diagnostic yield	Specificity of the test	FP, Estimated false positives	DY adjusted
Urine AlereLAM	61.2% (1550/2531)	91%	140	59% (1411/2392)
Sputum Xpert	38.8% (982/2531)	98%	20	38.3% (962/2511)
Sputum smear microscopy (SSM)	21.8% (458/2099)	97%	14	21.3% (444/2085)
Urine AlereLAM + sputum Xpert	83.1% (2104/2531)	89%	153	82% (1950/2378)
Urine AlereLAM+SSM	82.1% (1724/2099)	88%	202	80.2% (1522/1897)
Tuberculosis cases missed	15.7% (398/2531)			

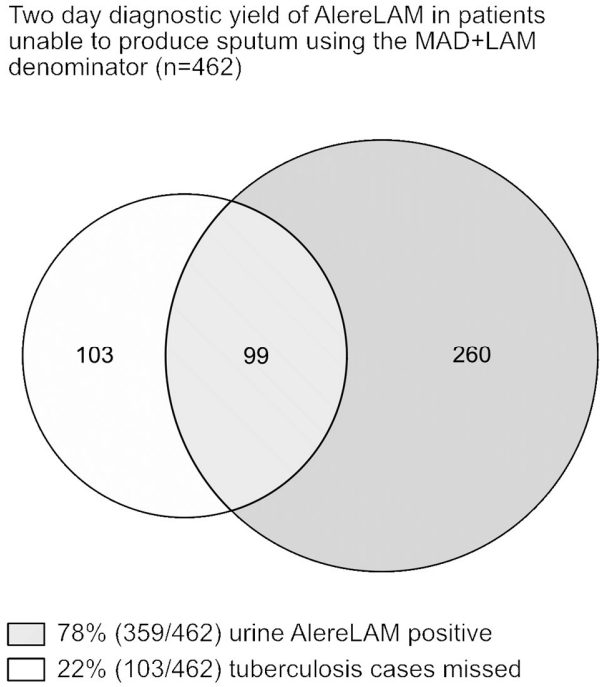
Specificities are based on the estimates from meta-analyses by Bjerrum et al. 2019[22] for AlereLAM, WHO's consolidated guidelines on tuberculosis for Xpert[23], and Steingart et al. 2006[24] for sputum smear microscopy.

Sputum smear microscopy (SSM) diagnostic yield for different microscopy methods

	Any SSM method			Fluorescence SSM			Ziehl-Neelsen SSM			Unspecified method		
Setting	**PT**	**MAD**	**DY**	**PT**	**MAD**	**DY**	**PT**	**MAD**	**DY**	**PT**	**MAD**	**DY**
Any	431	1283	34%	281	708	40%	13	109	12%	137	466	29%
Inpatients	236	581	40%	194	497	39%	0	0	NA	42	84	50%
Outpatients	195	702	28%	87	211	41%	13	109	12%	95	382	25%

Sputum smear microscopy (SSM) diagnostic yield in MAD positive participants with TB symptoms for different microscopy methods and settings. PT=number of positive tests from the first sputum sample collected in the initial two days after enrolment. MAD=number of positive participants as defined by the harmonized meta-analysis denominator based on microbiologically confirmed tuberculosis. SSM=Sputum smear microscopy. Unspecified method=seven studies performed SSM but didn't specify the method used.

Urine AlereLAM diagnostic yield in patients unable to provide sputum



Urine AlereLAM diagnostic yield in patients unable to provide sputum. Number of detected tuberculosis patients by AlereLAM on the first urine specimen obtained within two days of enrolment in all patients using microbiologically confirmed tuberculosis as the meta-analysis denominator including LAM (MAD–LAM). AlereLAM=Alere Determine TB LAM Ag assay. MAD=Meta-analysis denominator. LAM=Lipoarabinomannan.

Overview of sensitivity analyses

Analysis	Urine AlereLAM PT / Denominator	Predicted diagnostic yield	Sputum Xpert PT / Denominator	Predicted diagnostic yield	Sputum smear microscopy PT / Denominator	Predicted diagnostic yield
Main analysis using random effects meta-analysis (see Figure 2)	**660/1615**	**41% (95%CrI 15-66)**	**982/1615**	**61% (95%CrI 25-88)**	**434/1339**	**32% (95%CrI 10-55)**
GLMM (random and fixed effects) using MAD	660/1615	41%	982/1615	61%	434/1339	33%
GLMM after excluding studies with high or unclear Risk of Bias	590/1386	43%	811/1386	58%	407/1228	33%
GLMM after excluding studies that did not perform microbiological testing on samples other than sputum and urine	122/319	38%	163/319	51%	86/284	30%
GLMM including LAM in the Denominator (MAD–LAM, see above)	1550/2531	61%	982/2531	39%	458/2099	22%
Diagnostic yield including LAM in the Denominator (MAD–LAM) with adjustment for test specificity	1411/2392	59%	962/2511	38%	444/2085	21%

Comparison of diagnostic yield to calculated diagnostic yields based on previous meta-analyses

	Urine AlereLAM			Sputum Xpert				Sputum smear microscopy		
	Sn (95%CI)	**Sp (95%CI)**	**Reference**	**Sn (95%CI)**	**Sp (95%CI)**	**Comment**	**Reference**	**Sn (95%CI)**	**Sp (95%CI)**	**Reference**
Performance in PLHIV reported in other meta-analyses	42% (31-55)	91% (85-95)	Bjerrum et al. 2019[22]	77% (71-82) 81% (75-86) 75% (59-86) 79% (70-86) 88% (75-904)	98% (98-99) 98% (97-99) 100% (99-100) 98% (96-99) 93% (82-97)	Xpert MTB/RIF, result A Xpert MTB/RIF, result B Xpert MTB/RIF Xpert MTB/RIF Xpert Ultra	Horne et al. 2019[25] Horne et al. 2019[25] Zifodya et al. 2021[26] Steingart et al. 2014[27] Zifodya et al. 2021[26]	53% (43-63) for FM ZN 10% below FM	96% (86-99)	Chang et al. 2016[28] Steingart et al. 2006[24]
Sample provision reported in this IPD-MA	98%			82%				82%		
Calculated diagnostic yield based on previous meta-analysis sensitivity and sample provision from this IPD-MA (Sensitivity x sample provision)	41%			62-66% Xpert MTB/RIF 72% Xpert Ultra				43% FM 35% ZN		
Diagnostic yield reported in this IPD-MA	**41%**			**61%**				**32%**		
Difference	**0%**			**-1 to -5 % Xpert MTB/RIF**				**-11 % FM -3 % ZN**		

Comparison of diagnostic yield to calculated diagnostic yields based on sensitivity data from previous meta-analyses. The green row summarizes performance data from meta-analysis that informed WHO policy. The yellow row includes two-day sample provision percentages from this IPD-MA. The light blue row includes calculated diagnostic yields based on the sensitivities from the green row and the sample provision from the yellow row to estimate expected diagnostic yields based on previously published evidence. The dark blue row compares the diagnostic yields from this IPD-MA to the expected diagnostic yields based on previously published meta-analyses. For AlereLAM the diagnostic yield from this meta-analysis equal, for sputum Xpert the yield from this IPD-MA is 1 to 5 % lower, and for ZN sputum smear microscopy the yield is 3 % lower compared to what would be than what would be expected based on previous meta-analyses. FM=Fluorescence microscopy. ZN=Ziehl-Neelsen microscopy. MA=Meta-analysis. IPD-MA=Individual participant data meta-analysis. Sn=Sensitivity. Sp=Specificity.

Supplementary References

1. Broger T, Sossen B, du Toit E, Kerkhoff AD, Schutz C, Ivanova Reipold E, et al. Novel lipoarabinomannan point-of-care tuberculosis test for people with HIV: a diagnostic accuracy study. Lancet Infect Dis [Internet]. 2019 Aug;19(8):852–61. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309919300015

2. Lawn SD, Kerkhoff AD, Burton R, Schutz C, van Wyk G, Vogt M, 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 [Internet]. 2015;13(1):192. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4537538&tool=pmcentrez&rendertype=abstract

3. Lawn SD, Kerkhoff AD, Burton R, Schutz C, Boulle A, Vogt M, 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 [Internet]. 2017 Dec 21;15(1):67. Available from: http://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-017-0822-8

4. Calligaro GL, Zijenah LS, Peter JG, Theron G, Buser V, McNerney R, et al. Effect of new tuberculosis diagnostic technologies on community-based intensified case finding: a multicentre randomised controlled trial. Lancet Infect Dis [Internet]. 2017 Apr;17(4):441–50. Available from: http://dx.doi.org/10.1016/S1473-3099(16)30384-X

5. Ciccacci F, Welu B, Ndoi H, Karea I, Orlando S, Brambilla D, 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 [Internet]. 2021;104(PG-329-334):329–34. Available from: NS -

6. Cummings MJ, Bakamutumaho B, Owor N, Kayiwa J, Byaruhanga T, Namagambo B, 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.

7. Floridia M, Ciccacci F, Andreotti M, Hassane A, Sidumo Z, Magid NA, 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 Dec;65(11):1878–83.

8. García JI, Meléndez J, Álvarez R, Mejía-Chew C, Kelley H V, Sidiki S, 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 [Internet]. 2020 Dec 19;17(1):62. Available from: https://aidsrestherapy.biomedcentral.com/articles/10.1186/s12981-020-00318-8

9. Gupta-Wright A, Corbett EL, van Oosterhout JJ, Wilson D, Grint D, Alufandika-Moyo M, 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 [Internet]. 2018 Jul;392(10144):292–301. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673618312674

10. Huerga H, Ferlazzo G, Bevilacqua P, Kirubi B, Ardizzoni E, Wanjala S, 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 Jan;12(1).

11. Huerga H, Rucker SCM, Cossa L, Bastard M, Amoros I, Manhica I, et al. Diagnostic value of the urine lipoarabinomannan assay in HIV-positive, ambulatory patients with CD4 below 200 cells/mu l in 2 low-resource settings: A prospective observational study. PLOS Med [Internet]. 2019 Apr 30;16(4):e1002792. Available from: https://dx.plos.org/10.1371/journal.pmed.1002792

12. Huerga H, Rucker SCM, Bastard M, Dimba A, Kamba C, Amoros I, 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. JAIDS-JOURNAL Acquir IMMUNE Defic Syndr. 2020 Jan;83(1):E28.

13. Huerga H, SC MR, Bastard M, Mpunga J, I AQ, Kabaghe C, et al. Urine Lipoarabinomannan Testing for All HIV Patients Hospitalized in Medical Wards Identifies a Large Proportion of Patients With Tuberculosis at Risk of Death. Open Forum Infect Dis [Internet]. 2021;8(2 PG-ofaa639):ofaa639. Available from: NS -

14. Kasaro MP, Chilyabanyama ON, Shah NS, Muluka B, Kapata N, Krüüner A, et al. Performance of Xpert ® MTB/RIF and

Determine™ TB-LAM Ag in HIV-infected adults in peri-urban sites in Zambia. Public Heal Action [Internet]. 2020 Dec 21;10(4):134–40. Available from: https://www.ingentaconnect.com/content/10.5588/pha.20.0010

15. Lawn SD, Kerkhoff AD, Vogt M, Wood R, SD L, AD K, et al. 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 [Internet]. 2012 Mar [cited 2014 Sep 16];12(3):201–9. Available from: NS -

16. Peter JG, Zijenah LS, Chanda D, Clowes P, Lesosky M, Gina P, 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 [Internet]. 2016 Mar;387(10024):1187–97. Available from: http://dx.doi.org/10.1016/S0140-6736(15)01092-2

17. Theron G. NCT03187964 - Xpert Ultra and Xpert HIV-VL in People Living With HIV [Internet]. Available from: https://clinicaltrials.gov/ct2/show/NCT03187964

18. Thit SS, Aung NM, Htet ZW, Boyd MA, Saw HA, Anstey NM, et al. The clinical utility of the urine-based lateral flow lipoarabinomannan assay in HIV-infected adults in Myanmar: an observational study. BMC Med [Internet]. 2017 Dec 4;15(1):145. Available from: http://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-017-0888-3

19. Van Hoving DJDJ, Lahri S, Lategan HJ, Nicol MP, Maartens G, Meintjes G. Real-World Performance and Interobserver Agreement of Urine Lipoarabinomannan in Diagnosing HIV-Associated Tuberculosis in an Emergency Center. JAIDS-JOURNAL Acquir IMMUNE Defic Syndr [Internet]. 2019 May 1;81(1):E10–4. Available from: https://journals.lww.com/00126334-201905010-00019

20. Wake RM, Govender NP, Omar S V., Ismail F, Tiemessen CT, Harrison TS, et al. Rapid urine-based screening tests increase the yield of same-day tuberculosis diagnoses among patients living with advanced HIV disease. AIDS [Internet]. 2022 Jan 24;Publish Ah. Available from: https://journals.lww.com/10.1097/QAD.0000000000003177

21. Yoon C, Semitala FC, Asege L, Katende J, Mwebe S, Andama AO, et al. Yield and Efficiency of Novel Intensified Tuberculosis Case-Finding Algorithms for People Living with HIV. Am J Respir Crit Care Med [Internet]. 2018 Sep 7;199(5):1–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30192649

22. Bjerrum S, Schiller I, Dendukuri N, Kohli M, Nathavitharana RR, Zwerling AA, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in people living with HIV. Cochrane Database Syst Rev [Internet]. 2019 Oct 21;2019(10). Available from: http://doi.wiley.com/10.1002/14651858.CD011420.pub3

23. World Health Organization. Consolidated Guidelines on Tuberculosis - Module 3: Diagnosis [Internet]. Geneva, Switzerland; 2020. Available from: https://apps.who.int/iris/rest/bitstreams/1284627/retrieve

24. Steingart KR, Henry M, Ng V, Hopewell PC, Ramsay A, Cunningham J, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis [Internet]. 2006 Sep;6(9):570–81. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1473309906705783

25. Horne DJ, Kohli M, Zifodya JS, Schiller I, Dendukuri N, Tollefson D, et al. Xpert MTB/RIF and Xpert MTB/RIF Ultra for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev [Internet]. 2019 Jun 7;(6). Available from: http://doi.wiley.com/10.1002/14651858.CD009593.pub4

26. Zifodya JS, Kreniske JS, Schiller I, Kohli M, Dendukuri N, Schumacher SG, et al. Xpert Ultra versus Xpert MTB/RIF for pulmonary tuberculosis and rifampicin resistance in adults with presumptive pulmonary tuberculosis. Cochrane Database Syst Rev [Internet]. 2021 Feb 22;2021(5). Available from: http://doi.wiley.com/10.1002/14651858.CD009593.pub5

27. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Steingart KR, editor. Cochrane database Syst Rev [Internet]. 2014 Jan 21;1(1):CD009593. Available from: http://doi.wiley.com/10.1002/14651858.CD009593.pub3

28. Chang EW, Page A-L, Bonnet M. Light-emitting diode fluorescence microscopy for tuberculosis diagnosis: a meta-analysis. Eur Respir J [Internet]. 2016 Mar;47(3):929–37. Available from: http://dx.doi.org/10.1183/13993003.00978-2015

PRISMA-IPD Checklist

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Title			
Title	1	Identify the report as a systematic review and meta-analysis of individual participant data.	Done
Abstract			
Structured summary	2	Provide a structured summary including as applicable:	Done
		Background: state research question and main objectives, with information on participants, interventions, comparators and outcomes.	
		Methods: report eligibility criteria; data sources including dates of last bibliographic search or elicitation, noting that IPD were sought; methods of assessing risk of bias.	
		Results: provide number and type of studies and participants identified and number (%) obtained; summary effect estimates for main outcomes (benefits and harms) with confidence intervals and measures of statistical heterogeneity. Describe the direction and size of summary effects in terms meaningful to those who would put findings into practice.	
		Discussion: state main strengths and limitations of the evidence, general interpretation of the results and any important implications.	
		Other: report primary funding source, registration number and registry name for the systematic review and IPD meta-analysis.	
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known.	Done
Objectives	4	Provide an explicit statement of the questions being addressed with reference, as applicable, to participants, interventions, comparisons, outcomes and study design (PICOS). Include any hypotheses that relate to particular types of participant-level subgroups.	Not done
Methods			
Protocol and registration	5	Indicate if a protocol exists and where it can be accessed. If available, provide registration information including registration number and registry name. Provide publication details, if applicable.	Done
Eligibility criteria	6	Specify inclusion and exclusion criteria including those relating to participants, interventions, comparisons, outcomes, study design and characteristics (e.g. years when conducted, required minimum follow-up). Note whether these were applied at the study or individual level i.e. whether eligible participants were included (and ineligible participants excluded) from a study that included a wider population than specified by the review inclusion criteria. The rationale for criteria should be stated.	Done
Identifying studies -	7	Describe all methods of identifying published and unpublished studies including, as applicable: which bibliographic databases were searched with dates of coverage; details of any hand searching including of conference proceedings; use of study registers	Done

information sources		and agency or company databases; contact with the original research team and experts in the field; open adverts and surveys. Give the date of last search or elicitation.	
Identifying studies - search	8	Present the full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Done
Study selection processes	9	State the process for determining which studies were eligible for inclusion.	Done
Data collection processes	10	Describe how IPD were requested, collected and managed, including any processes for querying and confirming data with investigators. If IPD were not sought from any eligible study, the reason for this should be stated (for each such study).	Done
		If applicable, describe how any studies for which IPD were not available were dealt with. This should include whether, how and what aggregate data were sought or extracted from study reports and publications (such as extracting data independently in duplicate) and any processes for obtaining and confirming these data with investigators.	
Data items	11	Describe how the information and variables to be collected were chosen. List and define all study level and participant level data that were sought, including baseline and follow-up information. If applicable, describe methods of standardising or translating variables within the IPD datasets to ensure common scales or measurements across studies.	Done
IPD integrity	A1	Describe what aspects of IPD were subject to data checking (such as sequence generation, data consistency and completeness, baseline imbalance) and how this was done.	Done
Risk of bias assessment in individual studies.	12	Describe methods used to assess risk of bias in the individual studies and whether this was applied separately for each outcome. If applicable, describe how findings of IPD checking were used to inform the assessment. Report if and how risk of bias assessment was used in any data synthesis.	Done
Specification of outcomes and effect measures	13	State all treatment comparisons of interests. State all outcomes addressed and define them in detail. State whether they were pre-specified for the review and, if applicable, whether they were primary/main or secondary/additional outcomes. Give the principal measures of effect (such as risk ratio, hazard ratio, difference in means) used for each outcome.	Done
Synthesis methods	14	Describe the meta-analysis methods used to synthesise IPD. Specify any statistical methods and models used. Issues should include (but are not restricted to): • Use of a one-stage or two-stage approach. • How effect estimates were generated separately within each study and combined across studies (where applicable). • Specification of one-stage models (where applicable) including how clustering of patients within studies was accounted for. • Use of fixed or random effects models and any other model assumptions, such as proportional hazards. • How (summary) survival curves were generated (where applicable). • Methods for quantifying statistical heterogeneity (such as I^2 and τ^2). • How studies providing IPD and not providing IPD were analysed together (where applicable). • How missing data within the IPD were dealt with (where applicable).	Done

Exploration of variation in effects	A2	If applicable, describe any methods used to explore variation in effects by study or participant level characteristics (such as estimation of interactions between effect and covariates). State all participant-level characteristics that were analysed as potential effect modifiers, and whether these were pre-specified.	Done
Risk of bias across studies	15	Specify any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to not obtaining IPD for particular studies, outcomes or other variables.	Done
Additional analyses	16	Describe methods of any additional analyses, including sensitivity analyses. State which of these were pre-specified.	Done
Results			
Study selection and IPD obtained	17	Give numbers of studies screened, assessed for eligibility, and included in the systematic review with reasons for exclusions at each stage. Indicate the number of studies and participants for which IPD were sought and for which IPD were obtained. For those studies where IPD were not available, give the numbers of studies and participants for which aggregate data were available. Report reasons for non-availability of IPD. Include a flow diagram.	Done
Study characteristics	18	For each study, present information on key study and participant characteristics (such as description of interventions, numbers of participants, demographic data, unavailability of outcomes, funding source, and if applicable duration of follow-up). Provide (main) citations for each study. Where applicable, also report similar study characteristics for any studies not providing IPD.	Done
IPD integrity	A3	Report any important issues identified in checking IPD or state that there were none.	Not done
Risk of bias within studies	19	Present data on risk of bias assessments. If applicable, describe whether data checking led to the up-weighting or down-weighting of these assessments. Consider how any potential bias impacts on the robustness of meta-analysis conclusions.	Done
Results of individual studies	20	For each comparison and for each main outcome (benefit or harm), for each individual study report the number of eligible participants for which data were obtained and show simple summary data for each intervention group (including, where applicable, the number of events), effect estimates and confidence intervals. These may be tabulated or included on a forest plot.	Done
Results of syntheses	21	Present summary effects for each meta-analysis undertaken, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified, and report the numbers of studies and participants and, where applicable, the number of events on which it is based.	Done
		When exploring variation in effects due to patient or study characteristics, present summary interaction estimates for each characteristic examined, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified. State whether any interaction is consistent across trials.	
		Provide a description of the direction and size of effect in terms meaningful to those who would put findings into practice.	
Risk of bias across studies	22	Present results of any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to the availability and representativeness of available studies, outcomes or other variables.	Done

Additional analyses	23	Give results of any additional analyses (e.g. sensitivity analyses). If applicable, this should also include any analyses that incorporate aggregate data for studies that do not have IPD. If applicable, summarise the main meta-analysis results following the inclusion or exclusion of studies for which IPD were not available.	Done
Discussion			
Summary of evidence	24	Summarise the main findings, including the strength of evidence for each main outcome.	Done
Strengths and limitations	25	Discuss any important strengths and limitations of the evidence including the benefits of access to IPD and any limitations arising from IPD that were not available.	Done
Conclusions	26	Provide a general interpretation of the findings in the context of other evidence.	Done
Implications	A4	Consider relevance to key groups (such as policy makers, service providers and service users). Consider implications for future research.	Done
Funding			
Funding	27	Describe sources of funding and other support (such as supply of IPD), and the role in the systematic review of those providing such support.	Done

Statistical analysis plan

Diagnostic yield of urine lipoarabinomannan and sputum tuberculosis tests: a systematic review and meta-analysis of individual patient data

Statistical Analysis Plan (SAP)

Version: 1.0

Date: 04.09.2021

Tobias Broger[1§]

Ioana Diana Olaru[1]

Lisa Koeppel[1]

Claudia M Denkinger[1]

[1]Section Clinical Tropical Medicine, Heidelberg University Hospital, Heidelberg, Germany

[§]tobias.broger@uni-heidelberg.de, +41 78 843 21 12

1. Acronyms

Alere Determine TB LAM Ag test	AlereLAM
Fujifilm SILVAMP TB LAM	FujiLAM
Generalized linear mixed model	GLMM
GeneXpert MTB/RIF or GeneXpert MTB/RIF Ultra	Xpert
Individual patient data meta-analysis	IPD-MA
Intention-to-test	ITT
Interquartile range	IQR
Lipoarabinomannan	LAM
Modified Intention-to-test	MITT
Mycobacterium tuberculosis	*Mtb*
Per Protocol Population	PP
Prediction Interval	PI
Meta-analysis denominator	MAD
Sputum Smear Microscopy	SSM
Statistical Analysis Plan	SAP
Positive Test	PT
Tuberculosis	TB

2. Background, introduction and purpose of SAP

This document describes the statistical analysis plan for the following systematic review: Diagnostic yield of urine lipoarabinomannan and sputum tuberculosis tests: a systematic review and meta-analysis of individual patient data. The systematic review was specified in the Systematic Review Protocol from February 6th 2021 and pre-registered on February 14th on PROSPERO (available under: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=230337). The purpose of this document is to define the data analysis strategy and to describe the statistical methods that will be used. This document will also be used to gather feedback from investigators of included studies to reach consensus on the analysis.

3. Review questions and analyses

The primary question of this study is: What are the TB diagnostic yields of (1) urine LAM point-of-care tests on the first available urine sample and (2) sputum nucleic acid amplification tests (NAATs) on the first available sputum sample and (3) sputum smear microscopy (SSM) on the first available sputum sample against a harmonized, meta-analysis denominator (MAD, see definition below) across studies meeting the inclusion criteria?

Analyses, subgroup analyses and sensitivity analyses will be done for four different populations as defined in the Analysis populations section below.

The primary and secondary objectives and related analyses are summarized in the following table:

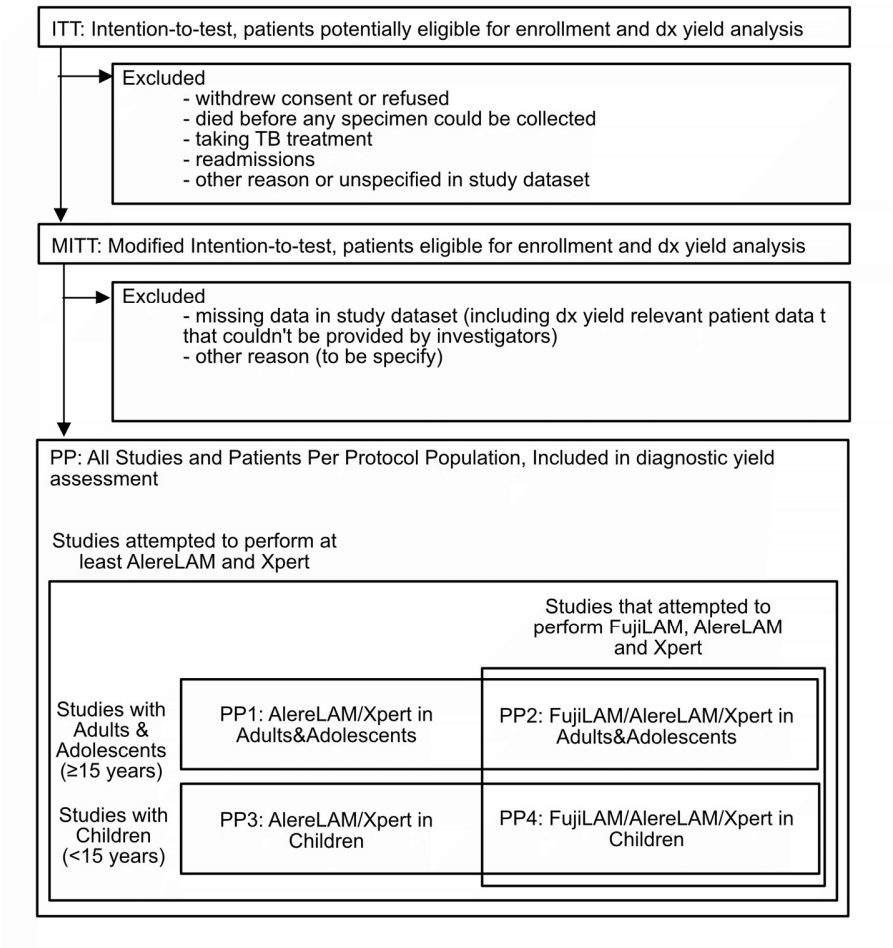
Primary Objectives	**Analysis**
1.1 To determine the TB diagnostic yield of urine AlereLAM, urine FujiLAM, sputum Xpert*, and SSM# from the first baseline diagnostic sample collection attempt§ against the reconstructed meta-analysis denominator (MAD)	Pooled estimates of diagnostic yield and 95% prediction intervals (PI) of AlereLAM, FujiLAM, Xpert, and SSM using fixed or random-effect meta-analysis using a GLMM method. Further details are provided in the statistical section below.
Secondary Objectives	**Analysis**
2.1 To determine the combined TB diagnostic yield of urine AlereLAM and sputum Xpert* from the first baseline diagnostic sample collection attempt§ against the reconstructed meta-analysis denominator (MAD)	Pooled estimates of combined diagnostic yield and 95% prediction intervals (PI) using fixed or random-effect meta-analysis using a GLMM method. Further details are provided in the statistical section below.
2.2 To determine the combined TB diagnostic yield of urine FujiLAM and sputum Xpert* from the first baseline diagnostic sample collection attempt§ against the reconstructed meta-analysis denominator (MAD)	As in 2.1
2.3 To determine the combined TB diagnostic yield of urine AlereLAM and SSM# from the first baseline diagnostic sample collection attempt§ against the reconstructed meta-analysis denominator (MAD)	As in 2.1
2.4 To determine the combined TB diagnostic yield of urine FujiLAM and SSM# from the first baseline diagnostic	As in 2.1

	sample collection attempt[§] against the reconstructed meta-analysis denominator (MAD)	
2.5	Adults: To determine the proportion of patients that can provide a baseline urine. Separately proportion of patients that can provide a non-induced sputum baseline sample. Children: To determine the proportion of patients that can provide a baseline urine, non-induced sputum, induced sputum, gastric aspirate (GA), nasopharyngeal aspirate (NPA), gastric lavage (GL) (if data can be obtained from primary studies).	Pooled estimate of percentage and 95% prediction intervals per sample collection method.

* Xpert MTB/RIF and Xpert MTB/RIF Ultra will be combined and treated equivalent but subgroup analysis of Xpert MTB/RIF vs. Xpert MTB/RIF Ultra is planned. Reporting on other NAAT's will be done separately if relevant studies will be identified.
[#] Ziehl–Neelsen and Fluorescence Microscopy will be combined and treated equivalent. This will be communicated as a study limitation
§ the first baseline diagnostic sample collection attempt is defined as the first attempt to collect a urine or sputum sample within the first 2 days of including a patient in the study. Typically, the attempts to collect these samples are done on the day of enrolment.

4. Analyses Populations

The prototype flow diagram below defines four analyses populations which will all be analysed separately to determine diagnostic yield. Currently it is planned to have separate publications for Studies with Children <15 years of age (PP3 and PP4) and Adults&adolescents ≥15 years (PP1 and PP2).



5. Denominator

Meta-analysis denominator (MAD)

The MAD is defined in the table below. LAM will initially be excluded from the denominator but a sensitivity analysis will be conducted with LAM in the denominator.

Adults and Adolescents	Children
Microbiologically confirmed TB will be used as a denominator and defined as follows: - Any culture (liquid or solid) positive for *Mtb* from any sample type or - any Xpert positive for *Mtb* from any sample type Only microbiologically confirmed TB will be deemed TB positive. Possible TB or clinical definitions of TB will be deemed negative because harmonisation across studies is not feasible. This will be mentioned as a limitation. Unclear cases will be discussed with the authors from the primary study.	All confirmed TB or unconfirmed TB cases based on the study definition. To the extent possible and in close discussions with authors of the primary study the case definition will be harmonized towards the consensus definitions from Graham et al. 2015. (1) In summary Graham et al. definitions are: **Confirmed TB** is defined as microbiologically confirmed TB (culture or Xpert from at least 1 respiratory specimen). **Unconfirmed TB** is defined as bacterial confirmation NOT obtained, and at least 2 of the following: (i) symptoms/signs suggestive of TB (persistent cough, weight loss/failure to thrive, persistent unexplained fever, persistent unexplained lethargy or reduced playfulness) (ii) chest radiograph consistent with TB, (iii) Close TB exposure or immunologic evidence of *Mtb* infection, (iv) positive response to TB treatment Justification: In contrast to adults, microbiological diagnosis of TB in children is difficult and yield is low.

6. Description of analyses and statistical methods

PRISMA Flow Diagram

The number of studies identified, screened, eligible and included will be described in a PRISMA Flow Diagram.

Descriptive statistics

The number of participants included and excluded will be reported using a STARD-like flow diagram (see prototype diagram above). Descriptive statistics tables will be generated to summarize the characteristics of the participants in the PP populations like in the following prototype table:

	Overall, N	**Study 1, N**	**…**	**Study k, N**
	median (IQR)	median (IQR)	median (IQR)	median (IQR)
Country				
South Africa	N (%)	N (%)	N (%)	N (%)
Kenya	N (%)	N (%)	N (%)	N (%)
Malawi	N (%)	N (%)	N (%)	N (%)
Mozambique	N (%)	N (%)	N (%)	N (%)
Zimbabwe	N (%)	N (%)	N (%)	N (%)
Tanzania	N (%)	N (%)	N (%)	N (%)
Uganda	N (%)	N (%)	N (%)	N (%)
Zambia	N (%)	N (%)	N (%)	N (%)
…	N (%)	N (%)	N (%)	N (%)
Age group				
Children: Age <2	N (%)	N (%)	N (%)	N (%)
Children: Age ≥2, <5	N (%)	N (%)	N (%)	N (%)
Children: Age ≥5, <15	N (%)	N (%)	N (%)	N (%)
Age ≥15, <18	N (%)	N (%)	N (%)	N (%)
Age ≥18	N (%)	N (%)	N (%)	N (%)

Sex				
Female	N (%)	N (%)	N (%)	N (%)
Male	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
HIV Positive	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
CD4 cell count, if HIV positive	median (IQR)	median (IQR)	median (IQR)	median (IQR)
CD4 group, if HIV positive				
CD4 ≤ 100 cells/mm^3	N (%)	N (%)	N (%)	N (%)
CD4 101-200 cells/mm^3	N (%)	N (%)	N (%)	N (%)
CD4 > 200 cells/mm^3	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
HIV-associated immunosuppression (for children, by WHO age -specific CD4% or count cut-offs, if HIV positive)				
Severe[$]	N (%)	N (%)	N (%)	N (%)
Not severe[$]	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
On ART, if HIV positive	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
Prior TB History	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
WHO TB symptom screen				
Symptomatic	N (%)	N (%)	N (%)	N (%)
Asymptomatic	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
Recruitment Setting				
Inpatient	N (%)	N (%)	N (%)	N (%)
Outpatient	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
Baseline Sputum 1§ available	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
Sputum induced	N (%)	N (%)	N (%)	N (%)
Unknown	N (%)	N (%)	N (%)	N (%)
Sputum Xpert 1§ Positive	N (%)	N (%)	N (%)	N (%)
Sputum Culture 1§ Positive	N (%)	N (%)	N (%)	N (%)
Sputum Smear 1§ Positive	N (%)	N (%)	N (%)	N (%)
Baseline Urine 1§ available	N (%)	N (%)	N (%)	N (%)
Urine AlereLAM 1§ Positive	N (%)	N (%)	N (%)	N (%)
Urine FujiLAM 1§ Positive	N (%)	N (%)	N (%)	N (%)
Meta-analysis denominator (MAD) Positive	N (%)	N (%)	N (%)	N (%)
Meta-analysis denominator including LAM (MAD–LAM) Positive	N (%)	N (%)	N (%)	N (%)

[$]HIV-associated severe immunosuppression by WHO recommended age-appropriate CD4% or count cut-off: <12 months: <25%, 12-35 months: <20%, >36 months: <15%) or, in absence of CD4 % data, in terms of CD4 count (age <12 months: <1500 cells/mm^3, 12-35 months: <750 cells/mm^3, >36 months <350 cells/ mm^3)

§ the first baseline diagnostic sample collection attempt is defined as the first attempt to collect a urine or sputum sample within the first 2 days of including a patient in the study. Typically, the attempts to collect these samples are done on the day of enrolment.

Diagnostic yield

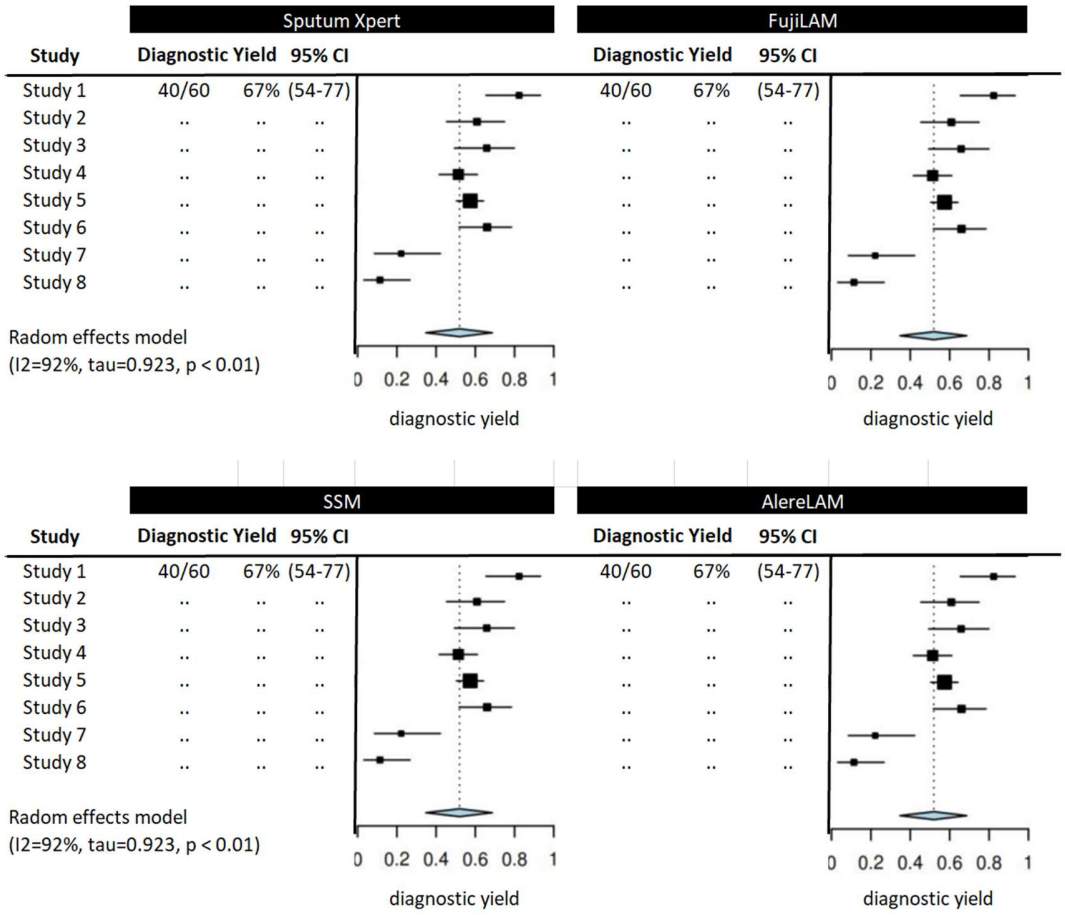
Point estimates (and 95% prediction intervals (PI) respectively confidence intervals (CI's)) of diagnostic yield for each test (sputum Xpert, SSM, urine AlereLAM, urine FujiLAM) and each study will be derived on the PP populations. The diagnostic yield of a test is the proportion of TB cases identified by this specific diagnostic test among the TB cases identified by the denominator. Diagnostic yield is defined as follows:

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

Whereas:
DY: diagnostic yield
PT: Number of positive Tests
MAD: Number of positive patients as defined by the harmonized Meta-Analysis Denominator (MAD).

Diagnostic yield for the different tests will be presented for all studies, for the different per protocol populations (PP's), and denominators in forest plots like in the following prototype figure:

Population: PP1 / Donominator MAD



Individual patient data meta-analysis
We will follow a one-stage IPD meta-analysis approach. (2) We will either perform a random or mixed-effects IPD meta-analysis to estimate the pooled diagnostic yield. We envision using a Generalized linear mixed model (GLMM) in the R packages (i.e. meta and lme4, likely function lmer) with further possible refinement and revision of the approach as needed. Random effect models account for heterogeneity across studies and mixed effect models partially account for heterogeneity across studies. In case of non-convergence of the model we will consider a two-stage approach.

Heterogeneity
Heterogeneity will be explored by visual inspection of forest plots and by using meta-regression on the following pre-specified covariates:

- Recruitment setting
- CD4 counts
- HIV status (if sufficient patient numbers)
- WHO TB symptom screen
- Proportion of patients with sputum available (defined below)
- Sputum provision (defined below)

Data checking, cleaning and merging

For each individual dataset, an overview analysis will be conducted to compare the information from the dataset with that in the original publication. If discrepancies are found which cannot be resolved, the database contributor will be contacted for clarifications. An overview table recording the presence or absence in each of the datasets of critical and secondary variables, relevant for the final analysis will be created.

During the data cleaning stage, variables will be renamed and grouped in a standardized way to enable merging across datasets. The individual datasets will be identified using indicator variables. Further indicator variables will be created to enable sub-group analyses (e.g. age group, setting, etc.).

Missing data and Data Imputation

The meta-analysis statistician will first identify the type of missing data (systematically (i.e., variables missing for an entire study dataset) and sporadically missing data (i.e. variables missing for certain patients within a study). We will train a logistic or linear regression model to predict missing values and impute them as appropriate.

Definition of subgroups

The following subgroups of study participants will be considered for a stratified analysis of the primary endpoint:

Sub-group	**Rational**	**Subgroup classification**
HIV status	Diagnostic yield is expected to be influenced by HIV status	-HIV+ -HIV- -Unknown
CD4 strata	To determine if LAM tests reach higher yields in patients with low CD4 counts (as is the case for AlereLAM)	-HIV+CD4 ≤ 100 -HIV+CD4 101-200 -HIV+ CD4 > 200 -HIV+ with unknown CD4 count -HIV- with unknown CD4 count
Sputum provision	Diagnostic yield of LAM-based tests is of particular interest in patients unable to expectorate sputum	-expectorated/non-induced first sample -induced first sample -Unknown
Recruitment Setting	Diagnostic yield and sample provision might be influenced by setting to which participants present for evaluation	-Inpatient -Outpatient -Unknown
WHO TB symptom screen	Diagnostic yield and sample provision might be influenced by current symptoms	-Symptomatic -Asymptomatic -Unknown
Proportion of patients with non-induced sputum available (Adults only)	Diagnostic yield is influenced by sample provision which could vary due to training and attempts by study staff	- studies where >50% of patients provided a non-induced sputum sample - studies where ≤50% of patients provided a non-induced sputum sample
Age group		< 2 years ≥ 2 years < 5 years ≥ 5 years <15 years ≥ 15 years <18 years ≥ 18 years

Denominator strength	There could be different yields depending on robustness of the denominator	- Unclear number of valid results from sputum-based Xpert or culture - ≤1 valid results from sputum-based Xpert or culture - ≥2 valid results from sputum-based Xpert or culture
Xpert version	Xpert Ultra has higher sensitivity than Xpert	-Sputum Xpert 1 using Xpert MTB/RIF -Sputum Xpert 1 using Xpert MTB/RIF Ultra

Sensitivity Analyses

Additional planned sensitivity analyses include:

- Including any LAM postive (either AlereLAM or FujiLAM positive) in the meta-analysis denominator (MAD)
- Excluding studies that didn't perform microbiological testing on samples other than sputum and urine
- Excluding studies with high or unclear risk of bias. I.e. studies with:
    - ≥2 risk of bias (RoB) questions answered as "high" or
    - 1 RoB questions answered as "high" and ≥1 RoB questions answered as "unclear" or
    - ≥2 RoB questions answered as "unclear"

Limitations

- Inferring diagnostic yield indirectly by assuming that the MAD represents all patients in the population with the disease and does not include false positives.
- The main analyses where only microbiologically confirmed TB is included in the denominator, will underestimate the diagnostic yield of LAM (i.e. in the studies with no extensive sampling) as patients with no Xpert/culture result (whether because they did not produce sputum or because these tests were not available) will be excluded from the analyses. The sensitivity analysis including LAM will overestimate the diagnostic yield of the LAM tests since all positive LAM results will be assumed to be true positives.
- Patients with high total burdens/disseminated TB usually have immunosuppression and poor ability to make sputum. So if a study wanted sputum (e.g. for reference standard microbiology) and excluded sputum scarce patients (by not doing induction) or by putting emphasis on sputum production, they are likely underestimating urine test yield.
- Sputum induction and for children NPA, GA, and GL might be very variable depending, in parts, on experience of study staff
- Xpert MTB/RIF and Xpert MTB/RIF Ultra will be combined and treated equivalent. However sub-group analysis is planned.
- Ziehl–Neelsen and Fluorescence Microscopy will be combined and treated equivalent.
- For adults only microbiologically confirmed TB will be deemed TB positive in the denominator. Possible TB or clinical definitions of TB will not be deemed positive because harmonisation of clinical definitions of TB across studies is not feasible. The MAD based on microbiologically confirmed TB might underdiagnose TB.

7. Statistical software

The analysis will be performed using the R statistical language (version 4.1.0 or higher), and Microsoft Excel 2017 (version 15.34 or higher) for a visual inspection of the data. The datasets and the code developed for the analysis will be versioned and will allow the reproducibility of the results at a later stage.

8. Timeline

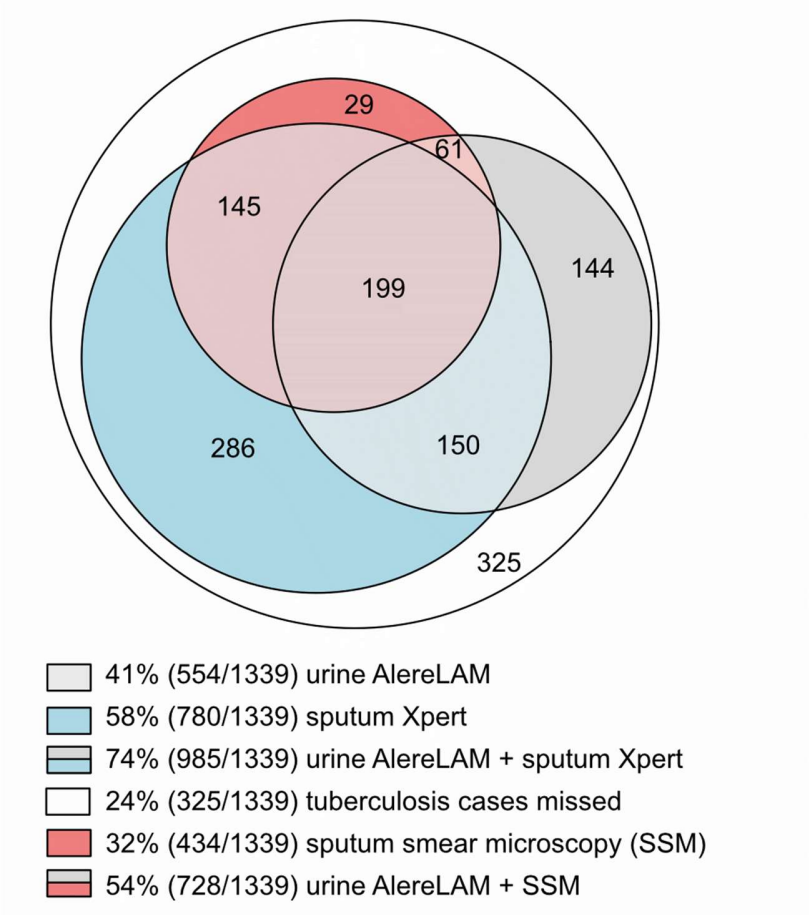
- Study datasets first deadline: 30.06.2021
- SAP draft sent to co-authors for comments: 31.07.2021
- Final study datasets shared last deadline: 31.08.2021
- Offer alignment meeting re SAP to co-authors: August 2021
- E-mails with co-authors to resolve data issues: August/September 2021
- Statistical analysis: September/October 2021
- Presentation of preliminary findings to co-authors: November 2021
- Presentation of draft manuscript to co-authors: November/December 2021
- Incorporation of feedback and submission of manuscript: December 2021

9. References

1. Graham SM et al. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update [Internet]. Clin. Infect. Dis. 2015;61(suppl 3):S179–S187.

2. Debray TPA et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology [Internet]. Res. Synth. Methods 2015;6(4):293–309.

Two day diagnostic yields in the subset of studies that assessed SSM



Two day diagnostic yield using the MAD in the subset of studies with results for all three tests SSM, urine AlereLAM and sputum Xpert. The MAD based on microbiologically confirmed tuberculosis was used as a denominator. AlereLAM=Alere Determine TB LAM Ag assay. Xpert=Xpert MTB/RIF or Xpert Ultra assay. SSM=sputum smear microscopy