Aus dem Zentrum für Infektiologie der Universität Heidelberg
Geschäftsführender Direktor: Herr Prof. Dr. med. Dr. h.c. Hans-Georg Kräusslich

Can Lipoarabinomannan Antigen Enable Point-of-Care TB Diagnosis? Exploring the Nexus of Clinical Concentration Ranges, Accuracy, and Diagnostic Yield

Inauguraldissertation
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)
an der
Medizinischen Fakultät Heidelberg
der Ruprecht-Karls-Universität

vorgelegt von
Tobias Broger

aus
Appenzell, Schweiz
2024

Dekan: Herr Prof. Dr. Michael Boutros

Doktormutter: Frau Prof. Dr. med. Claudia Maria Denkinger

Für Mija

CONTENTS

ABBREVIATIONS AND ACRONYMS

Abbreviation/ acronym	Meaning
ASSURED	Affordable, Sensitive, Specific, User-friendly, Rapid, Equipment-free, Delivered to end-users
AlereLAM	Alere Determine™ TB LAM (Abbott, Chicago, USA)
ART	Antiretroviral therapy
Ara4	Tetra-Arabinoside
Ara6	Hexa-Arabinoside
BMGF	Bill & Melinda Gates Foundation
CI	Confidence Interval
COVID-19	Coronavirus disease 2019
CrI	Credible interval
DMs	Decision makers
DNA	Deoxyribonucleic acid
DST	Drug susceptibility testing
ECL	Electrochemiluminescence
EclLAM	Laboratory-based ultrasensitive electrochemiluminescence LAM research assay
ELISA	Enzyme-linked immunosorbent assay
EPTB	Extra-pulmonary TB disease
EtD	Evidence-to-decision
FINDdx	Foundation for Innovative New Diagnostics
FujiLAM	Fujifilm SILVAMP TB LAM (Fujifilm, Tokyo, Japan)
GDG	Guideline Development Group
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HCPs	Health care providers
HIV	Human immunodeficiency virus
ICER	Incremental cost–effectiveness ratio
IPD	Individual participant data
IVD	*In vitro* diagnostic
LAM	Lipoarabinomannan

LED	Light-emitting diode
LFIA	Lateral flow immunoassay
LOD	Limit of detection
LMICs	Low- and middle-income countries
Man	Mannose
MGIT TTD	Mycobacterial growth indicator tube time to detection
MRS	Microbiological reference standard
Mtb	*Mycobacterium tuberculosis*
MTX	5-methylthio-d-xylofuranose
NAAT	Nucleic acid amplification test
NTM	Non-tuberculous Mycobacteria
PICO	Population, intervention, comparison, outcome
PLHIV	People living with the human immunodeficiency virus
POC	Point-of-care
QUADAS	Quality assessment of diagnostic accuracy studies
RCT	Randomized controlled trial
SDG	Sustainable development goals
SSM	Sputum smear microscopy
TB	Tuberculosis
TPP	Target product profile
uLAM	Urinary lipoarabinomannan
US	United States
USAID	United States Agency for International Development
WHO	World Health Organization
WRD	WHO-recommended rapid diagnostic test
WTP	Willingness to pay
Xpert	Xpert® MTB/RIF or Xpert® MTB/RIF Ultra (Cepheid, Sunnyvale, USA)

LIST OF TABLE AND FIGURES

1 PROLOGUE

In the meeting held on July 20, 2023, the Dr. sc. hum Doctoral Committee approved the submission of this publication-based cumulative dissertation. The results of this dissertation have already been published in three peer-reviewed articles. The first two publications are full-articles and I led study conception, design, setup, data collection, data processing and formal analysis. The third publication is a health policy review and guidance paper where I did the conception and literature research. For all three papers I wrote the first draft of the manuscripts and revised the manuscripts following co-author and peer review. Below is a detailed declaration of my contribution to the publications of this cumulative dissertation. The three publications of this publication-based dissertation are available in Chapter 3. A complete list of all my publications can be found in Chapter 8 - Personal publications and includes 18 first/senior author and 20 other publications. As of May 2024, my h-index is 21.

1.1 List of first-author publications relevant to the doctoral topic

1st Publication	**Form of Publication:**
Broger, T., Nicol, M. P., Sigal, G. B., Gotuzzo, E., Zimmer, A. J., Surtie, S., Caceres-Nakiche, T., Mantsoki, A., Reipold, E. I., Székely, R., Tsionsky, M., van Heerden, J., Plisova, T., Chikamatsu, K., Lowary, T. L., Pinter, A., Mitarai, S., Moreau, E., Schumacher, S. G., Denkinger, C. M., (2020). **Diagnostic accuracy of 3 urine lipoarabinomannan tuberculosis assays in HIV-negative outpatients**. Journal of Clinical Investigation *130 (11)*, 5756-5764, doi: 10.1172/JCI140461.	Original full-article, peer-reviewed, I am first-named shared first author (justification below), the Journal of Clinical Investigation is the number 2 of 139 journals in the area of Medicine, Research and Experimental, with an Impact Factor of 19.5 (2021 Journal Citation Report, Clarivate) Justification for shared co-first authorship: For the following reasons, I propose that the shared first authorship be considered equivalent to sole first authorship: This is a large international clinical multicentre study, where it is common to list the partner from the most important clinical centre (Prof. Dr. Mark Nicol from the University of Cape Town) as co-first author. I am listed in the first position, led the consortium, and had the conceptual and operational lead in the study and during the testing of the clinical samples.

2nd Publication **Broger, T**., Koeppel, L., Huerga, H., Miller, P., Gupta-Wright, A., Blanc, F. X., Esmail, A., Reeve, B. W. P., Floridia, M., Kerkhoff, A. D., Ciccacci, F., Kasaro, M. P., Thit, S. S., Bastard, M., Ferlazzo, G., Yoon, C., Van Hoving, D. J., Sossen, B., Garcia, J. I., Cummings, M. J., Wake, R. M., Hanson, J., Cattamanchi, A., Meintjes, G., Maartens, G., Wood, R., Theron, G., Dheda, K., Olaru, I. D., Denkinger, C. M. and Consortium. (2023). **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 *11 (6)*, e903-e916, doi: 10.1016/S2214-109X(23)00135-3.	**Form of Publication:** Original full-article, peer-reviewed, Individual participant data (IPD) meta-analysis, dataset with 10202 participants analysed with a Bayesian Model; I am first author; the Lancet Glob Health is the number 2 of 182 journals in the area of Public, Environmental & Occupational Health, with an Impact Factor of 38.9 (2021 Journal Citation Report, Clarivate)
3rd Publication **Broger, T**., Marx, F. M., Theron, G., Marais, B. J., Nicol, M. P., Kerkhoff, A. D., Nathavitharana, R., Huerga, H., Gupta-Wright, A., Kohli, M., Nichols, B. E., Muyoyeta, M., Meintjes, G., Ruhwald, M., Peeling, R. W., Pant Pai, N., Pollock, N. R., Pai, M., Cattamanchi, A., Dowdy, D. W., Dewan, P. and Denkinger, C. M. (2024). **Diagnostic yield as an important metric for the evaluation of novel tuberculosis tests – rationale and guidance for future research.** Lancet Glob Health (accepted).	**Form of Publication:** Health policy review and guidance paper. I am the first author, the Lancet Glob Health is the number 2 of 182 journals in the area of Public, Environmental & Occupational Health, with an Impact Factor of 38.9 (2021 Journal Citation Report, Clarivate)

1.2 Declaration of personal contribution to the publications

Procedure	**1st Publication**	**2nd Publication**	**3rd Publication**
Conception (%)	80%*	90%	90%
Literature research (%)	90%	75%	90%
Ethics request (%)	30%*	95%	Not relevant
(Animal experimentation permit) (%)	Not relevant	Not relevant	Not relevant
Experimental setup (%)	50%*	Not relevant	Not relevant
Data collection (%)	50%*	75%	Not relevant
Data evaluation (%)	90%	75%	Not relevant
Interpretation of results (%)	90%	90%	Not relevant
Writing of the manuscript text (%)	90%	90%	90%
Review and revision (%)	90%	90%	90%
Specify which illustrations/tables result from your doctoral thesis	Graphical Abstract Figure 1 Figure 2 Figure 3 Figure 4 Table 1 Table 2 Appendix	Figure 1 Figure 2 Figure 3 Table 1 Table 2 Box 1 Appendix	Figure 1 Panel 1 Figure 2 Figure 3 Panel 2 Table 1 Table 2

	1st Publication*	2nd Publication	3rd Publication
Contributions are described as follows in the original publication	**TB**, EIR, SGS, and CMD designed and oversaw the study. MPN, EG, SS, JVH, and TCN coordinated the individual study sites in South Africa and Peru. **TB**, GBS, MT, TP, TLL, AP, and EM contributed to EclLAM assay and reagent development. KC, RS, and SM coordinated measurement of AlereLAM and FujiLAM. AM and RS coordinated data collection and management. **TB** did the statistical analysis and **TB** and AJZ wrote the first manuscript draft. All authors contributed to interpretation of data and editing of the article and approved the final version of the manuscript. Authorship, including the order of co–first authors, was based on International Committee of Medical Journal Editors criteria.	**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.	**TB** and CMD wrote the first draft and all other authors contributed text, revised, commented and approved the final version of the manuscript. Authorship, including the order of authors was based on contribution and ICMJE criteria.

*The first publication is the result of a large, international, multi-centre clinical study. Parts of the work (i.e., conception, assay development, patient enrolment) began prior to the official start of the dissertation in March 2020 and were part of my long-term Tuberculosis research. The affiliation for all three publications is the University Hospital Heidelberg.

2 INTRODUCTION

2.1 The Tuberculosis pandemic

Tuberculosis (TB) caused 1.3 million deaths in 2022 and is the second leading infectious killer after COVID-19 as well as the 13th leading cause of death globally, despite being a preventable and curable disease (Vos et al. 2020; World Health Organization 2023a). Global targets for TB reduction have either been missed or remain off track, and if the SDG 2030 TB target are not achieved, an estimated 31.8 million TB deaths will likely occur from 2020 to 2050, corresponding to an economic loss of 17.5 trillion US Dollars (Silva et al. 2021).

2.2 The Tuberculosis diagnostic gap

A significant concern is the substantial number of people estimated to have TB who are currently not diagnosed or reported (Ismail et al. 2023). Of the estimated 10.6 million people who fell ill with TB in 2022 (incident cases), only 7.5 million were diagnosed, leaving a "TB diagnostic gap" of 3.1 million (29%) (World Health Organization 2023a). As shown in Figure 1, the COVID-19 pandemic has reversed years of progress made in the fight against TB and continues to have a damaging impact on access to TB services.

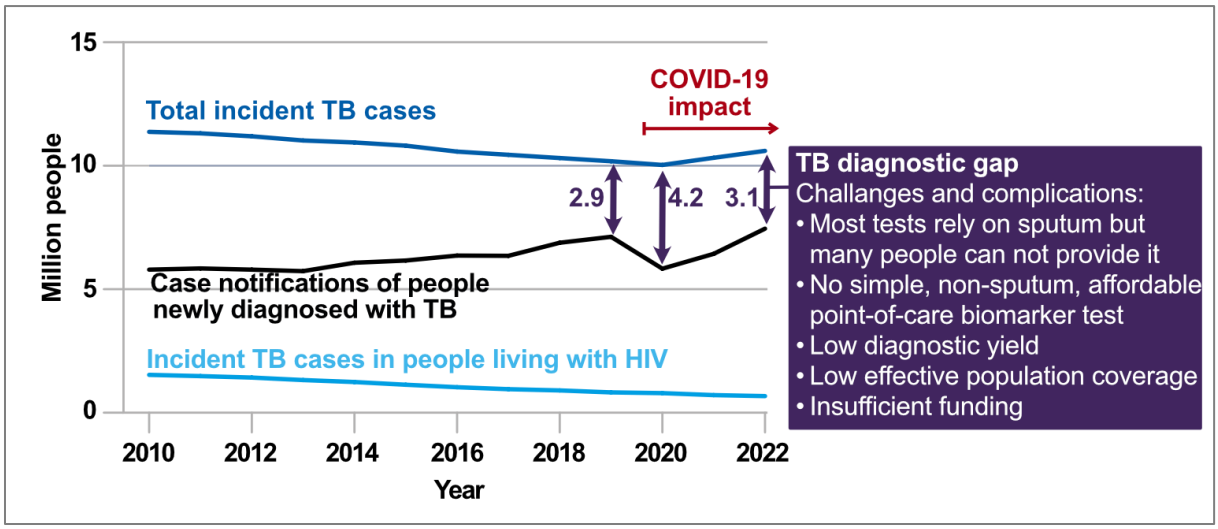

Figure 1: The TB diagnostic gap
The TB diagnostic gap is specified in millions of incident TB cases that were not diagnosed. Figure adapted from World Health Organization (2023a). TB=Tuberculosis. HIV=Human Immunodeficiency Virus.

Although the recent increase in case notification provides some hopeful evidence of a recovery in access to and provision of health services, COVID-19-related disruptions are estimated to have caused almost half a million excess deaths from TB in the three years from

2020-2022. The rebound of TB case notifications to beyond pre-COVID levels probably reflects diagnosis of both a sizeable backlog of people who developed TB in previous years but whose diagnosis was delayed due to COVID-19-related disruptions as well as an increase in the number of people falling ill with TB due to increased transmission (World Health Organization 2023a). There is an urgent need to increase bacteriological confirmation of TB by scaling up diagnosis. In 2022, only 47% (3.5 of the 7.5 million) of those newly diagnosed with TB used a WHO-recommended rapid diagnostic tests (WRDs). The use of WRDs remains far too limited due to the challenges and complications with sputum-based tests, low diagnostic yields, absence of effective diagnostic population coverage and insufficient funding.

2.3 Tuberculosis diagnostic challenges and complications

In 2014, the World Health Organization (WHO) called for the development of a "rapid biomarker-based non-sputum test capable of detecting all forms of TB by identifying characteristic biomarkers" and published a target product profile (TPP, (Denkinger et al. 2015; World Health Organization 2014)). Table 1 compares the TPP to currently available, frequently used and WHO-recommended TB tests and their fulfilment of the ASSURED (affordable, sensitive, specific, user-friendly, rapid, equipment-free, delivered to end-users) criteria. Mycobacterial culture is considered the reference standard but requires biosafety level 3 infrastructure that is typically only available in tertiary hospitals. Furthermore, culture takes 4 to 8 weeks until a diagnosis is available. Sputum smear microscopy (SSM) – a more than 100-year-old method – today only slightly improved with fluorescent dyes, remains the most widely used TB test in high-burden TB countries despite the well-known sensitivity limitations (Kik et al. 2014). Xpert MTB/RIF (Xpert, Cepheid, Sunnyvale, USA), a nucleic acid amplification test (NAAT), was the first WHO-recommended rapid diagnostic (WRD), endorsed in 2010, and revolutionised the diagnosis of TB (Ismail et al. 2023; World Health Organization 2020). Recently updated to the more sensitive Xpert MTB/RIF Ultra (Xpert Ultra), Xpert is a cartridge based NAAT for simultaneous rapid tuberculosis diagnosis and rapid antibiotic sensitivity testing to rifampicin, one of the most potent drugs for treating TB (Boehme et al. 2010; Dorman et al. 2018). This sputum-based test is simple to perform and produces results in approximately 90 minutes. Over the past decade, WHO has endorsed several similar molecular WRDs, including products suited to different contexts (Table 1, (World Health Organization 2020)). Culture, smear microscopy, and NAATs all use sputum which many people have difficulties to produce. Further, sputum is difficult to process; it is a heterogenous and usually viscous sample, all of which complicates the aforementioned diagnostic tests and therefore makes them expensive. The only non-sputum based test that received a WHO recommendation is the Alere Determine TB LAM Ag (AlereLAM, Abbott, Chicago, USA) based on the detection of mycobacterial lipoarabinomannan (LAM) antigen in urine (World Health Organization 2019). As further described in subsequent sections, the AlereLAM Lateral Flow Immunoassay (LFIA) is user-friendly, rapid, low-cost, and instrument free but has suboptimal

sensitivity which has limited its use to PLHIV where sensitivity and utility is highest (Bjerrum et al. 2019).

In summary, all these tests have major limitations and do not meet the minimum requirements of the WHO TPP, since they are either inaccurate, complex, expensive, based on sputum, or slow – and therefore inaccessible for many patients (Denkinger et al. 2015). As a result, only 47% of TB patients were tested with a WRD as an initial test in 2022 (Ismail et al. 2023; World Health Organization 2023a). To close the TB diagnostic gap, there is an urgent need to upscale WRDs as well as novel tests that come close to the WHO TPP.

The work presented herein investigates the relationship between clinical LAM concentration, test accuracy, and diagnostic yield and determines whether upscaling of the AlereLAM and development of an improved point-of-care LAM test could help find the “missing millions” of people with TB. Beyond that, this dissertation gives a comprehensive status update on 15 areas of research and development that have been pursued as part of the LAM development strategy that I defined in 2015 and further refined after a BMGF meeting with key experts in January 2016 in Seattle (USA) (Unpublished).

Table 1: Currently available and frequently used TB tests

Test (Product names)	**Principle**	**Healthcare system level**	**Specimen**	**WHO recom-mended**	**A** **Affordable**	**S** **Sensitive**	**S** **Specific**	**U** **User-friendly**	**R** **Rapid**	**E** **Equipment free**	**D** **Delivered to end-users**
Solid and Liquid Culture (BD MGIT)	Mycobacterial growth detection followed by speciation.	L3	Sputum	no	no	yes	yes	no	no	no	no
Sputum Smear Microscopy (SSM)	LED fluorescence microscopy after Auramine O staining	L1, L2, L3	Sputum	no	yes	no	yes	no	no	no	yes
Nucleic Acid Amplification Tests (NAATs) (Xpert MTB/RIF Ultra, Truenat, TB-LAMP, Abbott Realtime MTB, BD MAX MDR-TB, Cobas MTB, FluoroType MTB)	DNA detection with nucleic acid amplification	L2, L3	Sputum	yes	no	yes	yes	yes	no	no	no
Lipoarabinomannan Lateral Flow Immunoassay (LFIA) (Alere Determine TB LAM)	Detection of Lipoarabinomannan antigen	L0, L1, L2, L3	Urine	yes	yes	no	no	yes	yes	yes	no
Target Product Profile (TPP) Minimal Requirement		L1, L2, L3	Non-sputum		<$6	>65%	>98%	yes	<1 hour	yes	yes

Test criteria that meet the TPP requirements (specified at the bottom row in the table) are highlighted in green and those highlighted in red do not meet the criteria. For healthcare system levels, L0 = local community health post or care by local healthcare worker, L1 = microscopy centre or primary healthcare centre, L2 = district hospital or community health centre, L3 = reference or tertiary hospital or laboratory. Target product profile minimal requirements are based on Denkinger et al. (2015) and World Health Organization (2014). WHO recommended tests are based on World Health Organization (2020). ASSURED criteria are based on Mabey et al. (2004). LED= Light-emitting diode. DNA= Deoxyribonucleic acid. WHO=World Health Organization.

2.4 Lipoarabinomannan – a highly promising biomarker with potential to meet the target product profile

For a test to meet the TPP, it would need to be able to diagnose active TB and have a high specificity in order to allow initiation of antibiotic treatment at the same clinical encounter or on the same day (World Health Organization 2014). Such a point-of-care (POC) test would reduce diagnostic delays, help minimize TB transmission with appropriate therapy, address many of the current gaps in global TB control and ultimately decrease the TB diagnostic gap (Figure 1). To meet the TPP, a biomarker test would ideally be instrument free or feasible with limited instrumentation and would utilize easily accessible samples, such as blood, urine, saliva, swab, or exhaled breath (Bulterys et al. 2019). Non-DNA-based biomarker tests are more likely to meet the operational and cost targets of the TPP due to the availability of low-cost detection methodologies (such as LFIAs) compared to DNA-based tests. Therefore, TB biomarker research is an area of high activity, but its translational impact thus far has been limited. In 2019 (prior to this doctoral thesis), I co-led a large systematic review of biomarkers to detect active TB to identify the most promising lead candidates (MacLean et al. 2019). It included screening of 7631 reports and data extraction from 443 publications that fulfilled the inclusion criteria. Reports included host markers (blood cells or haematological markers, RNA, cytokines or chemokines, and antibodies) as well as pathogen markers (mycobacterial antigens, volatile organic compounds, pathogen RNA, *M. tuberculosis* metabolites and mycolic acids). LAM, a pathogen antigen, was the biomarker with the strongest high-quality evidence in the review. However, the review confirmed the already well-documented low sensitivity of urinary LAM (uLAM) detection methods in general non-HIV patient populations. The review concluded that "promising early-stage LAM studies suggest that LAM is present in general TB populations, including HIV-negative patients" but that the clinical performance of existing assays is hampered by their limited ability to detect sub-nanomolar ranges of LAM so that patients with low analyte concentrations are missed (MacLean et al. 2019). I proposed the development of better high-affinity antibodies which are essential for improved LFIAs, establishment of clinical concentrations ranges of LAM in non-sputum samples (i.e. urine), and investment in the development of point-of-care platforms able to detect LAM at concentrations of less than 10 pM (MacLean et al. 2019).

2.5 Urine Lipoarabinomannan antigen in active TB disease

2.5.1 Lipoarabinomannan

A detailed understanding of the LAM antigen was and remains essential to develop better antibodies and ultimately better performing tests. LAM has many attractive features as a biomarker – it is bacterially derived, highly-abundant in the cell wall of *Mycobacterium tuberculosis (Mtb)*, representing up to 15% of the bacterial mass, immunogenic, small enough to be filtered across the glomerular basement membrane into urine, and potentially has TB-

specific epitopes (Correia-Neves et al. 2019). These characteristics have helped inspire a quarter of a century of research on LAM's diagnostic potential. Despite these efforts, and those of basic scientists working to understand the structure and function of the glycolipid, many unanswered questions remain.

It is known that LAM is present in measurable quantities in the sputum, serum, and urine of different subgroups of TB patients (Brock et al. 2020; Broger et al. 2019b; Kawasaki et al. 2019; Sigal et al. 2018). Both HIV co-infected individuals who are severely immunocompromised and those with disseminated TB are particularly likely to have detectable levels of LAM in urine (Gupta-Wright et al. 2016). While the utility of LAM-based non-invasive diagnostics is especially welcome for this HIV population with a high-risk of TB-related morbidity and mortality, restriction of LAM testing to these populations limits the potential impact on transmission interruption and disease control in general populations. Therefore, what was not clear in 2016 is whether LAM's potential as a TB biomarker can be extended to the general symptomatic population with better reagents, detection platforms, or specimen processing.

Lipoarabinomannan (LAM) is a temperature-stable bacterial lipopolysaccharide produced by actinomycetes. Different forms of LAM and its precursors, lipomannan and the phosphatidylinositol mannosides (PIMs), are found in all mycobacterial species and are the most potent, non-peptidic molecules to modulate the host immune response (Besra and Brennan 1997). LAM is a major component of *Mtb*'s cell wall with multiple lipid-based molecules that create the bacteria's thick 'waxy' surface (Bulterys et al. 2019). The antigen is interspersed in the mycobacterial cell wall, firmly but non-covalently attached to the plasma membrane, and extends to the exterior of the cell wall, where it interacts as a potent virulence factor that modulates the host immune response and plays an important role in the pathogenesis of *Mtb* infection (Briken et al. 2004). The exact structure of *in vivo* LAM is unclear. *In vitro* LAM is heterogeneous in size, branching pattern, acylation, and phosphorylation on the arabinan and mannan portions with a peak molecular weight of 17.4 kDa and a reported size distribution of 6 kDa (Venisse et al. 1993).

LAM molecules have three structural domains (Figure 2): (i) the glycophospholipid anchor binds the molecule non-covalently to the plasma membrane, (ii) the attached mannan core is highly conserved across mycobacterial species, and (iii) the variable branching arabinan side chains with its variable mannose capping give rise to the diversity of LAM molecules (Lawn 2012). LAM can be classified into three major structural families according to the capping motifs (Guerardel et al. 2002):

a) ManLAM family (Figure 2A): The arabinan termini in the pathogenic, slow-growing strains (*Mycobacterium tuberculosis, Mycobacterium leprae, Mycobacterium avium, and Mycobacterium kansasii)* are modified with a single Man*p*, a dimannoside or a trimannoside.
b) PILAM family (Figure 2B) In the fast-growing, less pathogenic species (e.g. *Mycobacterium smegmatis, Mycobacterium fortuitum*), branches of the terminal arabinan are terminated by phospho-myo-inositol caps.

c) AraLAM family (not shown in Figure 2): LAM molecules devoid of both the manno-oligosaccharide and inositol phosphate caps have been identified in *Mycobacterium chelonae.*

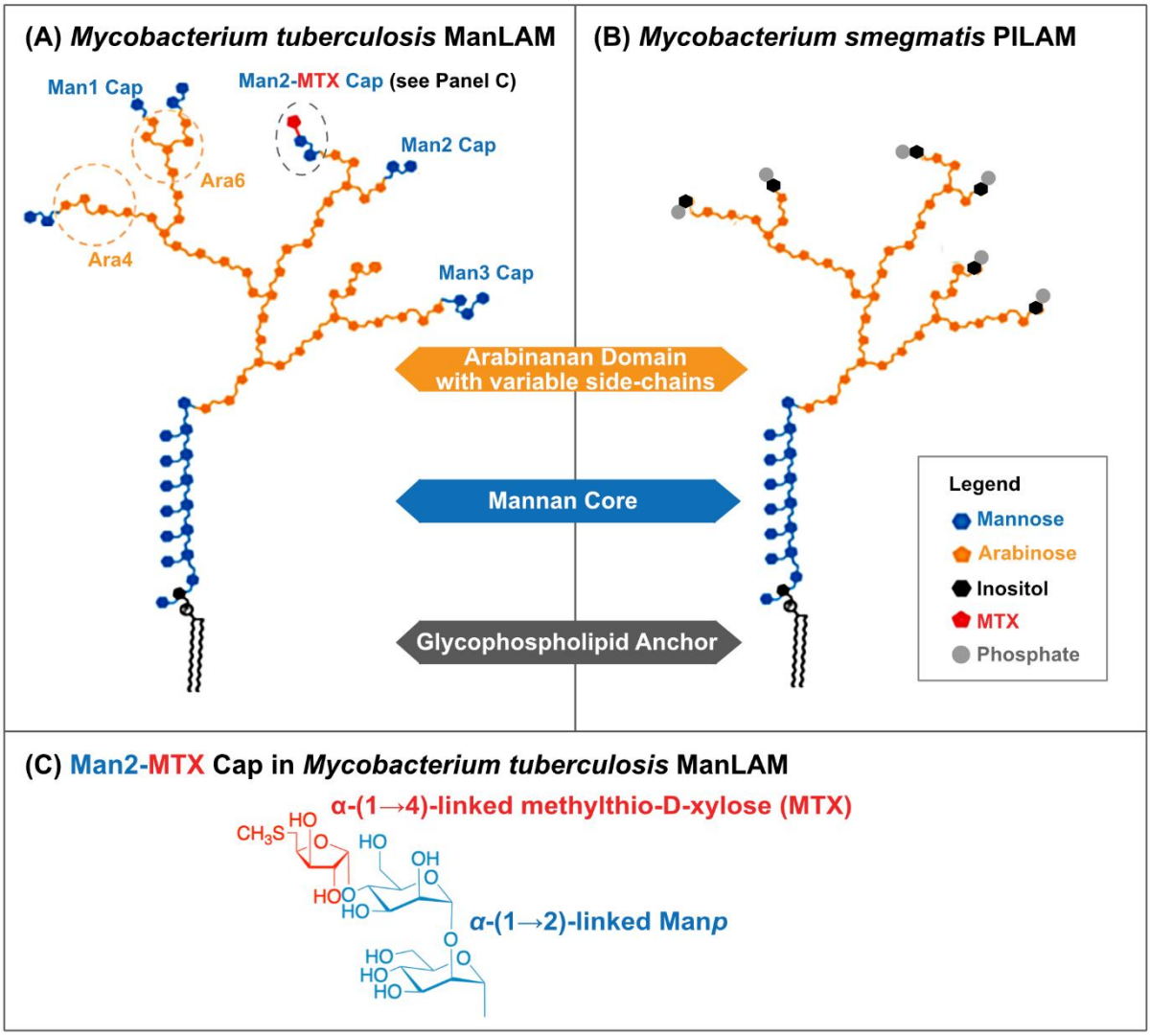

Figure 2: Lipoarabinomannan structure
(A) ManLAM as commonly found in Mtb and other pathogenic slow-growing mycobacteria with mannose caps. (B) PILAM as commonly found in rapid growing mycobacteria such as M. smegmatis with phospho-myo-inositol caps. (C) Representation of the nonreducing termini of Mtb ManLAM with two α-(1 →2)-Manp-linked residues giving rise to mannosylated LAM (ManLAM) further substituted with an α-(1→4)-linked methylthio-D-xylose (MTX) residue. Figure adapted from Angala et al. (2017); Fraaij et al. (2011); and Stalford et al. (2009).

AraLAM and PILAM have a strong proinflammatory effect within the human host whereas ManLAM molecules have potent anti-inflammatory properties (Lawn 2012). The proportion of LAM non-reducing termini capped with mannose varies among different species of slow-growing mycobacteria and among strains of *Mtb*, with fully virulent *Mtb* laboratory strains having up to 70% of mannose capping. Some of these are further modified by a single 5-methylthioxylofuranose (MTX) residue (Figure 2A with a detailed structure in Figure 2C) that played an essential role in my development of improved TB diagnostic tests, as will be

described below (Afonso-Barroso et al. 2013; Angala et al. 2017; Joe et al. 2006). Even within one species or strain, the degree of mannose capping may vary with growth conditions. Differences in mannose cap numbers between laboratory strains and clinical isolates are likely to affect ManLAM-associated biological functions and, particularly, their binding to C-type lectins (Nigou et al. 2003). *In vitro* ManLAM is predominantly associated with downregulation of host immunity through a number of pathways including suppression of *Mtb*-induced apoptosis (Rojas et al. 1997; Rojas et al. 2000), interleukin 12 secretion (Knutson et al. 1998), phagolysosomal fusion in macrophages (Fratti et al. 2001; Maruvada et al. 2004), maturation of dendritic cells (Geijtenbeek et al. 2002) and expansion of regulatory T cells (Garg et al. 2008). Unlike its structural analogues in avirulent mycobacteria, ManLAM is thought to contribute to the pathogenicity of virulent mycobacteria by undermining a protective host response and evading host immunity (Källenius et al. 2015; Sarkar et al. 2014; Torrelles et al. 2006). There is conflicting data about the contribution of mannose caps to mycobacterial virulence *in vivo* (Afonso-Barroso et al. 2013). While Pan and colleagues (2014) showed that aptamers binding ManLAM resulted in an inhibition of virulent *Mtb* infection in mice and rhesus monkeys, Afonso-Barroso et al. (2013) found that mutant BCG and *Mtb* strains lacking mannose caps showed similar growth in macrophages as wild type strains and did not show reduced virulence in experimentally infected mice. These data are not necessarily contradictory since virulence of *Mtb* might not be solely attributable to ManLAM immunomodulatory effects (Vergne et al. 2014). A comprehensive summary of the current knowledge on LAM structure and its ability to manipulate the endocytic pathway as well as phagocyte functions is given elsewhere (Vergne et al. 2014).

As replicating *Mtb* degrades, LAM circulating in the blood is filtered across the glomerular basement membrane of the kidneys and into urine, which lays the basis for LAM-based TB diagnosis (Bulterys et al. 2019). The presence of LAM in urine can also be a result of renal *Mtb* infection, as has been shown in autopsy studies (Cox et al. 2015).

2.5.2 First generation LAM tests

Urine LAM (uLAM) has been explored as a diagnostic biomarker for TB since 1997 (Hamasur et al. 2001). In 2003, Chemogen Inc. (Maine, USA) developed a urine ELISA for LAM detection, which was later commercialized by Inverness Inc. as the Clearview® TB ELISA assay (Lawn 2012). In 2010, Inverness became Alere and developed the first rapid POC LAM test: the Alere Determine TB LAM (AlereLAM) LFIA. In 2017, Alere was acquired by Abbott Diagnostics (Lake Bluff, USA). Based on the high observed agreement between the two assays, the AlereLAM test very likely uses the same polyclonal antibodies as its predecessor, the Clearview ELISA, (Lawn et al. 2012a).

AlereLAM meets most characteristics of the WHO TPP (affordable, user-friendly, rapid, equipment-free) for a biomarker-based, non-sputum TB test, but falls short on test accuracy, particularly in general populations without HIV infection (Table 1) (Bulterys et al. 2019). In a 2019 meta-analysis by Bjerrum et al. 2019 on HIV-positive patients, researchers compared

pooled sensitivity and specificity against a microbiological reference standard (MRS), defined as a positive TB culture or TB nucleic acid amplification test (NAAT). When compared against the MRS, pooled sensitivity and specificity were 42% (95% CI: 31–55%) and 91% (95% CI: 85–95%), respectively (Bjerrum et al. 2019). Among symptomatic participants with CD4 counts >200 cells/µL, pooled sensitivity and specificity were 16% (95% CI: 8–31%) and 94% (95% CI: 81–97%), respectively (Bjerrum et al. 2019). Among symptomatic participants with CD4 counts ≤200 cells/µL, pooled sensitivity and specificity were 45% (95% CI: 31–61%) and 89% (95% CI: 77–94%), respectively (Bjerrum et al. 2019). Among children, data were limited and ranged from 42 to 56% for sensitivity and 80 to 95% for specificity (Bjerrum et al. 2019). Based on this, the WHO updated the guidelines for use of AlereLAM in 2019 and currently recommends its use to assist in the diagnosis of active TB in PLHIV with symptoms of TB, those with advanced HIV diseases, those who are seriously ill, in inpatients with CD4 count <200 cells/µL irrespective of TB symptoms, and in outpatients with CD4 count <100 cells/µL (World Health Organization 2019). Despite WHO's recommendation and evidence from randomized controlled trials (RCTs) that the use of AlereLAM for TB diagnosis in PLHIV reduces tuberculosis-related mortality (Gupta-Wright et al. 2018; Peter et al. 2016), adoption and uptake of the test have been slow (Kraef et al. 2022; Singhroy et al. 2020).

2.5.3 Next-generation LAM tests

Between 2015 and 2019, when I was Senior Technical Officer at the Foundation for Innovative New Diagnostics (FINDDx, Geneva, Switzerland), and as part of the aforementioned LAM strategy, I led the development programme of the next-generation Fujifilm Silvamp TB LAM test (FujiLAM, Fujifilm, Tokyo, Japan). FujiLAM employs two high affinity monoclonal antibody pairs directed towards largely *Mtb*-specific LAM epitopes (i.e. MTX-Man2 and MTX-Man3, Figure 2C) and is an instrument-free POC test, with results available in less than one hour (Broger et al. 2019a). It uses a silver-amplification step that increases the visibility of the test and control lines, which facilitates an approximately 30 times lower cut-off than that of AlereLAM which uses conventional LFIA technology (Mitamura et al. 2013). In an individual participant data meta-analysis of the diagnostic accuracy of FujiLAM and AlereLAM in 1595 adult PLHIV, I showed that overall sensitivity for TB detection was 70.7% (CI 59.0%–80.8%) for FujiLAM compared to 34.9% (CI 19.5%–50.9%) for AlereLAM against a microbiological reference standard (MRS) (Broger et al. 2020c). Using the MRS, the specificity of FujiLAM was 90.9% (CI 87.2%–93.7%), and that of AlereLAM was 95.3% (95% CI 92.2%–97.7%).

2.6 Overarching research question for this dissertation: Is LAM a suitable biomarker for TB diagnosis in general populations?

In 2015, while working at the Foundation for Innovative New Diagnostics (FINDdx, Geneva, Switzerland) and with guidance from Prof. Dr. med. Claudia M. Denkinger (supervisor of this

thesis), Dr. Mark D. Perkins (Co-founder of FINDdx), and Dr. Jennifer L. Gardiner (former BMGF, responsible for TB biomarkers), I developed a systematic strategy towards answering the overarching question:

Is LAM a suitable biomarker for TB diagnosis in general populations?

This research strategy lays at the foundation of this thesis. To answer it, I embarked on a research programme to advance the LAM biomarker field towards a more sensitive, next generation of LAM tests by building on the existing AlereLAM assay. Studies clearly showed that a key reason for AlereLAMs slow uptake is its clinical utility which is limited to PLHIV (Mwaura and Engel 2021). Ideally, a next-generation LAM test would have broader utility, extending beyond individuals co-infected with HIV. This enhanced test could be applicable in general populations suspected of having TB, and potentially even for screening in people without TB symptoms. The strategy included work in four interlinked research streams on 15 areas as illustrated in Figure 3.

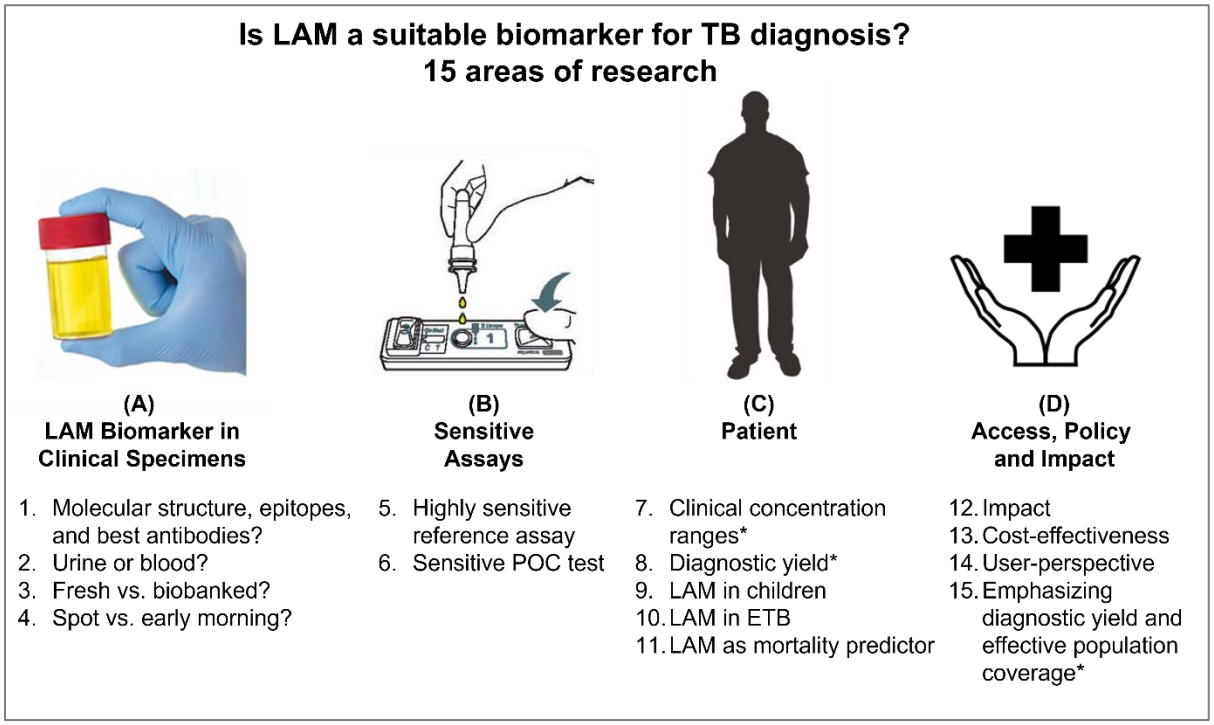

Figure 3: Is LAM a suitable biomarker for TB diagnosis – 15 areas of research and development
*Focus areas of this cumulative dissertation are areas 7, 8, and 15 (marked with *). LAM=Lipoarabinomannan. POC=point-of-care. ETB=extrapulmonary TB. TB=Tuberculosis.*

The first stream (A) focuses on the biomarker: research towards a better understanding of the LAM biomarker in clinical specimens. The second stream (B) focuses on the assay: development of highly sensitive LAM antigen detection assays. The third stream (C) focuses on the patient: research to better understand clinical LAM concentrations and performance in different patient populations. Finally, the fourth stream (D) focuses on the impact of a LAM-based assay: is work towards increasing access, estimating impact, and ensuring policy development. The strategy included rapid publication of research findings in open access

journals to create momentum around the LAM biomarker and catalyse other research activities and the development of next generation LAM tests. Several of the research areas have been addressed by my own research (total of 24 publications on the topic, listed in Chapter 8) or through others. The 15 research areas are addressed in the integrated discussion of this dissertation. The focus of this dissertation was to tackle research area 7 (clinical concentration ranges), 8 (diagnostic yield), and 15 (emphasizing diagnostic yield and effective population coverage) from Figure 3, which led to three peer-reviewed publications in high-impact journals that form the core of this cumulative dissertation. Addressing these three interconnected areas contributes significantly to the overarching research question:

- Firstly, understanding the range of clinical concentrations in general populations (as explored in Publication 1) enables the prediction of accuracy and diagnostic yield for next-generation LAM tests and paves the way for their development.
- Secondly, the distinction between diagnostic yield and sensitivity, researched in Publication 2, is vital as diagnostic yield uniquely benefits from the higher availability of urine samples compared to sputum, which is often difficult for many patients to produce. Better specimen availability is one of the main reasons for the push to develop non-sputum based tests as articulated in the WHO TPP.
- Lastly, the full potential of novel, innovative non-sputum tests, such as uLAM tests, hinges on recognizing the significance of diagnostic yield in the WHO guideline development process, particularly in terms of improved sample availability (Publication 3).

My research employed a diverse array of methodologies:

- **In publication 1,** for investigating LAM concentration ranges, a clinical multi-centre diagnostic accuracy study was conducted in South Africa and Peru, both high TB burden countries. To determine LAM concentrations, an improved version of a highly sensitive quantitative LAM immunoassay was used, and the assay was benchmarked against the WHO-recommended existing AlereLAM POC test and a more sensitive FujiLAM POC test, the next generation LAM test furthest along in the development.
- **In publication 2,** to evaluate the diagnostic yield of urine LAM against sputum-based TB tests, I conducted an extensive, systematic Individual Patient Data (IPD) meta-analysis, integrating datasets from existing TB studies. The main outcomes were the tuberculosis diagnostic yields of urine lipoarabinomannan tests, sputum NAATs, and sputum smear microscopy (SSM). Diagnostic yields were predicted using Bayesian random-effects and mixed-effects.
- **In publication 3,** which advocates for a greater focus on diagnostic yield in policy-making, I collaborated and reached a consensus with a large group of eminent TB diagnostic experts. I reviewed the history of TB policy development and formulated a comprehensive proposal to incorporate and evaluate diagnostic yield in this process.

The following sections contextualize these three research areas within the current scientific landscape in their respective fields. However, in the integrated discussion, I chose to not just

cover these three core questions but to discuss all 15 areas of Figure 3 to provide a thorough overview of the progress of my own and other's research. With this, I aim to enable readers to better understand the interconnectedness of individual publications. By discussing the interrelations and defining next steps, I aim to further advance the LAM field.

2.6.1 Research area No. 7 addressed in publication 1 – LAM clinical concentrations ranges

As described above, the first generation AlereLAM test reaches 42% sensitivity in HIV-positive patients, with higher sensitivity in immunocompromised people, but only 16% sensitivity in immunocompetent (defined as having a CD4 counts >200 cells/µL) people living with HIV (PLHIV) (Bjerrum et al. 2019). AlereLAM's cut-off for uLAM concentrations is around 0.5 ng/ml which suggests that more than four out of five PLHIV with TB have concentrations below that cut-off. In 2015 when I embarked on the LAM strategy, there was only very limited information on LAM concentrations in different specimens from people with TB, owing to the absence of quantitative reference assays. In particular there was no information on uLAM concentrations in HIV-negative people with TB. However, it is essential to understand clinical concentration ranges of an antigen biomarker as they directly define whether development of a more sensitive LAM point-of-care test is feasible. Classical LFIAs, like AlereLAM use colloidal gold as a label and reach cut-offs around 0.5 ng/ml (Nakiyingi et al. 2015). However, more sensitive LFIAs with lower cut-offs have been developed in other fields, like malaria diagnosis using malaria histidine-rich protein II (HRP2), or to detect low pg/ml amounts of troponin in patients with suspected acute coronary syndrome (Das et al. 2017). Yet, it was unclear what the cut-off requirement would be for a next-generation LAM test to achieve WHO's TPP target of >65% sensitivity in broad populations, i.e. in HIV-negative as well as HIV-positive people. This limited information on LAM concentrations "below the tip of the iceberg" detectable by first generation AlereLAMs leaves developers of next-generation LAM tests in the dark regarding the necessary cut-off and its relationship to diagnostic sensitivity (Figure 4). This uncertainty underscores the urgent need for further research to establish these clinical LAM concentrations in different patient populations.

To shed light on a more appropriate cut-off value, and as part of research area 1 (molecular structure, epitopes, and best antibodies, Figure 3), I led the development of a highly sensitive reference assay for LAM detection (research area 5; Sigal et al. 2018b, work done prior to this dissertation). It included screening of 100 monoclonal antibody pairs, targeting various LAM epitopes, on a sensitive electrochemiluminescence (ECL) platform to enhance diagnostic accuracy. Antibody screening was done directly with urine from TB patients as *in vitro* LAM is structually different from *in vivo* uLAM. The top twelve antibody pairs from the initial screening underwent testing in a retrospective case-control study with biobanked urine samples from 75 adults presumed to have TB. Using the optimal antibody pair, the laboratory-based ultrasensitive electrochemiluminescence LAM research assay (EclLAM) achieved a limit of detection (LOD) of 11 pg/ml and demonstrated an overall clinical sensitivity of 93% (95% CI,

80%–97%) and specificity of 97% (95% CI, 85%–100%). Notably, in HIV-negative subjects who were TB-positive, the test attained a sensitivity of 80% (CI, 55%–93%), surpassing the WHO TPP target of 65%. This contrasts with AlereLAM's overall sensitivity of 33% (CI, 20%–48%) in all patients and sensitivity of 13% (CI, 4%–38%) in HIV-negative subjects, based on the same sample set which is close to the sensitivity of 16% in immunocompetent HIV-positive patients (above and Figure 4).

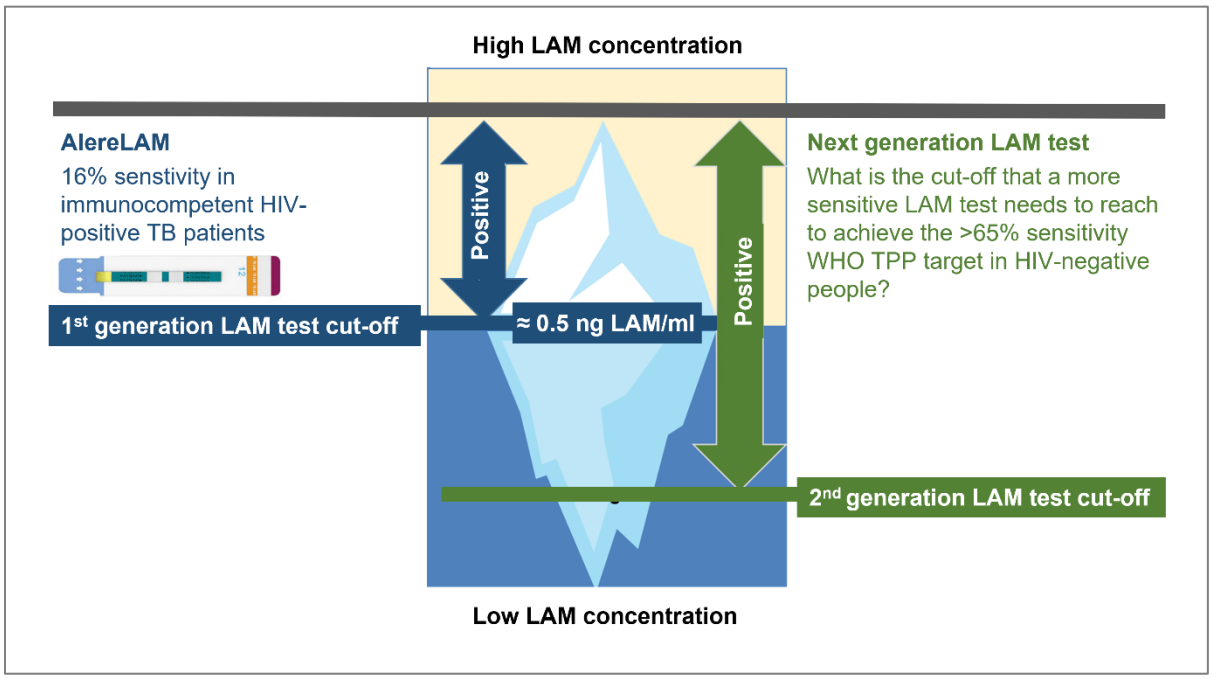

Figure 4: Relationship between LAM clinical concentration and test sensitivity
The AlereLAM test comes back TB positive for uLAM concentrations above ≈ 0.5 ng LAM/ml and in 16% of immunocompetent TB patients. It is unclear how low the cut-off of a more sensitive LAM antigen test needs to be to achieve the required WHO TPP target of >65% sensitivity in general populations including HIV-negative people. AlereLAM=Alere Determine TB LAM assay. LAM=lipoarabinomannan. TB=Tuberculosis. HIV=human immunodeficiency virus. TPP=Target product profile. WHO=World Health Organization.

While this initial study suggested that a test with a cutoff at 11 pg/ml could diagnose 80% of HIV-negative individuals, it has major limitations affecting the generalizability of the findings. The study's relatively small size and case-control design might have led to an overestimation of sensitivity due to selection bias and the focus on smear-positive TB patients with a higher bacterial load in sputum. The paper concluded that accuracy and clinical concentration should be assessed in adequately powered blinded cohort studies within relevant populations where the test would be clinically indicated for TB diagnosis with a further optimized reference assay, which I did in Publication 1 of this dissertation.

Publication 1 further aimed to assess the performance of the novel, more sensitive POC test, the FujiLAM, that I developed in close collaboration with Fujifilm prior to this dissertation (Broger et al. 2019a). All initial studies were done in PLHIV, therefore, Publication 1 also addressed research area 6 in Figure 3 regarding the question on whether a LAM POC test that meets the WHO TPP in general populations including HIV-negatives can be developed.

2.6.2 Research areas No. 8 addressed in publication 2 – Diagnostic yield of urine LAM and sputum Xpert

Diagnostic yield has become a metric in evaluating the utility of various diagnostic tests and procedures across different medical specialties. Many of the initial publications focused on cancer screening (Corral et al. 2019), but it is increasingly used in infectious diseases as well.

Sputum has been used to diagnose tuberculosis for over a century and is the most used sample type (Table 1). However, sputum can be difficult to obtain, particularly in PLHIV, 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 (Gupta et al. 2015). As laid out in detail above, the WHO TPP encourages the development of non-sputum biomarker tests with the ultimate aim of enabling rapid TB treatment initiation to address these gaps. Yet, diagnostic accuracy, used in the majority of studies that inform WHO guidelines, 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 (Ricks et al. 2020; Ryckman et al. 2022). 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 as well as the sensitivity of the test. In Publication 2 of this dissertation, I aimed to do an individual participant data (IPD) meta-analysis to compare the TB diagnostic yield of three tests – 1) uLAM POC tests on the first available urine sample, 2) sputum NAATs on the first available sputum sample, and 3) SSM on the first available sputum sample within 2 days of enrolment – against a harmonised denominator. To my knowledge, this is the first IPD meta-analysis assessing diagnostic yield of TB tests.

2.6.3 Research area No. 15 addressed in publication 3 – Emphasizing diagnostic yield and effective population coverage

During the writing and peer-review of publication 2, it became clear to me and others (Pai et al. 2023) that better explanation and consideration of diagnostic yield in WHO's guideline development process is of paramount importance as otherwise new, potentially highly impactful non-sputum TB tests would be undervalued. In publication 3, a health policy review and guidance paper, I emphasize that while better TB tests are required, an exclusive focus on improving accuracy alone will not be sufficient to close the diagnostic gap for TB.

The example in Figure 5, that is based on the meta-analysis from publication 2, illustrates how an accessible test with moderate sensitivity performed using an easily obtainable non-sputum specimen like urine could be more useful than a sensitive test reliant on sputum.

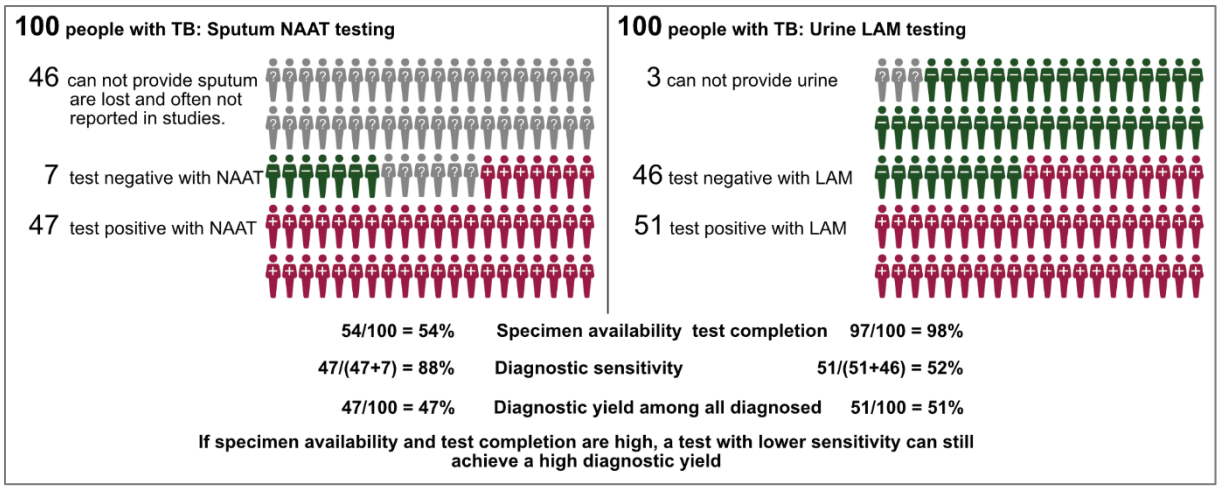

Figure 5: Comparison of diagnostic yield and sensitivity

Example based on the diagnostic yield meta-analysis from publication 2 in hospitalized people living with HIV that often have difficulty expectorating sputum. An accessible test with moderate sensitivity performed using an easily obtainable non-sputum specimen like urine can reach similar yields like a sensitive NAAT reliant on sputum. NAAT=Nucleic Acid Amplification Test. TB=Tuberculosis. LAM=Lipoarabinomannan.

In the example, the AlereLAM test only reaches a sensitivity of 52% but is feasible for 98% of people, leading to a diagnostic yield among all diagnosed of 51% with a simple POC antigen test within 25 minutes. In contrast, the sputum Xpert NAAT, with a much higher sensitivity of 88% is only possible for half of the patients since many of the hospitalized PLHIV cannot provide sputum, leading to a yield of 47% with this more complex and costly NAAT that takes 90 minutes. Diagnostic yield is especially relevant in populations and settings where people often cannot provide sputum. Examples include PLHIV, especially those with advanced disease, children, and minimally symptomatic or asymptomatic individuals undergoing TB screening in high prevalence settings. Timely specimen availability is a critical aspect related to yield. For example, if sputum cannot be obtained, sputum induction might be necessary which requires skill, instruments, and time, resulting in in delays in obtaining a specimen. Thus, impaired yield also negatively affects time to diagnosis – yet despite the challenges around its collection, sputum is still the most common TB diagnostic specimen.

Currently, policy makers like the World Health Organization (WHO), diagnostic test regulators, test manufacturers, and the scientific community are almost exclusively focused on test sensitivity and specificity as the main proxy of clinical utility and patient-important outcomes. Critically, this approach neglects key utility considerations such as patient/population accessibility, specimen obtainability, successful test completion, and whether test results are available in time to guide treatment and public health decision making. The poor availability of sputum is usually not captured in test performance assessments and many TB studies only enrol sputum producers, but there continues to be very little awareness on the influence of this patient-selection bias in nearly all TB diagnostic performance evaluation studies.

Since 2007, the World Health Organization (WHO) has been applying the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) process for guideline

development (Schünemann et al. 2008). GRADE prioritizes randomized controlled trials (RCTs) that directly evaluate the impact of a diagnostic test on patient-important outcomes under 'real-life' conditions (Schünemann et al. 2019). However, in the absence of direct evidence from RCTs, WHO's guideline development groups (GDGs) usually link accuracy studies to patient-important outcomes, such as cure, mortality, time to diagnosis, and time to treatment, and integrate these into GRADE's evidence to decision (EtD) framework to infer likely impact of tests and develop recommendations (Schünemann et al. 2016). If a test is not likely to improve patient-important outcomes or population health, then the healthcare system has no reason to use it, whatever its accuracy may be (Schünemann et al. 2008). The objective of publication 3 is to emphasize diagnostic yield and effective population coverage for policy development. In addition to the suggested refinements in health policy development, the paper provides clear definitions for diagnostic yield including an Excel calculator, lessons learned from other diseases, guidance on how to design and report diagnostic yield studies, and a comprehensive discussion of the strengths and limitations of diagnostic yield.

3 PUBLICATIONS

3.1 Publication 1 - Diagnostic accuracy of 3 urine lipoarabinomannan tuberculosis assays in HIV-negative outpatients

Diagnostic accuracy of 3 urine lipoarabinomannan tuberculosis assays in HIV-negative outpatients

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

[1]FIND, Geneva, Switzerland. [2]Division of Tropical Medicine, Center for Infectious Diseases, Heidelberg University Hospital, Heidelberg, Germany. [3]Division of Infection and Immunity, School of Biomedical Sciences, University of Western Australia, Perth, Australia. [4]Division of Medical Microbiology, University of Cape Town, Cape Town, South Africa. [5]Meso Scale Diagnostics LLC, Rockville, Maryland, USA. [6]Cayetano Heredia University, Lima, Peru. [7]Department of Mycobacterium Reference and Research, Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Tokyo, Japan. [8]Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada. [9]Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan. [10]Public Health Research Institute Center, New Jersey Medical School, Rutgers University, Newark, New Jersey, USA.

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

METHODS. In this multicenter diagnostic test accuracy study, we recruited HIV-negative adults with symptoms suggestive of pulmonary TB presenting to outpatient health care centers in Peru and South Africa. Urine samples were tested using FujiLAM, AlereLAM, and EclLAM, and the diagnostic accuracy was assessed against a microbiological reference standard (MRS) and a composite reference standard.

RESULTS. Three hundred seventy-two HIV-negative participants were included and the prevalence of microbiologically confirmed TB was 30%. Compared with the MRS, the sensitivities of AlereLAM, FujiLAM, and EclLAM were 10.8% (95% confidence interval [CI] 6.3%–18.0%), 53.2% (95% CI 43.9%–62.1%), and 66.7% (95% CI 57.5%–74.7%), respectively. The specificities of AlereLAM, FujiLAM, and EclLAM were 92.3% (95% CI 88.5%–95.0%), 98.9% (95% CI 96.7%–99.6%), and 98.1% (95% CI 95.6%–99.2%), respectively. Positive likelihood ratios of AlereLAM, FujiLAM, and EclLAM were 1.4, 46.2, and 34.8, respectively, and positive predictive values were 37.5%, 95.2%, and 93.7%, respectively.

CONCLUSION. Compared with AlereLAM, FujiLAM detected 5 times more patients with TB in HIV-negative participants, had a high positive predictive value, and has the potential to improve rapid diagnosis of TB at the point-of-care. EclLAM demonstrated that additional sensitivity gains are possible, which highlights LAM's potential as a biomarker. Additional research is required to assess FujiLAM's performance in prospective cohorts, its cost-effectiveness, and its impact in real-world clinical settings.

FUNDING. Global Health Innovative Technology Fund, the UK Department for International Development, the Dutch Ministry of Foreign Affairs, the Bill and Melinda Gates Foundation, the Australian Department of Foreign Affairs and Trade, the German Federal Ministry of Education and Research through Kreditanstalt für Wiederaufbau, and the NIH and National Institute of Allergy and Infectious Diseases.

▶ Related Commentary: p. 5671

Authorship note: TB and MPN contributed equally and are co-first authors. EM, SGS, and CMD contributed equally.
Conflict of interest: TB, AM, EIP, RS, EM, SGS, and CMD were previously or are currently employed by FIND. GBS, MT, and TP are employed by Meso Scale Diagnostics LLC and received funding from FIND. AP is an author on patents in the field of lipoarabinomannan (LAM) detection, "Novel anti-lam and anti-pim6/lam monoclonal antibodies for diagnosis and treatment of mycobacterium tuberculosis infections" (US20190038747) and "Methods for dual detection and differentiation of infection by mycobactierum tuberculosis complex and nontuberculous mycobacteria" (WO2020018806). TB is an author on a patent in the field of LAM detection, "Antibody or antibody combination and method using same for detection of an antigen related to mycobacterium in a urine sample of a subject" (WO2019186486).
Role of funding source: The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Submitted: May 20, 2020; **Accepted:** July 16, 2020; **Published:** September 28, 2020.
Reference information: *J Clin Invest*. 2020;130(11):5756–5764. https://doi.org/10.1172/JCI140461.

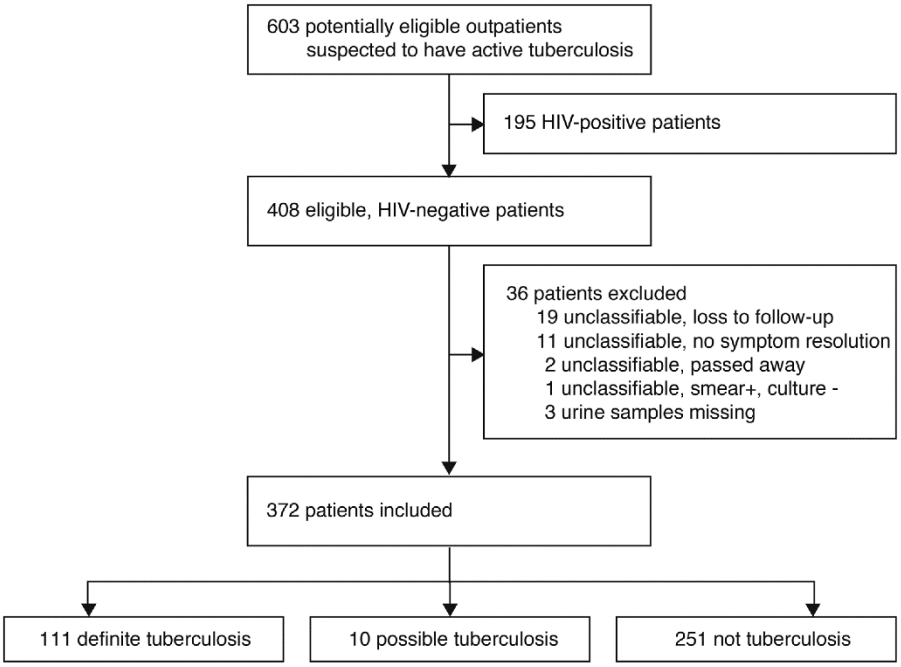

Figure 1. Study flow diagram.

Introduction

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

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

Fujifilm recently developed a next-generation POCT, the Fujifilm SILVAMP TB LAM test (FujiLAM). To improve sensitivity while maintaining high specificity, FujiLAM uses a pair of high-affinity monoclonal antibodies selected to detect LAM presenting the *Mycobacterium tuberculosis*–specific (*Mtb*-specific) 5-Methylthio-D-xylofuranose epitope (MTX-LAM), and employs a silver-amplification step (16–18). A recent meta-analysis of 1595 HIV-positive inpatients and outpatients confirmed FujiLAM's superiority, demonstrating a sensitivity of 71%, twice that of AlereLAM (19). Further, FujiLAM showed good sensitivity for the detection of extrapulmonary TB (EPTB) ranging from 47% to 94% across different forms of ETB (20) and could have rapidly diagnosed TB in up to 89% of HIV-positive inpatients who died within 12 weeks (21).

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

Results

Between February 9, 2017, and October 4, 2017, 603 potentially eligible participants were screened. A total of 408 HIV-negative participants met inclusion criteria and 372 were included in the analyses (Figure 1). Of these, 30% (111/372) were classified as definite TB, 3% (10/372) as possible TB, and 67% (251/372) as not TB (Table 1). Prevalence of definite TB was higher in Peru (43%) compared with South Africa (17%). Most participants were young adults (median age 32 years) and 14% had a history of prior TB disease. In participants with definite TB, 68% (76/111) had at least one positive fluorescence sputum smear microscopy (SSM) result. Peruvian participants with TB had shorter mycobacterial growth indicator tube (MGIT) liquid culture time to detection and

Table 1. Demographic and clinical characteristics of the study participants

	All participants (*N* = 372)	Cape Town, South Africa (*n* = 187)	Lima, Peru (*n* = 185)
Demographic or clinical characteristic			
Median age, years (IQR)	32 (25–47)	33 (26–48)	31 (24–42
Sex			
Female, no. (%)	158/372 (42)	96/187 (51)	62/185 (34)
Male, no. (%)	214/372 (58)	91/187 (49)	123/185 (66)
History of TB, no. (%)	52/372 (14)	22/187 (12)	30/185 (16)
Distribution in diagnostic categories			
Definite TB, no. (%)	111/372 (30)	32/187 (17)	79/185 (43)
Possible TB, no. (%)	10/372 (3)	1/187 (1)	9/185 (5)
Not TB, no. (%)	251/372 (67)	154/187 (82)	97/185 (52)
Reference standard			
MRS positive, no. (%)	111/372 (30)	32/187 (17)	79/185 (43)
CRS positive, no. (%)	121/372 (33)	33/187 (18)	88/185 (48)
Distribution in MRS-positive patients			
Sputum smear microscopy positive (any of 3 tests on 3 sputum samples positive), no. (%)	76/111 (68)	15/32 (47)	61/79 (77)
Sputum smear microscopy positive (1 test on first sputum sample), no. (%)	68/111 (61)	12/32 (38)	56/79 (71)
Blood culture positive, no. (%)	0/111 (0)	0/32 (0)	0/79 (0)
Urine Xpert positive, no. (%)	4/111 (4)	0/32 (0)	4/79 (5)
Sputum Xpert positive (any of 3 tests on 3 sputum samples positive), no. (%)	102/111 (92)	27/32 (84)	75/79 (95)
Sputum Xpert positive (1 test on first sputum sample), no. (%)	82/111 (74)	22/32 (69)	60/79 (76)
Sputum Xpert result (from testing on first sputum sample, Xpert was repeated once using the same sample in case of an indeterminate result)			
Negative, no. (%)	25/111 (23)	10/32 (31)	15/79 (19)
Very low, no. (%)	19/111 (17)	4/32 (13)	15/79 (19)
Low, no. (%)	32/111 (29)	6/32 (19)	26/79 (33)
Medium, no. (%)	28/111 (25)	7/32 (22)	21/79 (27)
High, no. (%)	7/111 (6)	5/32 (16)	2/79 (3)
Average sputum MGIT time to detection			
MGIT negative (only Xpert positive), no. (%)	6/111 (5)	3/32 (9)	3/79 (4)
Greater than 14 days, no. (%)	40/111 (36)	13/32 (41)	27/79 (34)
0–14 days, no. (%)	65/111 (59)	16/32 (50)	49/79 (62)

a larger proportion of patients had positive SSM and Xpert results compared with patients from South Africa, suggesting higher mycobacterial load in sputum (Table 1). None of the patients with definite TB had a positive blood culture and only 4% (4/111) had a positive urine Xpert result, with all of the latter also having positive sputum Xpert and culture results.

The index test failure rate number was very low (1 repeat FujiLAM and no repeats for AlereLAM). Results of the diagnostic accuracy of urine LAM-based assays (AlereLAM, FujiLAM and EclLAM), sputum-based assays (Xpert and SSM), and combinations of these assays are shown in Figure 2 and in Supplemental Figure 3 (supplemental material available online with this article; https://doi.org/10.1172/JCI140461DS1). Overall, compared with the microbiological reference standard (MRS), the sensitivities of AlereLAM, FujiLAM, and EclLAM were 10.8% (12/111; 95% confidence interval [CI] 6.3%–18.0%), 53.2% (59/111; 95% CI 43.9%–62.2%), and 66.7% (74/111; 95% CI 57.5%–74.7%), respectively (Figure 1). The sensitivity of urine FujiLAM and sputum Xpert in combination was 82.0% (91/111; 95% CI 73.8%–88.0%) and the sensitivity of urine FujiLAM in combination with a single SSM was 70.3% (78/111; 95% CI 61.2%–78.0%) and higher than Xpert alone (Figure 2, B and C). Using the composite reference standard (CRS), the sensitivities of all assays were not substantially changed (FujiLAM 48.8%, EclLAM 62.0%, AlereLAM 12.4%) (Supplemental Table 3).

All tests, except AlereLAM, reached specificities of 98% or higher in the MRS-based analysis (AlereLAM 92.3% (241/261; 95% CI 88.5%–95.0%), FujiLAM 98.9% (258/261; 95% CI 96.7%–99.6%), EclLAM 98.1% (256/261; 95% CI 95.6%–99.2%). When comparing results from the CRS-based analysis to the MRS-based analysis, specificity remained largely unchanged for all assays (AlereLAM 93.2 [234/251], FujiLAM 98.8% [248/251], EclLAM 98.4% [247/251]) (Supplemental Table 3). Against the MRS at study prevalence of 30%, the positive predictive values (PPVs) of AlereLAM, FujiLAM, and EclLAM were 37.5%, 95.2%, and 93.7%, respectively. When assuming a lower pretest probability of 20%, the PPV of AlereLAM, FujiLAM and EclLAM were 26.1%, 92.0%, and 89.7% respectively (Table 2). Positive likelihood ratios (LR+) were 46.2 for FujiLAM and 1.4 for AlereLAM. When assuming 20% pretest probability, the negative predictive values (NPVs) of AlereLAM, FujiLAM, and EclLAM were 80.5%, 89.4%, and 92.2%, respectively (Table 2). Negative likelihood ratios (LR–) were 0.47 for FujiLAM and 0.97 for AlereLAM. Fagan nomograms illustrating pretest and posttest probabilities are available in Supplemental Figure 4.

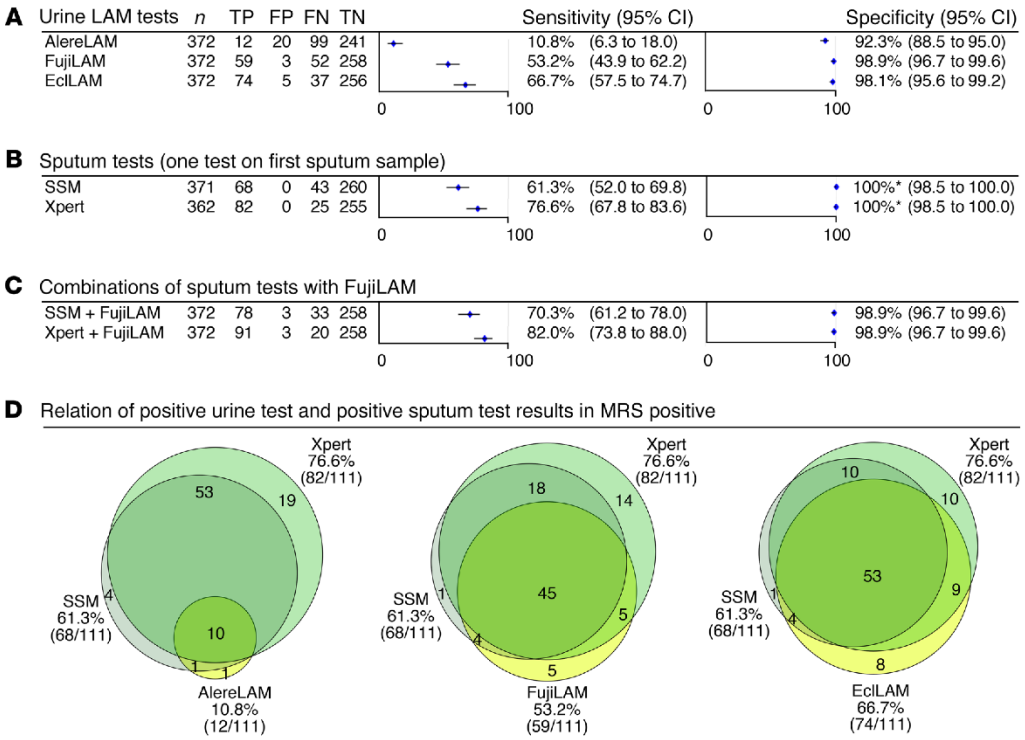

Figure 2. Diagnostic accuracy against microbiological reference standards. (**A**) Urine LAM tests, (**B**) sputum tests, (**C**) combinations of sputum tests with FujiLAM, and (**D**) positive urine LAM tests in relation to positive sputum tests. *SSM and Xpert were part of the microbiological reference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

Figure 3 shows the receiver operating characteristic (ROC) curve of the quantitative EclLAM assay and highlights the point estimates of AlereLAM, FujiLAM, and EclLAM (at a cutoff of 5.2 pg/mL) in comparison with the TPP performance target. Using the conversion scale based on the EclLAM calibration curve (Supplemental Figure 2) in Figure 3, we estimate that the LAM threshold of FujiLAM is 10–20 pg/mL and at least 10 times below the threshold of AlereLAM. Data suggest that a threshold around 5 pg/mL or below is required to meet the TPP sensitivity target.

Subgroup analysis revealed that FujiLAM sensitivity was higher in Peru (64.6%) compared with South Africa (25.0%), and this trend was also observed for AlereLAM and EclLAM (Figure 4A and Supplemental Figures 5 and 6). Subgroup analyses per MGIT TTD (Figure 4B), SSM status (Figure 4C), and semiquantitative GeneXpert MTB/RIF (Xpert) results (Figure 4D) indicate that FujiLAM sensitivity increases as mycobacterial load in sputum increases, and again this trend was confirmed with AlereLAM and EclLAM (Supplemental Figures 5 and 6).

Importantly, FujiLAM fails to detect 31.6% (24/76) of SSM-positive patients when using 3 SSM results as the basis (Figure 4C) or 27.9% (19/68) when using 1 SSM result as the basis (Figure 2D). On the other hand, FujiLAM detected 23.3% (10/43) of single SSM-negative patients with definite TB (Figure 2D). This increases the sensitivity from 61.3% for a single SSM to 70.3% when FujiLAM and a single SSM are used in combination (Figure 2, C and D). Even the more sensitive EclLAM fails to detect 16.2% (11/68) of single SSM-positive patients but would have detected 39.5% (17/43) of single SSM-negative patients with TB (Figure 2D).

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

Discussion

In this multicenter cohort study of 372 HIV-negative outpatients with respiratory symptoms suggestive of pulmonary TB from high-burden TB settings in Peru and South Africa, the FujiLAM POCT was 98.9% specific and identified 53.2% of positive TB cases, representing a 5-fold increase in sensitivity among HIV-negative patients compared with AlereLAM. FujiLAM was designed as a rule-in TB diagnostic test to allow rapid treatment initiation and reached a PPV of 95.2%. Together with its high sensitivity for TB diagnosis in people living with HIV (19) (sensitivity of 70.7% across CD4 strata), the FujiLAM might have considerable impact on the TB epidemic when scaled-up widely for use in near-patient settings. This is supported by a recent impact modeling analysis. The analysis focusing on LAM-based assays concluded that, relative to the status quo, using a urine-based LAM assay (with 70% sensitivity in people living with HIV and 30% sensitivity in HIV-negative people) in all people presenting to care with TB

Table 2. Predictive values and likelihood ratios of urine LAM tests, sputum tests, and combinations of sputum tests with FujiLAM against the microbiological reference standard

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

PPV and NPV were recalculated at an assumed TB prevalence of 20%. [A]SSM and Xpert were part of the microbiological reference standard and therefore PPV is 100% and LR+ not defined. AlereLAM, Alere Determine TB LAM Ag assay; FujiLAM, Fujifilm SILVAMP TB LAM assay; EclLAM, electrochemiluminescence-based LAM detection assay; SSM, sputum smear microscopy; Xpert, GeneXpert MTB/RIF assay; PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; LR–, negative likelihood ratio.

symptoms would avert 30% of TB deaths and 18% of incident TB cases between 2020 and 2035 in South Africa (22). The FujiLAM does meet these targets in the study analyzed here. While the study might overestimate sensitivity because of the high burden of disease at the study sites, it might at the same time underestimate sensitivity as the study did not consider patients with extrapulmonary TB or patients that have a hard time producing a sputum (e.g., children). FujiLAM's NPV is 83.2% and a negative FujiLAM result alone should not be used to rule-out TB; additional microbiological testing is required.

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

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

When comparing the 2 study sites, the performance of all LAM tests was lower in South Africa compared with Peru. The FujiLAM PPV in South Africa was 72.7% compared with 100% in Peru, which was partially due to the lower TB prevalence in South Africa. Assuming a similar pretest probability (prevalence) like in Peru, which could be achieved with optimized TB screening (e.g., with x-ray), the PPV of FujiLAM would increase to greater than 90%. This is sufficiently high to initiate treatment and substantially higher than the clinical diagnosis that is often used in today's clinical practice to initiate empiric treatment. Various indicators suggest that late medical consultation resulting in more advanced disease in Peru, or patient selection bias, are possible reasons that explain the large differences between sites. This is further supported by a relatively high TB prevalence, high SSM positivity rate, and more patients with higher mycobacterial loads in sputum, as indicated by 62% of patients with short MGIT time to detection (TTD) in Peru compared with patients in South Africa (Table 1). Subgroup analyses showed that LAM positivity is associated with surrogate markers of body mycobacterial load, such as shorter MGIT TTD, SSM positivity and Xpert semiquantitative result (Figure 4). This finding is in line with an earlier study showing that urine LAM likely reflects total mycobacterial body burden (25). On the other hand, FujiLAM was negative in a subset of patients with smear-positive disease, suggesting that mycobacterial burden in the sputum is likely not fully reflective of total mycobacterial body burden. Another factor that could impact diagnostic accuracy as a function of geography are structural differences of LAM in different TB strains but there is no scientific evidence of such differences, and further research is needed.

A high specificity (≥98%) of a POCT is necessary to avoid overtreatment. Specificity of FujiLAM in this study with HIV-negative patients was 98.9%, higher than AlereLAM's specificity at 92.3% (Figure 2). Earlier studies reported a lower specificity for FujiLAM (16, 19, 26) and the result from this study underlines the importance of a very comprehensive reference standard for a proper specificity assessment of urine biomarker tests (27).

Our study further shows the potential of LAM as a TB diagnostic biomarker. Using preconcentration of urine samples and the ultrasensitive EclLAM assay, which exploits high-affinity mono-

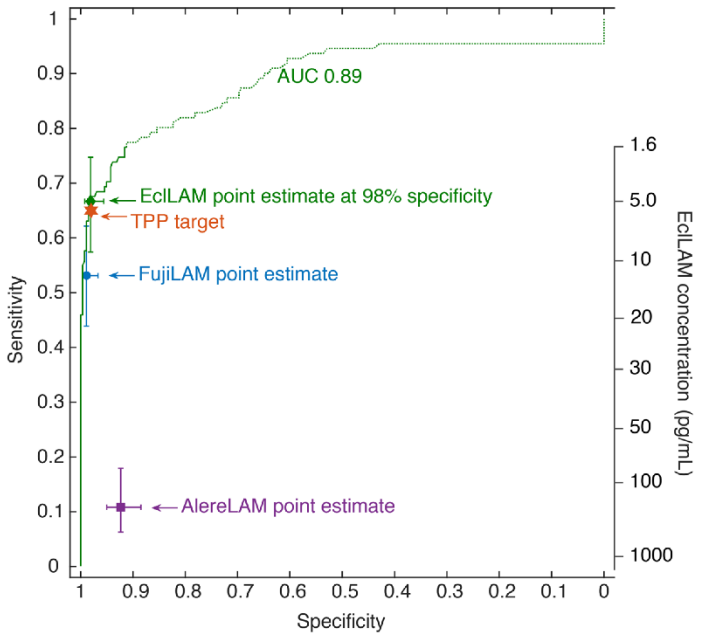

Figure 3. ROC analysis of the EclLAM concentration data compared with FujiLAM, AlereLAM, and EclLAM performance. The EclLAM concentration for the ROC curve is indicated on the secondary *y* axis. The ROC curve was restricted by the LOD (1.6 pg/mL) of the EclLAM assay, meaning that lower concentrations could not be measured so that the upper part of the ROC curve (green dotted line) should be treated with caution. n = 372 patients. TPP, target product profile.

clonal antibodies directed toward *Mtb*-specific lipoarabinomannan epitopes (14), we demonstrated that sensitivity increments compared with FujiLAM are feasible. However, this currently requires specialist laboratory equipment allowing electrochemiluminescent-based detection and sophisticated assay protocols. The EclLAM assay reached 66.7% sensitivity at 98.1% specificity, showing that a threshold around 5 pg/mL LAM or below is required to meet the TPP sensitivity target. Other recent research studies (25, 28–30) indicated that lower detection limits will translate into higher diagnostic sensitivity. We also showed this in our earlier small case control study that uses an earlier version of the EclLAM and reached 80% sensitivity in HIV-negative SSM-positive patients at a threshold of 11 pg/mL (14). In this study, despite the lower threshold due to urine preconcentration, the sensitivity of EclLAM is lower, which is likely a result of the case control design of the earlier study, whereas this study used a rigorous cohort design with a low risk of bias.

It is important to mention that our threshold is an estimate based on the EclLAM assay and nonstandardized LAM calibration material and might not be generalizable to other LAM assays with different antibodies, detection technologies, or LAM calibration material. Establishing biological reference materials, as has been done by the WHO for other diseases (31), is an urgent priority to support the development, validation, and comparison of current and future LAM assays.

EclLAM is a research assay employing laboratory equipment and is not designed for use at the POC. In addition, preconcentration was required to increase sensitivity. Therefore, key challenges in the development of next generation LAM POCT's are to reach a high analytical sensitivity with thresholds in the low pg/mL range while keeping the test simple, affordable, and highly specific. Today, the most sensitive POC lateral flow immunoassays detect antigens, like the Malaria histidine-rich protein II or LAM in case of FujiLAM, in the low picogram per milliliter range (16, 32) and sample concentration, signal amplification, and/or reagent optimization will likely be needed for POCT's to reach sensitivities like the EclLAM.

Taken together, these results suggest that LAM is present in the urine of most HIV-negative patients and that improved assay methods and reagents for LAM detection will lead to increased diagnostic accuracy. The results also suggest that as the detection limits for high sensitivity laboratory-based tests for LAM continue to improve, centralized urine or blood-based (33) TB antigen detection could also provide a high-throughput complement to nucleic acid tests for TB.

The strengths of this study were the consecutive enrollment of a cohort of HIV-negative patients from 2 epidemiologically diverse TB endemic settings in Africa and South America, the comparison of 3 independent LAM assays, the rigorous study design and the comprehensive reference standard. A limitation of the study is that patients unable to provide sputum and patients in whom the disease was thought to be only extrapulmonary and who might benefit from non–sputum-based testing (20) were excluded, which could have decreased the sensitivity of the LAM assays. Also, the SSM proportion, Xpert, and MGIT TTD results suggest more advanced disease in the Peruvian cohort but relatively low burden in South Africa, which could have artificially influenced the sensitivity of the assays. FujiLAM was designed as a POCT and can be used with fresh, unprocessed urine. The use of frozen urine samples for LAM testing in this study could have lowered LAM concentrations, as a recent study showed that the use of fresh samples leads to minor sensitivity increases in FujiLAM (34). Centrifugation of urine is not necessary before FujiLAM testing, but it was a standard procedure in this study.

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

Methods

Study design and participants. In this multicenter diagnostic accuracy study, we consecutively enrolled adults aged 18 or older with symptoms of pulmonary TB (at least 2 weeks of persistent cough and at

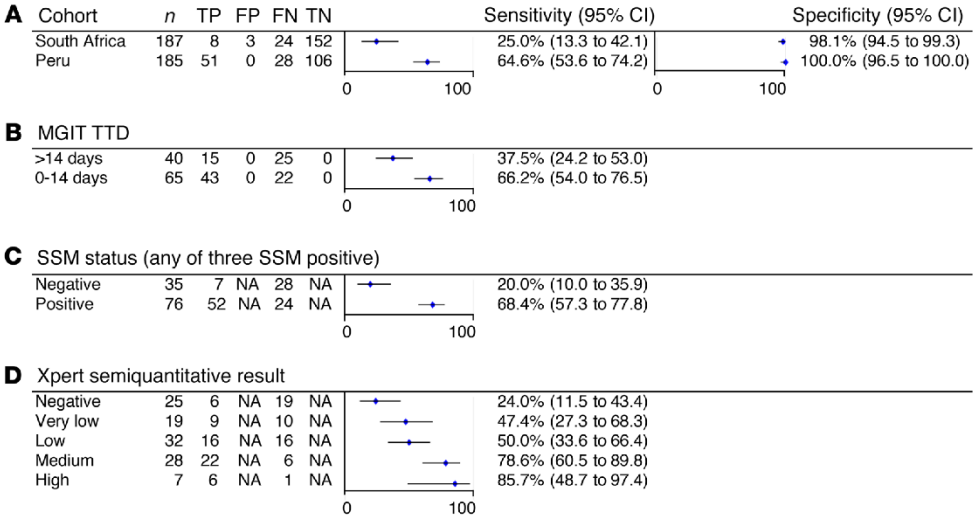

Figure 4. Subgroup analysis of FujiLAM. Sensitivity and specificity against microbiological reference standard of (**A**) FujiLAM by study site, (**B**) FujiLAM by MGIT TTD, (**C**) FujiLAM by SSM status, and (**D**) FujiLAM by semiquantitative Xpert result.

least one additional finding such as hemoptysis, weight loss, fever, night sweats, malaise, contact with an active case, chest pain, or loss of appetite) able to produce sputum. In South Africa, outpatients were enrolled at the Town Two and Nolungile primary health care facilities in the Khayelitsha township between February 9, 2017, and August 31, 2017. In Peru, outpatients were enrolled in 28 primary health care DOTS (directly observed treatment, short-course) treatment centers in high TB prevalence areas in the suburbs of Lima, and referred to the Universidad Peruana Cayetano Heredia between March 22, 2017, and October 4, 2017. Patients in whom the disease was thought to be only extrapulmonary or who received anti-TB treatment in the 60 days before enrollment were excluded (Supplemental Table 1).

Procedures. Three sputum (2 on day 1 and a third sputum on day 2), 1 blood, and 2 urine specimens were collected at enrollment and one sputum specimen was collected at 2 months follow-up for reference standard testing. Details on the specimen collection and testing flow are provided in Supplemental Figure 1. Reference standard testing was performed in the reference laboratories of the University of Cape Town and the Universidad Peruana Cayetano Heredia on all sputum specimens and included Xpert (Cepheid; Xpert testing predated roll-out of Xpert Ultra), smear fluorescence microscopy after Auramine O staining, MGIT liquid culture (Becton Dickinson), and solid culture on Löwenstein-Jensen medium. Blood cultures from all participants were done in BACTEC Myco/F Lytic culture vials (Becton Dickinson). On average 9.4 valid liquid or solid sputum cultures and 2.7 sputum Xpert results were available per patient, 74% of patients had a valid blood culture and 97% had a valid urine Xpert result. The presence of *Mtb* complex in solid and liquid culture was confirmed with MPT64 antigen (Tauns) detection and/or the MTBDRplus line probe assays (Hain Lifesciences). WHO prequalified in vitro diagnostics were used for HIV testing (rapid diagnostic tests) and CD4 cell counting (flow cytometry). Urine was immediately put on ice after collection and processed within 4 hours. Urine was centrifuged (2000*g* at 4°C for 10 minutes), aliquoted on the day of collection, and stored at −80°C until batch testing of the liquid fraction with LAM assays. For urinary Xpert testing, 30–40 mL urine was centrifuged (3000*g* at 4°C for 15 minutes), and following removal of supernatant the pellet was resuspended in 1 mL PBS and tested using Xpert on the day of collection. Clinical information, index test results, and comparator test results were not available to the assessors of the reference standard.

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

For AlereLAM and FujiLAM testing at RIT-JATA, frozen urine aliquots were thawed to ambient temperature and mixed by inversion. Samples that were not immediately used for testing were stored at 4°C for a maximum of 4 hours. Testing with FujiLAM was performed according to the manufacturers' instructions using urine from the same aliquot as that used for AlereLAM. The 5-step FujiLAM test procedure is illustrated in an online video (35) and takes 50–60 minutes from start to end result. The FujiLAM assay does not use a reference scale card and any visible test line is considered positive. The AlereLAM test was used according to the tests package insert and the four-grade Reference Scale Card, with the Grade 1 cutoff point as the positivity threshold. FujiLAM and AlereLAM were independently read by 2 readers, each blinded to all other clinical, demographic, and test data associated with the samples. After the initial test interpretation, the 2 readers compared results and, in the event of discordance, established a final consensus result by mutual agreement. In case of test failure, the test was repeated, and the first valid result was used for the analysis.

Blinded EclLAM testing at Meso Scale Diagnostics followed a previously established assay protocol, except for the addition of a preconcentration step to improve the limit of detection (14). To preconcentrate, frozen urine aliquots were thawed to ambient temperature, mixed by inversion, and a 490 μL sample was added to Amicon Ultra-0.5 mL centrifugal filters (MilliporeSigma) with a 3 kDa cutoff.

After ultrafiltration for 20 minutes at 14000*g*, approximately 25 μL deionized water was added to the retentate (~45 μL) to get a total of 70 μL and an estimated concentration factor of 7. Prior to analysis, samples were heat treated at 85°C for 10 minutes. The immunoassays for LAM used the same antibodies as FujiLAM in a sandwich immunoassay format employing ECL detection as described in the Supplemental Methods and elsewhere (14, 17, 36). The research team performing EclLAM had no access to all other clinical, demographic, and test data associated with the samples.

Statistics. Before data analysis, clinical investigators, who were masked to index test results, categorized patients as having definite TB, possible TB, not TB, and unclassifiable, using a combination of clinical and laboratory findings (Supplemental Table 2). Patients with definite TB had microbiologically confirmed *Mtb* (any culture or any Xpert positive for *Mtb* during admission). Patients defined as not TB had all microscopy, culture, and Xpert test results negative for *Mtb*, had not started TB treatment, and were alive or had improvement in clinical tuberculosis symptoms at a 2-month follow-up. Patients defined as possible TB did not satisfy the criteria for definite TB but had clinical features suggestive of TB and were started on TB treatment. Patients who did not fall into any of these categories were defined as unclassifiable and were removed from the main analyses. Definite TB and not TB categories were used to allocate patients into positive and negative, respectively, for both the MRS and CRS. The possible TB group was considered negative by MRS, but positive by CRS as previously proposed in a study guidance publication (27).

Simple descriptive statistics were used to characterize cohorts. We calculated the point estimates and 95% Wilson CIs for the sensitivity, specificity, PPV, NPV, LR+, and LR– for FujiLAM, AlereLAM, EclLAM, Sputum Smear Microscopy (SSM), Xpert, and combinations of FujiLAM+SSM and FujiLAM+Xpert by comparison with the MRS and CRS. The R package eulerr (37) was used to generate area-proportional Euler diagrams to illustrate the number of positive test results by test in the definite TB group. Fagan nomograms were used to illustrate pretest and posttest probabilities of FujiLAM and AlereLAM.

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

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

Author contributions

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

Acknowledgments

The authors thank Mark D. Perkins and Ranald Sutherland for helping with the conceptualization of this work; Michelle A. Bulterys, Amanda Jackson, Celeste Worship, Meagan McMaster, Aurélien Macé, Stefano Ongarello, and André Trollip for data management; and the clinical and laboratory teams, including Jimena Collantes, Cesar Ugarte, Nchimunya Hapeemla, and Widaad Zemanay, at the partner sites for their efforts in the implementation, conduct, and timely completion of the study. This work was funded by the Global Health Innovative Technology Fund (grant G2015-201), the UK Department for International Development (grant 300341-102), the Dutch Ministry of Foreign Affairs (grant PDP15CH14), the Bill and Melinda Gates Foundation (grant OPP1151258), the Australian Department of Foreign Affairs and Trade (grant 70957), the German Federal Ministry of Education and Research through Kreditanstalt für Wiederaufbau (grant 2020 60 457), and the National Institute of Allergy and Infectious Diseases, NIH (grant 1R01AI104589).

Address correspondence to: FIND, Emmanuel Moreau, Chemin des Mines 9, 1202 Geneva, Switzerland. Phone: 41.22.710.0590; Email: emmanuel.moreau@finddx.org. Or to: Tobias Broger, Metlenweg 4, 9050 Appenzell, Switzerland. Phone: 41.78.843.2112; Email: tobias.broger@gmail.com.

1. WHO. Global tuberculosis report 2019. https://www.who.int/tb/publications/global_report/en. Updated October 17, 2019. Accessed August 17, 2020.
2. WHO. High-priority target product profiles for new tuberculosis diagnostics. Report of a consensus meeting. https://www.who.int/tb/publications/tpp_report/en. Accessed August 17, 2020.
3. Denkinger CM, et al. Defining the needs for next generation assays for tuberculosis. *J Infect Dis*. 2015;211 Suppl 2:S29–S38.
4. MacLean E, et al. A systematic review of biomarkers to detect active tuberculosis. *Nat Microbiol*. 2019;4(5):748–758.
5. Bulterys MA, et al. Point-of-care urine LAM tests for tuberculosis diagnosis: a status update. *J Clin Med*. 2019;9(1):E111.
6. Bjerrum S, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in people living with HIV. *Cochrane Database Syst Rev*. 2019;10:CD011420.
7. WHO. Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update (2019). https://www.who.int/tb/publications/2019/LAMPolicyUpdate2019/en. Accessed August 17, 2020.
8. WHO. WHO consolidated guidelines on tuberculosis module 3: diagnosis - rapid diagnostics for tuberculosis detection. https://www.who.int/publications/i/item/who-consolidated-guidelines-on-tuberculosis-module-3-diagnosis---rapid-diagnostics-for-tuberculosis-detection. Updated June 30, 2020. Accessed August 17, 2020.

9. Peter JG, et al. Effect on mortality of point-of-care, urine-based lipoarabinomannan testing to guide tuberculosis treatment initiation in HIV-positive hospital inpatients: a pragmatic, parallel-group, multicountry, open-label, randomised controlled trial. *Lancet*. 2016;387(10024):1187–1197.
10. Gupta-Wright A, et al. Rapid urine-based screening for tuberculosis in HIV-positive patients admitted to hospital in Africa (STAMP): a pragmatic, multicentre, parallel-group, double-blind, randomised controlled trial. *Lancet*. 2018;392(10144):292–301.
11. Sahle SN, et al. Performance of point-of-care urine test in diagnosing tuberculosis suspects with and without HIV infection in selected peripheral health settings of Addis Ababa, Ethiopia. *BMC Res Notes*. 2017;10(1):74.
12. Nicol MP, et al. Urine lipoarabinomannan testing for diagnosis of pulmonary tuberculosis in children: a prospective study. *Lancet Glob Health*. 2014;2(5):e278–e284.
13. Kroidl I, et al. Performance of urine lipoarabinomannan assays for paediatric tuberculosis in Tanzania. *Eur Respir J*. 2015;46(3):761–770.
14. Sigal GB, et al. A novel sensitive immunoassay targeting the 5-methylthio-d-xylofuranose-lipoarabinomannan epitope meets the WHO's performance target for tuberculosis diagnosis. *J Clin Microbiol*. 2018;56(12):e01338-18.
15. Suwanpimolkul G, et al. Utility of urine lipoarabinomannan (LAM) in diagnosing tuberculosis and predicting mortality with and without HIV: prospective TB cohort from the Thailand Big City TB Research Network. *Int J Infect Dis*. 2017;59:96–102.
16. Broger T, et al. Novel lipoarabinomannan point-of-care tuberculosis test for people with HIV: a diagnostic accuracy study. *Lancet Infect Dis*. 2019;19(8):852–861.
17. Choudhary A, et al. Characterization of the antigenic heterogeneity of lipoarabinomannan, the major surface glycolipid of *Mycobacterium tuberculosis*, and complexity of antibody specificities toward this antigen. *J Immunol*. 2018;200(9):3053–3066.
18. Joe M, et al. The 5-deoxy-5-methylthio-xylofuranose residue in mycobacterial lipoarabinomannan. Absolute stereochemistry, linkage position, conformation, and immunomodulatory activity. *J Am Chem Soc*. 2006;128(15):5059–5072.
19. Broger T, et al. Diagnostic accuracy of a novel tuberculosis point-of-care urine lipoarabinomannan assay for people living with HIV: A meta-analysis of individual in- and outpatient data. *PLoS Med*. 2020;17(5):e1003113.
20. Kerkhoff AD, et al. Diagnostic sensitivity of SILVAMP TB-LAM (FujiLAM) point-of-care urine assay for extra-pulmonary tuberculosis in people living with HIV. *Eur Respir J*. 2020;55(2):1901259.
21. Sossen B, et al. 'SILVAMP TB LAM' rapid urine tuberculosis test predicts mortality in hospitalized HIV patients in South Africa [published online January 9, 2020]. *Clin Infect Dis*. https://doi.org/10.1093/cid/ciaa024.
22. Ricks S, Denkinger CM, Schumacher SG, Hallett TB, Arinaminpathy N. The potential impact of urine-LAM diagnostics for tuberculosis in South Africa: a modelling analysis. *PLoS Med*. In press.
23. Huerga H, et al. Incremental yield of including determine-TB LAM assay in diagnostic algorithms for hospitalized and ambulatory HIV-positive patients in Kenya. *PLoS ONE*. 2017;12(1):e0170976.
24. Sun AY, Pai M, Salje H, Satyanarayana S, Deo S, Dowdy DW. Modeling the impact of alternative strategies for rapid molecular diagnosis of tuberculosis in Southeast Asia. *Am J Epidemiol*. 2013;178(12):1740–1749.
25. Paris L, et al. Urine lipoarabinomannan glycan in HIV-negative patients with pulmonary tuberculosis correlates with disease severity. *Sci Transl Med*. 2017;9(420):eaal2807.
26. Bjerrum S, et al. Diagnostic accuracy of a novel and rapid lipoarabinomannan test for diagnosing tuberculosis among people with human immunodeficiency virus. *Open Forum Infect Dis*. 2019;7(1):ofz530.
27. Drain PK, et al. Guidance for studies evaluating the accuracy of biomarker-based nonsputum tests to diagnose tuberculosis. *J Infect Dis*. 2019;220(Suppl 3):S108–S115.
28. Shapiro AE, et al. A point-of-care test to concentrate and detect urine LAM for TB diagnosis: results from the first-in-human study of FLOW-TB. Abstract presented at: 50th World Conference on Lung Health of the International Union Against Tuberculosis and Lung Disease; October 30–November 2, 2019; Hyderabad, India. https://hyderabad.worldlunghealth.org/wp-content/uploads/2019/11/20191101_UNION2019_Abstracts_Final.pdf.
29. Wood A, et al. Ultrasensitive detection of lipoarabinomannan with plasmonic grating biosensors in clinical samples of HIV negative patients with tuberculosis. *PLoS ONE*. 2019;14(3):e0214161.
30. Hamasur B, Ignatowicz L, Ramachandraiah H. SpeClean: urine sample treatment method for ultra-sensitive LAM diagnostics. Abstract presented at: 50th World Conference on Lung Health of the International Union Against Tuberculosis and Lung Disease; October 30–November 2, 2019; Hyderabad, India. https://hyderabad.worldlunghealth.org/wp-content/uploads/2019/11/20191101_UNION2019_Abstracts_Final.pdf.
31. WHO. Providing International Biological Reference Preparations. https://www.who.int/biologicals/reference_preparations/en. Accessed March 31, 2020.
32. Das S, et al. Performance of a high-sensitivity rapid diagnostic test for *Plasmodium falciparum* malaria in asymptomatic individuals from Uganda and Myanmar and naive human challenge infections. *Am J Trop Med Hyg*. 2017;97(5):1540–1550.
33. Broger T, et al. Sensitive electrochemiluminescence (ECL) immunoassays for detecting lipoarabinomannan (LAM) and ESAT-6 in urine and serum from tuberculosis patients. *PLoS ONE*. 2019;14(4):e0215443.
34. Broger T, Muyoyeta M, Kerkhoff AD, Denkinger CM, Moreau E. Tuberculosis test results using fresh versus biobanked urine samples with FujiLAM. *Lancet Infect Dis*. 2020;20(1):22–23.
35. Fujifilm SILVAMP TB LAM test procedure. Fujifilm SILVAMP TB LAM test procedure. https://youtu.be/aK-QtzkLBug. Updated April 24, 2019. Accessed August 17, 2020.
36. Kawasaki M, et al. Lipoarabinomannan in sputum to detect bacterial load and treatment response in patients with pulmonary tuberculosis: Analytic validation and evaluation in two cohorts. *PLoS Med*. 2019;16(4):e1002780.
37. Larsson J. eulerr: Area-proportional Euler and Venn diagrams with ellipses. R Package version 6.1.0. https://cran.r-project.org/package=eulerr. Updated March 9, 2020. Accessed August 17, 2020.

3.2 Publication 2 - Diagnostic yield of urine lipoarabinomannan and sputum tuberculosis tests in people living with HIV: a systematic review and meta-analysis of individual participant data

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).

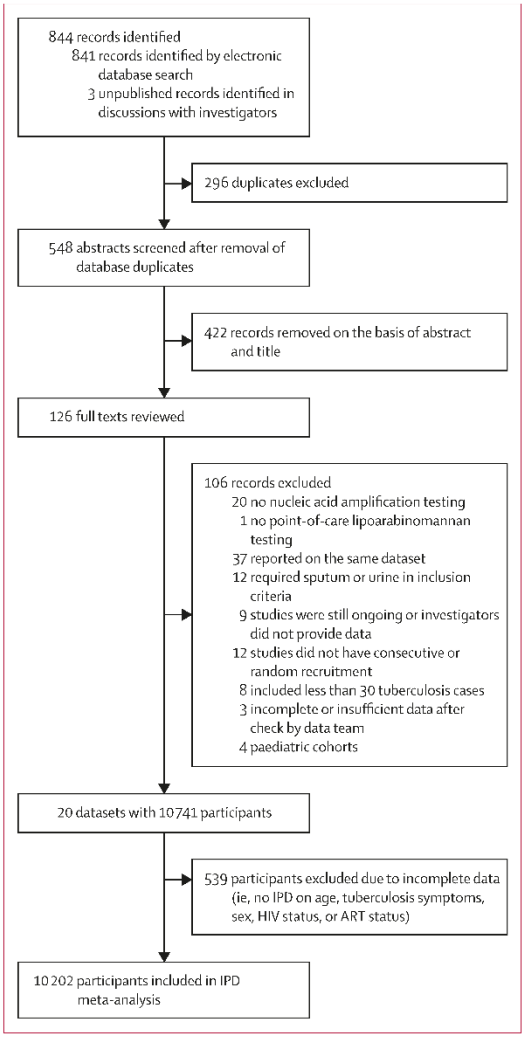

Figure 1: **PRISMA diagram of studies included in this meta-analysis and reasons for exclusion**
ART=antiretroviral therapy. IPD=individual participant data.

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%)

(Table 1 continues in next column)

	All participants (n=10 202)
(Continued from previous column)	
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 10741 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 10741 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.

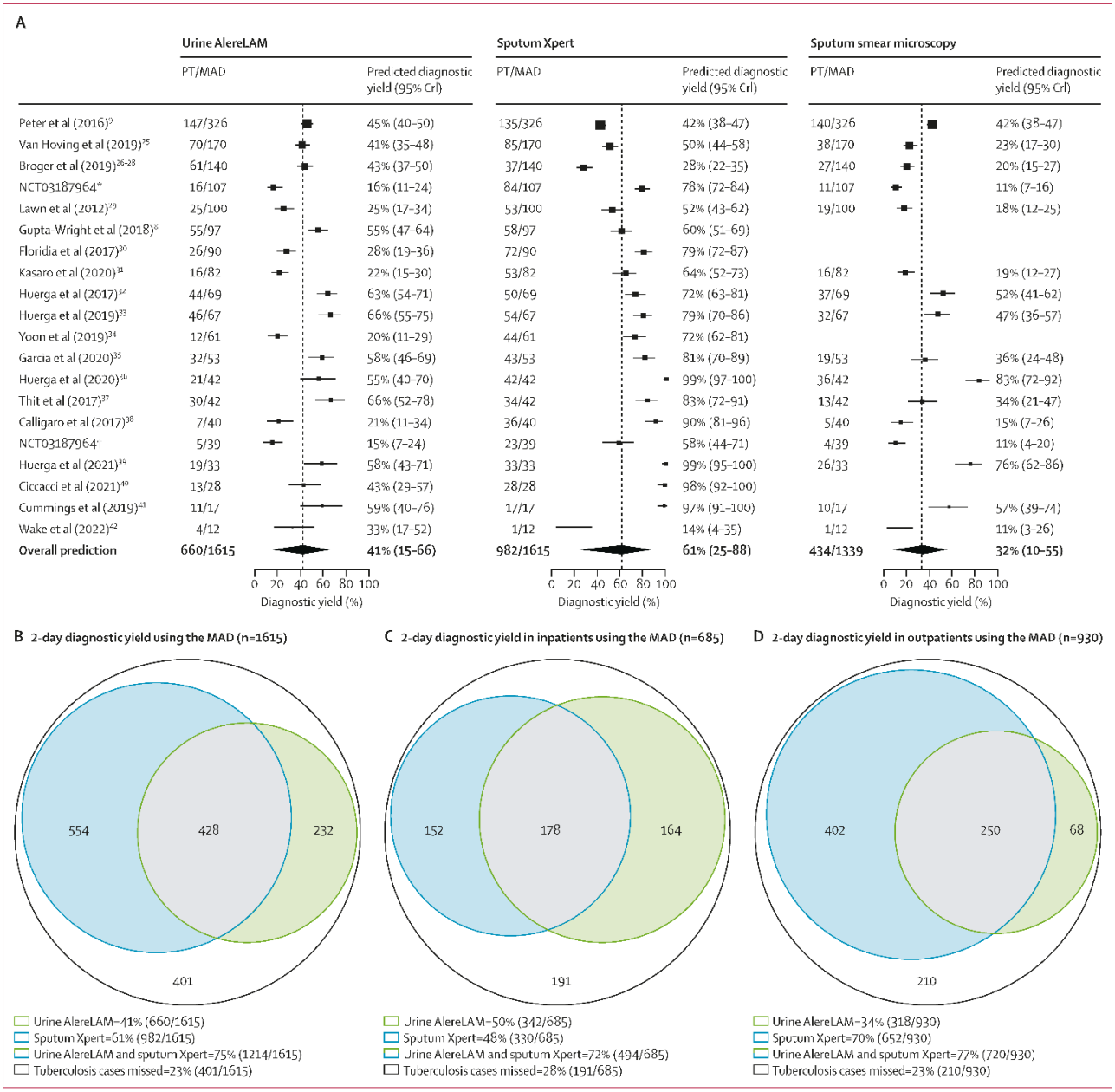

Figure 2: **Diagnostic yields of urine AlereLAM, sputum Xpert, and sputum smear microscopy**

(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.

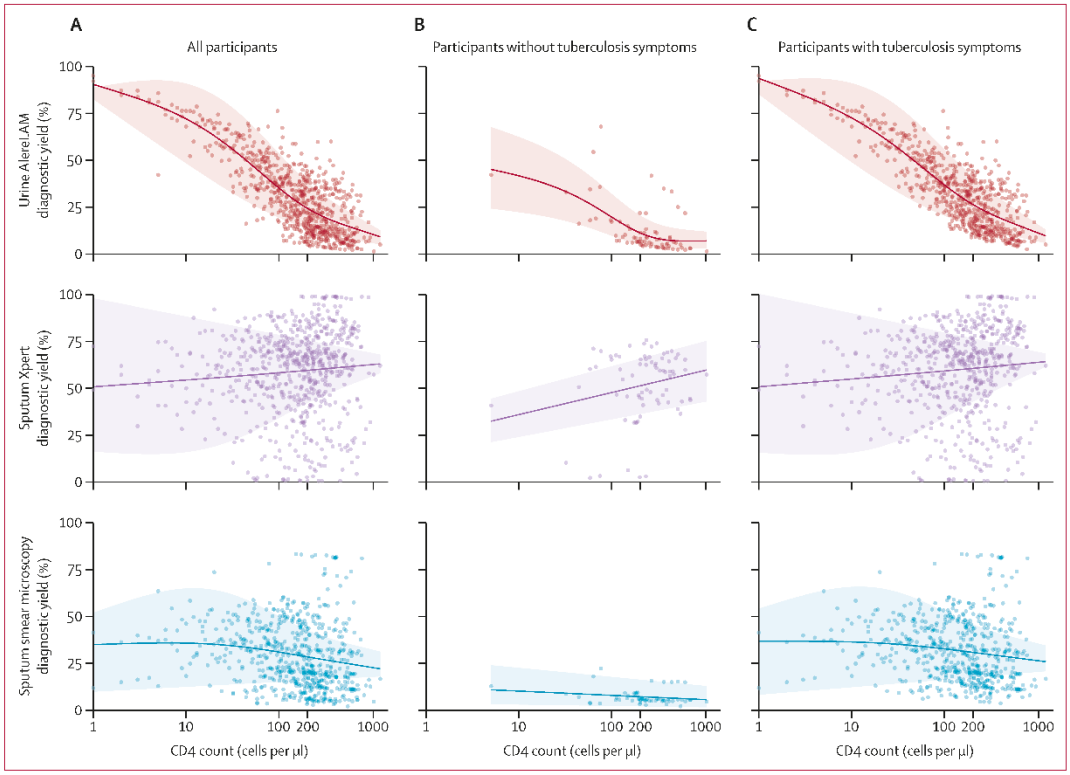

Figure 3: **Diagnostic yield predictions as a function of CD4 cell count**
Diagnostic yield predictions in participants with tuberculosis based on the MAD in all participants (A), participants without tuberculosis symptoms (B), and participants with tuberculosis symptoms (C). We used the number of positive 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,[19] 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[18,50,51] 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.[19] 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.[37] 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. CMe 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.

3.3 Publication 3 - Diagnostic yield as an important metric for the evaluation of novel tuberculosis tests – rationale and guidance for future research

Comments: The publication has been accepted by Lancet Global Health but is currently being put into the format of the journal. Number of Figures and Table is like in the accepted manuscript. It will be published under the Creative Commons Attribution (CC BY 4.0) license and can be copied and redistributed in any medium or format for any purpose. The paper uses the we form indicating the group of authors. My lead on the paper and contribution is clearly defined in the Prologue in Chapter 1.

Authors and Affiliation:

Name	Highest Degree	Email	Affiliation
Tobias Broger	MSc	tobias.broger@gmail.com	Department of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, Heidelberg, Germany
Florian M Marx	MD	florian.marx@uni-heidelberg.de	Department of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, Heidelberg, Germany DSI-NRF Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Faculty of Science, Stellenbosch University, Stellenbosch, South Africa
Grant Theron	PhD	gtheron@sun.ac.za	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
Ben J Marais	MD	ben.marais@sydney.edu.au	The University of Sydney Infectious Diseases Institute (Sydney ID) and the Children's Hospital at Westmead, Sydney, Australia
Mark P Nicol	MD	mark.nicol@uwa.edu.au	Division of Infection and Immunity, School of Biomedical Sciences, University of Western Australia, Perth, Australia
Andrew D Kerkhoff	MD	andrew.kerkhoff@ucsf.edu	Division of HIV, Infectious Diseases and Global Medicine, Zuckerberg San Francisco General Hospital and Trauma Center, and Center for Tuberculosis University of California San Francisco, San Francisco, CA, USA
Ruvandhi Nathavitharana	MBBS	rnathavi@bidmc.harvard.edu	Division of Infectious Diseases, Beth Israel Deaconess Medical Center, Boston, MA 02215
Helena Huerga	PhD	helena.huerga@epicentre.msf.org	Department of Epidemiology. Epicentre, Paris, France
Ankur Gupta-Wright	MD	ankurgw@outlook.com	Division of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, Heidelberg, Germany Department of Infectious Diseases, Imperial College London, London, UK

Mikashmi Kohli	PhD	mikashmi.Kohli@finddx.org	FIND, the global alliance for diagnostics, Geneva, Switzerland
Brooke E Nichols	PhD	brooke.nichols@finddx.org	FIND, the global alliance for diagnostics, Geneva, Switzerland Department of Medical Microbiology, Amsterdam University Medical Center, Amsterdam, the Netherlands
Monde Muyoyeta	PhD	monde.muyoyeta@cidrz.org	Centre for Infectious Diseases Research in Zambia, Lusaka, Zambia
Graeme Meintjes	PhD	graeme.meintjes@uct.ac.za	Department of Medicine, University of Cape Town, Cape Town, South Africa / Wellcome Centre for Infectious Diseases Research in Africa (CIDRI-Africa), Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa
Morten Ruhwald	MD	morten.ruhwald@finddx.org	FIND, the global alliance for diagnostics, Geneva, Switzerland
Rosanna W Peeling	PhD	rosanna.peeling@lshtm.ac.uk	London School of Hygiene &Tropical Medicine, Keppel Street, London WC1E 7HT, UK
Nitika Pant Pai	MD	nitika.pai@mcgill.ca	Department of Medicine, , Centre for Outcomes Research & Evaluation (RI-MUHC), McGill University, Montreal, Canada
Nira R Pollock	MD	nira.pollock@childrens.harvard.edu	Boston Children's Hospital, Boston, MA, USA
Madhukar Pai	MD	madhukar.pai@mcgill.ca	McGill International TB Centre, McGill University, Montreal, Canada
Adithya Cattamanchi	MD	adithya.cattamanchi@uci.edu	Department of Medicine, Division of Pulmonary Diseases and Critical Care Medicine, University of California Irvine, Irvine, USA Center for Tuberculosis, University of California San Francisco, San Francisco, USA
David W Dowdy	MD	ddowdy1@jhmi.edu	Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA
Puneet Dewan	MD	puneet.dewan@gatesfoundation.org	Bill & Melinda Gates Foundation, Seattle, WA, USA
Claudia M Denkinger*	MD	claudia.denkinger@uni-heidelberg.de	Department of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, Heidelberg, Germany AND German Center for Infection Research (DZIF), partner site Heidelberg University Hospital, Heidelberg, Germany

*Corresponding author: Claudia M. Denkinger, Department of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, Im Neuenheimer Feld 324, 69120, Heidelberg, Germany, email: claudia.denkinger@uni-heidelberg.de, phone: +49 6221 56-22999

Keywords: Tuberculosis; TB; Diagnostics; Yield; Accuracy; Effective Population Coverage; Universal Access

Summary

Better access to TB testing is a key priority for fighting tuberculosis (TB), the leading infectious disease killer of humankind. Despite the roll-out of molecular World Health Organization (WHO)-recommended rapid diagnostics (mWRDs) to replace sputum smear microscopy over the past decade, a large diagnostic gap remains. Of the estimated 10.6 million people who developed TB globally in 2022, over 3.1 million were not diagnosed.

Better TB tests are required, yet an exclusive focus on improving accuracy alone will not be sufficient to close the diagnostic gap for TB. Diagnostic yield, which we define as the proportion of people in whom a diagnostic test identifies TB among all people we attempt to test for TB, is an important metric not adequately explored. Diagnostic yield is particularly relevant for subpopulations unable to produce sputum such as young children, people living with HIV, and people with subclinical TB. As more accessible non-sputum specimens (e.g. urine, oral swabs, saliva, capillary blood, breath) are being explored for point-of-care TB testing, the concept of yield will be of growing importance.

Using the example of urine LAM testing, we illustrate how even tests with limited sensitivity may diagnose more people with TB if they enable increased diagnostic yield. Using tongue swab-based molecular TB testing as another example, we provide definitions and guidance for the design and conduct of pragmatic studies that assess diagnostic yield. Lastly, we show how diagnostic yield and other important test characteristics like cost and implementation feasibility are essential for increased effective population coverage, which are required for optimal clinical care and transmission impact.

We are calling for diagnostic yield to be incorporated into TB test evaluation processes, including the WHO Grading of Recommendations, Assessment, Development and Evaluations (GRADE) process, providing a crucial ‘real life’ implementation metric that complements traditional accuracy measures.

Introduction

TB diagnosis has relied on low sensitivity sputum smear microscopy for over 100 years. In 2022, of the estimated 10.6 million people who developed TB, 3.1 million were not diagnosed and reported (World Health Organization 2023a). The persistent TB diagnostic gap is closely associated with the inability of countries to reach the WHO standard of universal access to rapid molecular TB diagnostics (World Health Organization 2023b). In 2022, for example, only 47% of patients notified with TB were initially tested with a World Health Organization (WHO)-recommended rapid diagnostics (WRD) (Ismail et al. 2023; World Health Organization 2023a). Constraints in diagnostic access are central to these institutional failures, spanning the cascade

from being identified as needing a test, obtaining specimens, receiving testing to starting and completing treatment (Pai et al. 2022).

Figure 1 illustrates key moments in the last century that have been instrumental in defining how TB diagnostics are evaluated.

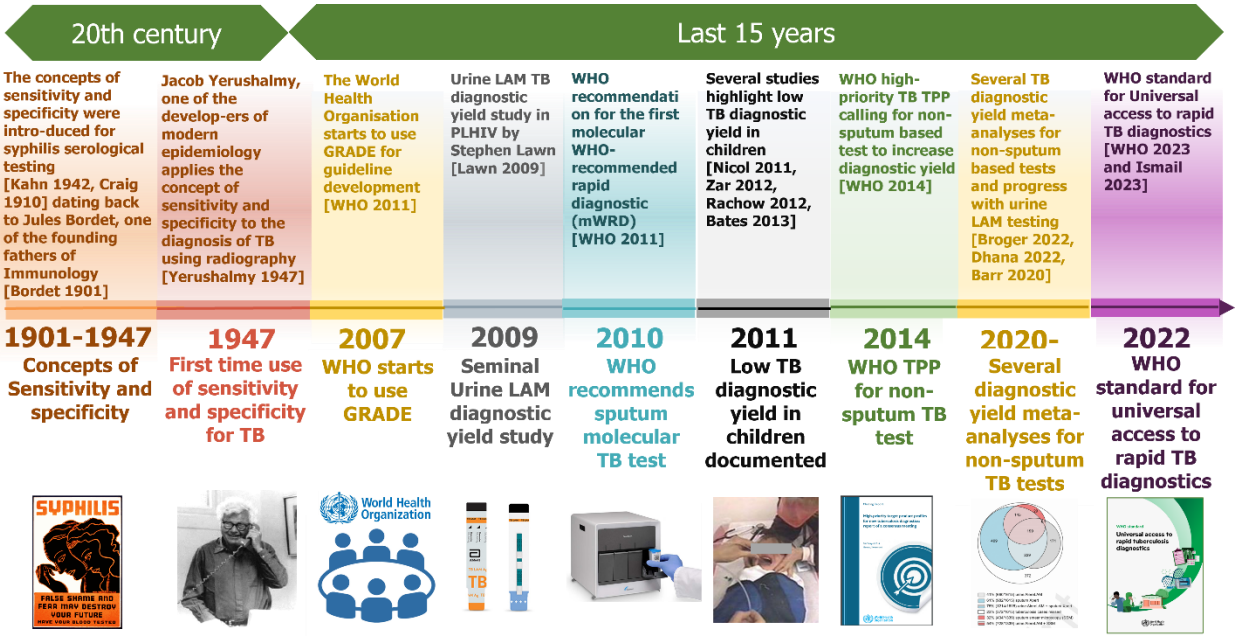

Figure 1: Key moments and seminal publications in the conceptualization of diagnostic accuracy, policy development and diagnostic yield that inform the current approach to evaluate TB diagnostics*. TB=Tuberculosis. GRADE=Grading of Recommendations, Assessment, Development and Evaluations. TPP=Target product profile. WHO=World Health Organization. LAM=Lipoarabinomannan. mWRD=molecular WHO-recommended rapid diagnostic. References (Barr et al. 2020; Bates et al. 2013; Bordet and Gengou 1910; Broger et al. 2023; Craig 1910; Dhana et al. 2022; Hsu et al. 2011; Ismail et al. 2023; Kahn 1942; Lawn et al. 2009; Nicol and Zar 2011; Rachow et al. 2012; World Health Organization 2013; World Health Organization 2014; Yerushalmy 1947; Zar et al. 2012)*

Since 2007, the World Health Organization (WHO) has been applying the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) (Schünemann et al. 2008) process for guideline development. GRADE prioritizes randomized controlled trials (RCTs) that directly evaluate the impact of a diagnostic test on patient-important outcomes under 'real-life' conditions (Schünemann et al. 2019). However, in the absence of direct evidence from RCTs, WHO's guideline development groups (GDGs) usually link accuracy studies to patient-important outcomes, such as cure, mortality, time to diagnosis, and time to treatment, and integrate these into GRADE's evidence to decision (EtD) framework to infer likely impact of tests and develop recommendation (Schünemann et al. 2016). If a test is not likely to improve patient-important

outcomes or population health, then the healthcare system has no reason to use it, whatever its accuracy may be (Schünemann et al. 2008).

A new pipeline of TB diagnostics is emerging, partly as a dividend from massive diagnostic infrastructure investments made during the COVID-19 pandemic (Yerlikaya et al. 2023). While traditional accuracy metrics and individual patient-important outcomes remain critical, population health-focused measures are equally important in defining how diagnostics can be deployed to achieve public health objectives. Such measures are currently less emphasized within GRADE, and accordingly recommendations for new diagnostics may fail to focus on the potential to identify more people with TB and close the diagnostic gap (Pai et al. 2023). In this article, we aim to emphasize the importance of diagnostic yield for emerging TB diagnostics and how yield relates to effective population coverage.

What is diagnostic yield?

Since the mid-20th century, diagnostic yield has become a metric in evaluating the utility of various diagnostic tests and procedures across different medical specialties. Many of the initial publications focused on cancer screening, but it is increasingly used in infectious diseases as well (Barr et al. 2020; Bates et al. 2013; Broger et al. 2023; Corral et al. 2019; Dhana et al. 2022; Krumholz 1966; Lawn et al. 2009; Mase et al. 2007; Meggi et al. 2017; Nicol and Zar 2011; Rachow et al. 2012; Zar et al. 2012).

For this article, diagnostic yield (DYT, Panel 1) is defined as the proportion of people identified with disease using a highly specific diagnostic test (thus resulting in mostly true positives – see further comment below; PT), out of all people eligible to be tested (D), irrespective of adequate specimen collection. Diagnostic yield is a comprehensive measure of the performance of a test because it considers access, specimen availability, sensitivity, and test completion. Considering the diagnostic literature more broadly, the common denominator is the total number of people who are attempted to test, which also includes people who were unable to provide a specimen and those for whom the test failed to deliver a result (due to indeterminates, invalids or errors which are linked to test robustness and user friendliness). Another definition (DYD) frequently used in TB literature (Barr et al. 2020; Boyles et al. 2018; Broger et al. 2019a; Broger et al. 2023; Dhana et al. 2022; Huerga et al. 2017; Lawn et al. 2012b; Lawn et al. 2014) only considers the diagnostic yield among those diagnosed with TB. The article will subsequently employ the DYT definition; nonetheless, we advocate for the assessment of both DYT and DYD in studies.

Usually, estimates of diagnostic yield are based on a single test attempt using a single diagnostic specimen from a single clinical encounter, as would happen in routine clinical care. Diagnostic yield may further include turnaround time such as “diagnostic yield at the first clinical encounter” or “24-hour diagnostic yield”. The concept of yield can also be extended to diagnostic algorithms involving more than one test as composite diagnostic yield.

Panel 1: Definition of diagnostic yield

Diagnostic yield among all tested (DYT)	**Definition:** **Proportion of people in whom a diagnostic test identifies TB among all people for whom testing is attempted.** **Formula:** **DYT=PT/D** DYT=Diagnostic yield among all people attempted to test. PT=Number of people with a positive diagnosis by the test. D=Denominator defined as the total number of people for whom TB testing is attempted. DYT may include a turnaround time component such as "diagnostic yield during the first clinical encounter" or "24-hour diagnostic yield". Usually, DYT is based on a single test attempt using a single specimen from a single clinical encounter, but it can consider a test series or standardised combination of tests. The crucial factor is that it should mimic real-world clinical practice in a TB endemic setting.	**Strengths:** -**Simplicity in study design**: The focus on positive results (PT) simplifies study design, making it feasible even in the absence of a comprehensive reference standard that would be required to distinguish between true and false positives. -**Pragmatic Assessment:** Enables studies that replicate real-world clinical practices by considering factors such as test completion, specimen viability, and timely result availability in all people attempted to test. **-Reflects real-world testing conditions**: By considering a single test attempt using a single specimen from a single clinical encounter, DYT mirrors real-world testing conditions. **Limitations:** -**Dependence on prevalence**: DYT is influenced by TB prevalence, which may limit its generalizability across populations with different prevalences. **-Specificity consideration:** Specificity requires careful consideration in light of the clinical goal, associated cost and prevalence. DYT includes false positives in PT. If test specificity has been well established (e.g. in previous accuracy studies) and is high (≥98.5%), the impact of false positives on DYT is low, particularly if TB prevalence is high. In this case, PT approximates the number of true positive results. DYT for tests with lower specificity should be adjusted (see Web Appendix p. 1) for formula and Excel Calculator) and be interpreted carefully.
Diagnostic yield among all diagnosed **(DYD)**	**Definition:** **Proportion of people in whom a diagnostic test identifies TB among TB positive people for whom testing is attempted.** **Formula:** **DYD=PT/DD** DYD=Diagnostic yield among TB positive people. PT=Number of people with a positive diagnosis by the test. DD=Denominator defined as the total number of people diagnosed with TB. Usually, a comprehensive microbiological reference standard that includes mycobacterial culture and/or NAAT from any specimen including sputum, urine, blood, and other extrapulmonary samples. The positive participants from the index test under assessment can be included in the denominator permitted specificity of the test is sufficiently high. As for DYT, DYD may include a turnaround time component and is usually based on a single test, but can include a test series or standardised combination of tests used in real world clinical practice	**Strengths:** -**Not depending on prevalence:** DYD is independent of TB prevalence, making it easier to compare across studies conducted in different settings or populations. **-Addressing false positives:** The comprehensive reference standard enables the exclusion of false positives in the count of people with a positive diagnosis (PT). -**Has been used in several TB studies** **Limitations:** -**Requires comprehensive reference standard:** DYD requires a comprehensive microbiological reference standard, with multiple tests for defining the total number of people diagnosed with TB in the denominator (DD). This may not be feasible in all settings, particularly in pragmatic studies. Even the most comprehensive reference standard may still miss people with TB.

Why is diagnostic yield important and how is it linked to effective population coverage?

Figure 2 maps TB testing and care metrics to the TB care cascade. The ultimate goal of a diagnostic test is to achieve population-level impact on patient outcomes which requires effective population coverage (Jannati et al. 2018; Murray and Evans 2003; Shengelia et al. 2005) and universal access (Ismail et al. 2023; World Health Organization 2023a). Diagnostic yield is an important part of effective population coverage as it measures a test's ability to deliver actionable positive diagnoses in those attempted to test (covering steps 3 and 4 of the TB care cascade in Figure 2).

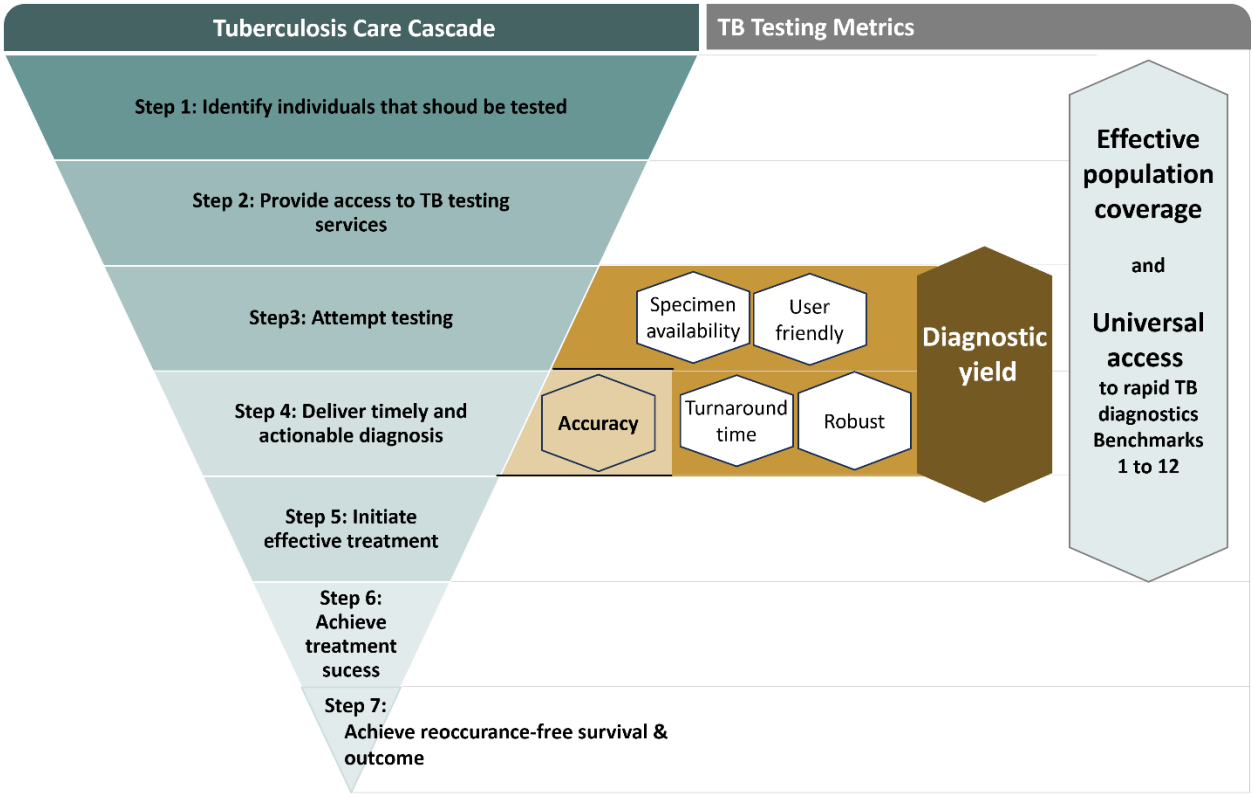

***Figure 2: TB testing metrics (right) mapped to the TB care cascade (left)**. Accuracy evaluates step 4, diagnostic yield steps 3 and 4 and effective population coverage steps 1 to 5 of the care cascade. TB testing and care metrics refer to the REASSURED criteria(Land et al. 2018; Mabey et al. 2004), which include Ease of specimen collection, Accuracy (Sensitive, Specific), User-friendly, Rapid & Robust, and are considered elements of diagnostic yield. REASSURED further includes: Affordable, Deliverable Equipment-free, and Real-time connectivity (not shown in figure) which are relevant for effective population coverage. Tuberculosis (TB) care cascade adapted from Subbaraman et al. (2019) and Ismail et al. (2023). Universal access benchmarks are described in the WHO standard (World Health Organization 2023b).*

In addition to test accuracy, diagnostic yield covers other key aspects by accounting for specimen availability, turnaround time, test robustness and failures, and user friendliness. All of these aspects collectively contribute to effective and timely result generation in a real-world clinical

setting in high endemic countries but are not usually covered by diagnostic accuracy studies. Specimen availability is particularly relevant for key subpopulations that are often unable to produce sputum, such as people living with HIV (PLHIV), children, people with extrapulmonary TB and people with subclinical TB who do not exhibit overt symptoms and signs associated with active TB (Broger et al. 2023; Huerga et al. 2023).

Sputum induction may be necessary for patients who cannot expectorate, but this puts healthcare workers at risk for infection and requires expertise, motivated staff, equipment, and time, potentially delaying specimen collection and time to diagnosis. Further, sputum is a complex viscous sample requiring complex sample processing which leads to longer turnaround times and high requirements on test robustness to avoid indeterminates and invalids. Despite these challenges, sputum remains the primary TB diagnostic specimen. Recognizing the limitations of sputum-based diagnostics triggered the explicit inclusion of non-sputum specimens as a priority element of TPPs for new TB diagnostics (Denkinger et al. 2015; World Health Organization 2014; World Health Organization 2024).

In principle, TB tests with only moderate sensitivity which utilize more easily accessible specimen have the potential to diagnose a higher number of people with TB than a more sensitive molecular test reliant on sputum as illustrated in the following two examples.

In the first example, in an individual participant data (IPD) meta-analysis (Broger et al. 2023) of TB testing among 3662 hospitalized PLHIV, 69% (2531/3662) had a sputum specimen obtainable in the first 2 days, whereas 98% (3585/3662) had a urine specimen obtained. Diagnostic yield was comparable, at 9.3% (342/3662) for urine LAM and 9.0% (330/3662) for sputum Xpert (Figure 3). This result was obtained despite the lower sensitivity of the LAM test used in the studies (42% for urine AlereLAM) relative to the sputum assay (77% for sputum Xpert). The comparability in diagnostic yield was purely attributable to higher urine specimen availability. While everyone can provide a urine sample, only a certain fraction of people can produce sputum samples.

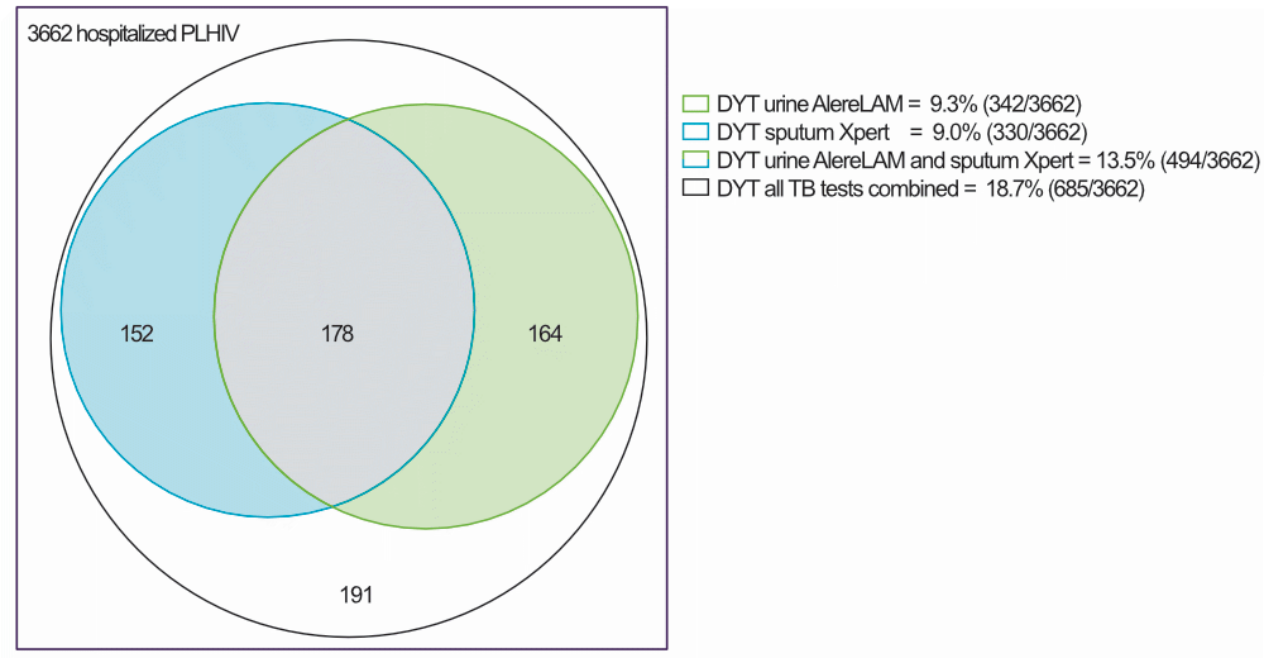

Figure 3: Comparison of TB diagnostic yield between urine LAM and sputum Xpert *among hospitalized PLHIV from the first sample collected in the initial 2 days after enrolment (Broger et al. 2023). Diagnostic yield calculated as number of patients with positive test results among all people attempted to test. DYT=Diagnostic Yield. AlereLAM=Alere Determine TB LAM Ag assay. Xpert=Xpert MTB/RIF or Xpert Ultra assay.*

The second is a hypothetical example. Consider replacing Xpert Ultra on a single spot sputum (assuming specimen availability 80%, sensitivity 91%, specificity 98.5%) with tongue swab specimen and a nucleic acid amplification test (NAAT) (assuming specimen availability 100%, sensitivity 73%, specificity 98.5%). When applied to the same population, the same number of TB patients would be identified by the two testing approaches, since the specimen collection advantage compensates for the loss in sensitivity when assessing tongue swab specimens.

An important consideration for the concept of diagnostic yield is test specificity. If test specificity has been well established in diagnostic accuracy studies and confirmed to be high (≥98.5%), then the number of positive results in a high incidence setting would be representative of true positive cases. This simplifies study design when highly specific tests are evaluated allowing more pragmatic studies, even in the absence of a comprehensive reference standard to distinguish between true and false positives. For tests with lower specificity, this is more complex, and adjustment could be performed using Bayesian Latent Class Analysis (or other methods with similar properties) to i) allow explicit incorporation of additional information about an individual and their probability of having TB, e.g. including chest X-Ray result, whether a clinical diagnosis was made, or if follow-up is possible - response to therapy; and ii) include appropriate

considerations of uncertainty about specificity estimates (Schumacher et al. 2016). For reporting, the diagnostic yield can be adjusted by considering test specificity from other studies or meta-analyses (see Web Appendix p. 1 for formula and Excel Calculator).

Using more readily accessible specimens can enable higher diagnostic yield for TB. Paired with near-patient testing, including home- and self-testing, this lays the basis for higher effective population coverage, including for disadvantaged and vulnerable groups. A diagnostic test with a high accuracy will have limited impact on the population if it does not also have high diagnostic yield and is not widely available and used. It is therefore important to not only emphasize diagnostic yield, but also coverage (Figure 2). Effective population coverage refers to the population at risk that is reached by a specific health intervention, such as a diagnostic test, and who are therefore able to benefit from it. As such it combines need, utilization, and quality (Jannati et al. 2018; Ng et al. 2014). For a test to reach high effective population coverage, it needs to be feasible to implement at scale, affordable, and cost-effective, and it should achieve high diagnostic yield within a clinically relevant turnaround time (Figure 2).

How to integrate diagnostic yield into diagnostic research

Panel 2 shows an example study approach to establish diagnostic accuracy and diagnostic yield for a tongue swab-based molecular TB test and Table 1 provides a checklist for the design of a diagnostic yield study. In general the same criteria used for high-quality accuracy studies apply (Bossuyt et al. 2015; Drain et al. 2019; Whiting 2011). A two-step process might be most appropriate with initial establishment of diagnostic accuracy (especially specificity) using a comprehensive and well validated reference standard from sputum (including methods that facilitate sputum production; detailed guidance on such studies is published elsewhere (Denkinger et al. 2019), followed by a 'real life' diagnostic yield study. In some instances, a pragmatic combined effectiveness-implementation study might be useful, if accuracy and yield assessment can be incorporated into a single study.

Panel 2: Example of the proposed two-step study approach and PICO (Population, Intervention, Comparison, Outcome) questions to establish diagnostic accuracy and yield for a tongue swab-based molecular TB test.

Accuracy study PICO

- Population: Patients at risk of TB (with/without symptoms) presenting to a health facility
- Intervention: Tongue swab-based molecular test
- Comparator: Sputum-based molecular test
- Primary Outcome: Diagnostic accuracy (sensitivity and specificity) in reference to a microbiological reference standard (including sputum culture and sputum molecular test),
- Secondary Outcomes: Proportion indeterminate, diagnostic accuracy against alternative reference standards (clinical reference standard, extended microbiological reference standard, or latent class modelling (Drain et al. 2019; Schumacher et al. 2016)), negative predictive value (NPV), positive predictive value (PPV), turn-around-time, quality and proportion indeterminate of the reference standard methods.

Diagnostic yield study PICO

- Population: Patients at risk of TB (with/without symptoms), ideally in healthcare or community settings
- Intervention: Tongue swab-based molecular test
- Comparator: Sputum-based molecular test
- Outcomes: Diagnostic yield (DYT) defined as the proportion of tongue swab positive people (numerator TP, Panel 1) among the total number of people attempted to test (denominator D, Panel 1).
- Secondary Outcome: composite yield of both tests or test algorithms, time to diagnosis, sample provision, proportion indeterminate, effectiveness of implementation, user-friendliness in programmatic setting with the intended user.

For the diagnostic accuracy study, demonstrating specificity against a comprehensive reference standard will be critical to ensure that additional cases identified in the diagnostic yield study are true positive cases. This may be of less concern in the example of a swab-based molecular test that specifically detects *Mtb* DNA (the caveat being presence of *Mtb* DNA after cure) (Bahr et al. 2018; Dorman et al. 2018). However, if there is less confidence in the specificity of a test, as is the case for LAM-detecting urine-based TB assays, this consideration gains critical importance and could be addressed by analytical means (e.g. Bayesian latent class analysis) as outlined above (Schumacher et al. 2016).

Diagnostic yield studies should be designed with the necessary statistical power and similar statistical methods used for accuracy sample size calculations (Drain et al. 2019; Newcombe 1998). For example, for a test with 20% diagnostic yield, 49 TB test positives would be required to achieve a 95% confidence interval (CI) width of less than 10%. Assuming a prevalence of 25% this would require the enrolment of 246 participants, ideally without selection bias, as is often observed in studies that only enrol highly symptomatic participants who can spontaneously expectorate sputum.

Table 1: Checklist for the design of a TB diagnostic yield study based on Drain et al. (2019).

Topic	**Recommendation**
Study design	• Use an inclusive and representative cross-sectional study design (minimal exclusions)
	• Include assessment of diagnostic yield as a primary or secondary objective
	• Ensure sample size calculation accounts for yield objectives, specimen provision and indeterminate results
Participants and Population	• Don't exclude participants based on inability to provide a specimen and describe the characteristics of this participant group well (e.g. did they later produce sputum or was sputum induction attempted, were they followed-up and what was the clinical picture, etc.)
	• Consider diagnostic yield assessment in settings and populations with minimally symptomatic individuals, risk factors for TB and people with difficulties to produce sputum such as PLHIV, children, community-based and facility-based TB screening for active case finding
	• Avoid selecting participants in whom TB has already been diagnosed for the current episode or who have already started TB treatment
	• Perform testing in intended use setting with intended user (pragmatic setting and staff is preferable)
Specimen collection	• Document number of specimen collection attempts, timing, and collection time per specimen
	• Document whether sputum was spontaneously expectorated or induced and describe the volume and quality of the sputum sample obtained. Document the level of coaching, effort and attempts to collect sputum
	• If sputum induction is done; provide clear instructions and training to healthcare workers, the effort to collect a specimen should be representative of the routine situation and well documented
	• Collect participant and healthcare worker feedback on which specimen they prefer

Index test and comparators	• Document time to result from first attempt to collect specimen
	• Link test result to specimen collection data
	• Document numbers and reasons for non-availability of results (no specimen, assay failure, etc.)
	• Consider specifics of the index test that may influence study design, ideal population, and setting
	• Include sputum smear microscopy and sputum Xpert as comparators. Consider including other WHO recommended tests as comparators (e.g., LAM, particularly if urine-based tests are evaluated).
Reporting	• Report on baseline diagnostic yield (1) among total number of people attempted to test (DYT) and (2) among all diagnosed with TB (DYD).
	• Report specimen availability, indeterminate results, invalids, and collection efforts
	• Consider subgroup analysis (e.g., based on specimen provision and result time, incl. time when a patient was started on treatment)
	• Consider reporting on composite yield of test regimens/algorithms (e.g., swab-based molecular and urine LAM in combination, or X-ray based triage and mWRD-based confirmation)
	• Publish participant level data along with a clear codebook to enable secondary research (meta-analyses, modelling, policy development)
	• Discuss the potential impact of false positives and disease prevalence on diagnostic yield. Report specificity-adjusted diagnostic yield (see Web Appendix p. 1 and Excel Calculator). Consider evaluating how the presence of diseases with similar symptoms in a population, context specific environmental factors, including healthcare infrastructure, accessibility to healthcare services, and socioeconomic conditions, might influence diagnostic yield.
	• Consider reporting on and discussing key aspects beyond diagnostic accuracy and yield relevant for easy test access and effective population coverage such as cost, time-to-diagnosis, user friendliness, training requirements, training, throughput, portability, etc. and their implications for practice and effective population coverage.

What lessons on diagnostic yield and effective population coverage can TB take from other diseases of public health importance?

For several diseases, diagnostic innovation has improved both yield and coverage. Table 2 lists diagnostic examples from different diseases to illustrate the interplay of test characteristics to achieve effective population coverage, including yield.

For syphilis diagnosis, for example, one study indicated that improvement in the sensitivity of antenatal syphilis tests without a corresponding increase in patient return rate would not yield any substantial gains in health outcomes, highlighting the importance of turnaround time (Aledort et al. 2006). Today, serological syphilis rapid diagnostic tests (RDTs), which are less accurate but more accessible than laboratory diagnostics, are part of the WHO's list of essential diagnostics (World Health Organization 2021a).

HIV self-testing also provides critical lessons for the TB community. HIV rapid test uptake is negatively impacted by the stigma and discrimination associated with visibility of testing in health

facilities. In 2012, when the US Food and Drug Administration (FDA) approved oral self-tests for HIV, the concerns about lower accuracy of oral self-tests compared to laboratory-based tests were overridden because of their potential to expand diagnostic yield due to ease of sampling (e.g. oral vs. finger prick blood), potential for expanded access to testing, and thus increase effective population coverage compared to laboratory tests. The WHO's release of self-testing guidelines in 2016 catalysed the global availability, accessibility, and impact of these tests (Giguère et al. 2021). Improved population coverage was achieved in Southern Africa by a successful rollout of oral self-tests for screening followed by blood-based tests for confirmation, resulting in a reduction of the proportion of undiagnosed individuals without knowledge of their HIV serostatus from 40-50% in 2000 to 16% in 2020 (Giguère et al. 2021).

For malaria, the development of antigen-based RDTs changed the landscape by offering accurate diagnosis while circumventing both venous blood collection and microscopy obstacles in peripheral health care settings, including cost of equipment, time to result, and the need for skilled personnel. The first malaria RDTs emerged in the early 1990s (Thepsamarn et al. 1997), and the WHO held its first meeting on rapid diagnostic testing in 1999 (World Health Organization 2000). Initial adoption was hampered by variable field performance, which led the WHO and other agencies to create an international quality control programme for malaria RDTs (Cunningham et al. 2019). In recent years, RDT testing has been significantly expanded around the world. In 2021, 413 million RDTs were sold by manufacturers and 262 million were distributed by national malaria programmes (World Health Organization 2021b).

Most recently, the COVID-19 pandemic also showed the benefits of shifting attention from the narrow focus on test accuracy to diagnostic yield and effective population coverage to address diagnostic gaps (Mina et al. 2020; Pai et al. 2023). Nasal rapid antigen tests (RATs) achieved high diagnostic yield and effective population coverage despite their reduced sensitivity relative to NAATs from nasopharyngeal swabs (Pavelka et al. 2021; World Health Organization 2022; Zhang et al. 2022).

Novel testing solutions for TB testing on the horizon that have the potential to improve diagnostic yield

Developing non-sputum-based rapid tests for TB presents significant challenges due to the anatomical location of TB infection that usually involves the lung parenchyma with resultant lung pathobiology. Few biomarkers progress towards tests with clinical utility (MacLean et al. 2019). For pathogen markers (e.g. antigen or DNA), abundance in non-respiratory specimens (like blood, or urine) is very low, complicating sensitive detection with low-cost point-of-care tests (Broger et al. 2020b). For host markers, detection has also proved challenging due to similarities in the immune response related to *Mtb* infection (without disease) and active disease (Broger et al. 2017). Nevertheless there are several exciting non-sputum tests based on urine, tongue swab,

breath, blood and stool on the TB testing horizon (Portevin et al. 2014; R2D2 2024; Treatment Action Group (TAG) 2023; Yerlikaya et al. 2023). It is important to highlight that two non-sputum-based tests are already recommended by the WHO and available today. These include the urine AlereLAM test to assist in TB diagnosis in PLHIV and Xpert stool testing in children (World Health Organization 2020). The greater appreciation of yield within the TB field will be critical to increase acceptance and uptake of these tests and will help to address major implementation gaps that continue to exist for these tests.

Table 2: Exemplar use-cases of diagnostic tests that achieved high effective population coverage

Example	Use-case	Sample type	Test characteristics	Accuracy	Diagnostic Yield	Population coverage	References
HIV	Community-based POC and self-testing	Oral mucosal transudate (OMT)	Rapid antibody testing	Reduced accuracy (sensitivity 98.7%, specificity 99.8%) against laboratory-based blood tests. Confirmatory testing is recommended.	High diagnostic yield due to accessible sample type and high acceptability.	High: In 2020, 84% of people living with HIV knew their HIV status. Expansive coverage was achieved by a successful roll-out funded by the UNITAID STAR Program and BMGF. Between 60-90% or participants opted for selft-testing depending on setting. Self-testing increased update by 145%.	(Hatzold et al. 2019) (Giguère et al. 2021) (Zachary et al. 2012) (Witzel et al. 2020) (Pai et al. 2012)
Syphilis	Screening and POC testing	Finger-stick	Rapid antibody testing	Reduced accuracy (sensitivity 75-100%, specificity 65-100%) for rapid tests in minimal to no infrastructure areas compared to laboratory tests.	High diagnostic yield: fingerpick capillary blood is feasible in community settings and health facilities during one encounter without laboratories. Minimal infrastructure tests could alleviate 47% of disease burden despite lower	High: Countries have begun using dual HIV/syphilis RDTs to increase effective population coverage for both, HIV and syphilis. Syphilis RDTs were added to WHO's list of essential diagnostics. Tests which require minimal infrastructure and return results in <2 hrs. can alleviate about 30%-50% of disease burden	(Aledort et al. 2006)

					sensitivity and specificity		
Malaria	POC testing	Finger-stick	Rapid antigen testing	Reduced accuracy (62% sensitivity, 99% specificity compared to laboratory-based molecular testing.	High diagnostic yield: fingerpick capillary blood is feasible in community settings and health facilities during one encounter without laboratories	High In 2021, 413 million RDTs were sold by manufacturers.	(Cunningham et al. 2019) (Yimam et al. 2022)
SARS-CoV-2	Screening and self-testing to identify infectious people which effectively limits further spread	Nasal swab	Rapid antigen testing	Reduced accuracy (76% sensitivity, 98.9% specificity) against reference standard PCR. Similar sensitivity for self-testing compared to professional testing.	High: accessible sample types (i.e. self-performed anterior nasal swab) Over 80% of users found rapid antigen tests easy to perform	High: Improved substantially initially through large publicly funded screening programs. Initially with assisted testing and further through lay self-testing	(Young et al. 2021) (Polechová et al. 2022) (Lindner et al. 2021a) (Lindner et al. 2021b) (Pavelka et al. 2021) (Mina et al. 2020) (Brümmer et al. 2022)

Conclusion

As part of urgent efforts to identify and treat people with TB who are currently missed and not receiving appropriate care, it is necessary to place a stronger focus on diagnostic yield and effective population coverage. Using more readily accessible specimens together with novel near-patient tests, including home- and self-testing, will improve diagnostic yield and coverage, especially in disadvantaged and vulnerable groups who are at the greatest risk of disease development and spread. Thus, we propose the inclusion of diagnostic yield as an additional metric when evaluating the value of novel diagnostic tests for TB, as well as in the GRADE evidence synthesis process that informs WHO policy decisions.

Contributors

TB and CMD wrote the first draft and all other authors contributed text, revised, commented and approved the final version of the manuscript. Authorship, including the order of authors was based on contribution and ICMJE criteria.

Declaration of interests

TB reports patent applications in the field of TB detection and is a shareholder of Avelo Ltd. MP serves as an advisor for non-profits such as BMGF, FIND, WHO and Stop TB Partnership. All other authors have no financial or industry-related conflicts.

Acknowledgements

We thank Miriam Compton for the manuscript review. This article was written without external funding. Consequently, there was no involvement of any funding source.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT4 for proofreading and shortening sections of text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Appendix

Specificity Adjusted Diagnostic Yield

Imperfect test specificity leads to an overestimation of diagnostic yield. Diagnostic yield can be adjusted for test specificity as follows:

$$Specificity\ Adjusted\ DYT = \frac{PT - ((1 - Prevalence) * (1 - Specificity) * Test\ Completion * D)}{D} \times 100\%$$

DYT: Diagnostic yield among all tested (as per Panel 1).

PT: Number of index test positive tests (as per Panel 1).

Prevalence: Estimate based on experience in study population. Alternatively, Bayesian latent class analysis with combination of available test results can be used.

Specificity: Test specificity of the test under assessment based on published evidence from accuracy studies or meta-analyses.

Test completion: Proportion of people in the diagnostic yield study with a valid positive or negative test results (N) over all people (D).

D: Total number of people for whom TB testing is attempted (as per Panel 1)

Separately, also evaluate against DD=Denominator defined as the total number of people diagnosed with TB to consider DYD.

Excel Diagnostic Yield Calculator

An Excel Tool for the calculation of Diagnostic Yield and Specificity Adjusted Diagnostic Yield, including an example based on literature values is attached.

References

Are integrated in the dissertation references below.

4 DISCUSSION

“Having a TB diagnostic is our top priority for diagnostics – there is many diagnostics that we want, but the one that would be the most dramatic would be to have something like a two dollars or less TB diagnostic” (Bill Gates from the Bill & Melinda Gates Foundation during a meeting with Indian Minister of Health, Mansukh Mandaviya in New Delhi in February 2024, NDTV 2024). The development and delivery of such a test would not just be a scientific advancement; it would be a critical step towards fulfilling the Sustainable Development Goal (SDG) 2030 target of reducing TB deaths by 90% and cutting TB incidence by 80%. The potential impact of this next-generation LAM TB test could be transformative, offering a more effective tool in the global fight against tuberculosis.

In alignment with Gates' emphasis, my dissertation delved into investigating the potential of LAM, one of the most promising biomarkers for fulfilling this crucial need. Consequently, the overarching question of my dissertation was: “Is LAM a suitable biomarker for TB diagnosis in general populations?”.

In the first section of this integrated discussion, I summarize the status of all 15 areas of research and development from Figure 3. This also includes an integrated discussion of the three focus areas that formed the core of this cumulative dissertation and led to the three publications presented in Chapter 3.

In the second section, I provide a nuanced answer to the overarching question whether LAM is a suitable biomarker for TB diagnosis in general populations by predicting TPP fulfilment of the currently feasible, hypothetical, next generation LAM test.

In the conclusion, I end by summarizing my main findings and provide some final thoughts and recommendations.

4.1 A status update of 15 LAM research and development areas

After a decade of dedicated research following my LAM strategy, I and others made progress on all 15 research and development areas (Figure 3), paving the way for a next-generation LAM-based TB diagnostic test. Table 2 summarizes the status and the following sections discusses all areas in detail.

Table 2: Is LAM a suitable biomarker for TB diagnosis – answers to the 15 areas of research and development from Figure 3.

No.	Question	Short answer	Status	Outstanding Questions and Future Considerations	References
Stream A - LAM Biomarker in clinical specimens					
1	What is the molecular structure of *in vivo* LAM and what are the most promising epitopes to target with existing and new antibodies to maximize performance?	A combination of two antibodies performs best for *in vivo* uLAM detection: (a) the S4-20 antibody targets the MTX-dependent Man2 or Man3 epitope in LAM, and (b) the A194-01 antibody possesses high affinity for linear Ara4, branched Ara6 and also binds Man-capped structures. See Figure 2 for LAM epitopes and structure.		• How is LAM degraded in the host and how does that impact LAM structure and ability to be detected in bodily fluids? • Are there alternative binding motifs to push sensitivity to the next level? • Can even better antibodies with higher affinity and specificity be developed? • The question: what is the molecular structure of *in vivo* LAM? is not yet fully answered.	Cantera et al. 2022; Cantera et al. 2023; Choudhary et al. 2018; De et al. 2021; Lowary and Achkar 2022; Sigal et al. 2018b; Zheng et al. 2017; Corrigan et al. 2022; Ishida et al. 2021
2	Is urine or blood the better specimen for LAM detection?	Urine		• The understanding of LAM concentration and structure in specimen other than urine is still limited and further research is required. • Comparisons across specimen types are influenced by reagents, structural differences of LAM in different sample types, absence of LAM reference material, sample treatment, assay designs, and detection platforms.	Brock et al. 2020b; Broger et al. 2019d
3	Is there a difference between LAM detection in fresh vs. biobanked urines?	There are only minor differences between LAM detection in fresh vs. biobanked urines. Urines from biobanks can be used for the development. Once the design is final, performance should be confirmed in fresh urines.		• Assessment of analyte stability and matrix effects should be part of every assay validation and it is essential that this is done with the assay under consideration, as different LAM assays use different reagents that bind to different LAM epitopes with different stabilities or availabilities. • Stability of LAM in urine has not been confirmed and studied with quantitative LAM assays and long-term freezing may diminish uLAM levels. • Differences could be assay/epitope specific, and reassessment of new LAM tests is needed, particularly if they use new antibodies.	Broger et al. 2020a; Connelly et al. 2019

No.	Question	Short answer	Status	Outstanding Questions and Future Considerations	References
4	Is there a difference between LAM detection in early-morning vs. spot urines?	There are no major differences between LAM detection in early-morning vs. spot urines.		• Only information from few studies. • Differences could be assay specific, and reassessment of new LAM tests will be needed.	Bjerrum et al. 2022; Székely et al. 2022
Stream B - Sensitive Assays					
5	Can a highly sensitive reference assay for LAM be developed?	Yes, I was able to develop a highly sensitive laboratory-based electrochemiluminescence LAM research assay (EclLAM). It achieved a limit of detection (LOD) of 1.6 pg/ml.		• Development of LAM reference material and well-titrated calibration samples will be required to compare the analytical sensitivities of reference assays moving forward. • LODs are influenced by detection methodologies, detection reagents, sample processing, etc. • LAM quantification is complicated by the differences of *in vitro* and *in vivo* uLAM and the unclear molecular structure of *in vivo* uLAM.	Broger et al. 2020c; Cantera et al. 2022; Hoa et al. 2023; Sigal et al. 2018b
6	Can low-cost, simple, and rapid platform alternatives be developed that substantially improve POC detection compared to conventional lateral-flow immunoassays (LFIAs)?	Yes, development of FujiLAM showed that improved alternatives to conventional gold nanoparticle-based LFIAs (such as AlereLAM) can be developed. This paved the way for the development of several other next-generation uLAM tests.		• The increased sensitivity requirement may come at a cost of longer time to result, higher cost, or higher complexity.	Broger et al. 2019a; Treatment Action Group (TAG) 2023
Stream C - Patient: population, use-case, feasibility for TB diagnosis in all people					
7*	Is LAM detectable in people with TB without HIV co-infection and what are concentration ranges and accuracy?	66.7 % of people with TB without HIV co-infection had uLAM concentrations ≥5 pg/ml when using the EclLAM assay (at a specificity of 98.1%). The predicted relationship between test performance and LAM concentration can be found in Broger et al. (2020, Figure 3) and in Table 3. .		• Does the predominant lineage of the *Mtb* strain play a role in uLAM levels and test sensitivity • Current concentration ranges might not be generalizable to other LAM assays with different antibodies, detection technologies, or LAM calibration/reference material. Can more generalizable concentrations be measured? • Measuring LAM concentration with analytical methods like mass spectrometry has been attempted but led to very different results.	Broger et al. 2020; Hoa et al. 2023; Amin et al. 2018

No.	Question	Short answer	Status	Outstanding Questions and Future Considerations	References
8*	What are the diagnostic yields of first and next-generation LAM tests alone and in combination with sputum smear microscopy or sputum Xpert?	Yields in PLHIV are: Sputum Smear Microscopy (SSM): 32% Urine AlereLAM: 41% Sputum Xpert: 61% Urine FujiLAM: 65% Sputum Xpert + Urine AlereLAM: 75% SSM+ Urine AlereLAM: 54%		• Study parallel use of TB diagnostic tests, including on different specimen to maximize diagnostic yield. • Report on sample provision and diagnostic yield in parallel to test accuracy.	Broger et al. 2023; Bjerrum et al. 2024
9	LAM in children	The 2019 WHO LAM recommendation to use AlereLAM to assist in the diagnosis of active TB in PLHIV also applies to adolescents and children living with HIV. For FujiLAM testing in children, Olbrich et al. (2022), conducted a systematic review and found three studies with sensitivities and specificities ranging from 42% to 63% and 84 to 93%, respectively, against a microbiological reference standard.		• Prospective studies in children. • Direct head-to-head comparison of AlereLAM to next generation LAM tests.	Bjerrum et al. 2019a; Comella-del-Barrio et al. 2021; Kroidl et al. 2015; LaCourse et al. 2018; Nicol et al. 2014; Nicol et al. 2020; Nkereuwem et al. 2020; Olbrich et al. 2022
10	LAM in extrapulmonary TB (EPTB)	uLAM tests can diagnose TB in EPTB patients. More sensitive uLAM detection with FujiLAM showed promising and substantially higher sensitivity over the AlereLAM for detecting extra-pulmonary TB in HIV positive inpatients.		• Future research is needed to better understand the mechanisms by which LAM enters the bloodstream and urine. • Research on uLAM detection in HIV-negative individuals with EPTB is needed. • The currently recommended Xpert may require an invasive sample to be collected, which limits its use for EPTB detection to hospitals. uLAM detection could be an important alternative given its rapid detection on easily obtainable urine and its increased performance in patients with disseminated TB.	Kerkhoff et al. 2020

No.	Question	Short answer	Status	Outstanding Questions and Future Considerations	References
11	LAM as a mortality predictor	uLAM positivity is a predictor of mortality. A next-generation uLAM test could rapidly diagnose TB in up to 89% of PLHIV who die while a negative uLAM result means a high probability of three-month survival (86%).		• The clinical impact, including the potential survival benefit of next-generation uLAM test (once their accuracy has been shown), should be prospectively assessed in different real-world settings and more diverse populations—including outpatients, children, and HIV-uninfected patients.	Sossen et al. 2020
Stream D - Impact & Implementation, public health, impact, access, policy					
12	What is the impact of a next-generation LAM POC test?	AlereLAM use has been shown to reduce all-cause mortality in high-risk groups of PLHIV. Hospitalized PLHIV who received AlereLAM testing had 15% lower risk of mortality at eight weeks compared to routine TB diagnostic testing without AlereLAM. Modelling suggests that the use of an improved, next-generation LAM test (sensitivity 30% in HIV negatives and 70% in HIV positives) amongst all people with symptoms, regardless of their HIV status, could avert 30% of TB deaths and 18% of TB cases within 15 years in South Africa.		• Evaluating the effect of LAM testing on patient-important outcomes in children. • LAM testing will not be able to detect drug resistance; how does this limit impact? • Study other patient-important outcomes including time to diagnosis and time to initiation of TB treatment (related to diagnostic yield and effective coverage).	Grant et al. 2020; Gupta-Wright et al. 2018b; Nathavitharana et al. 2021; Peter et al. 2016b
13	Is the use of a next-generation LAM test cost-effective?	Yes, the use of next-generation uLAM tests is cost-effective in combination with sputum Xpert in PLHIV.		• Cost-effectiveness of next-generation uLAM tests in broader populations including in HIV negatives. • Cost-effectiveness of next-generation uLAM tests alone (not in combination with sputum Xpert).	Reddy et al. 2021;Fekadu et al. 2023; Ockhuisen et al. 2024; Yakhelef et al. 2020
14	What is the user-perspective of new LAM tests and the intended level of healthcare facility and training required?	On the example of FujiLAM, a next-generation LAM test is implementable in remote areas and decentralized settings.		• Manufacturers and implementors of next generation uLAM tests should conduct usability studies in the intended use setting (health posts in remote areas with limited infrastructure) with intended users (community health workers, nurses, clinicians with little training or even lay users).	Herrmann et al. 2022

No.	Question	Short answer	Status	Outstanding Questions and Future Considerations	References
15*	Emphasizing diagnostic yield and effective population coverage for WHO policy development.	Broger et al. (2024) provides clear definitions, examples, and study guidance for diagnostic yield evaluation and makes a strong case for inclusion of diagnostic yield in Grading of Recommendations, Assessment, Development and Evaluations (GRADE) policy processes		• Diagnostic yield should be assessed for novel non-sputum TB test. • Diagnostic yield should be included in the policy development of TB tests.	Broger et al. 2024

*Status column:  =answered,  =mostly answered,  =not answered. The focus of this dissertation was to tackle research areas 7 - clinical concentration ranges, 8 - diagnostic yield, and 15 - emphasizing diagnostic yield and effective population coverage (marked with *). LAM=Lipoarabinomannan. uLAM=urinary LAM. AlereLAM=Alere Determine TB LAM. FujiLAM=Fujifilm Silvamp TB LAM. HIV=Human Immunodeficiency Virus. LOD=Limit of Detection. Mtb=Mycobacterium tuberculosis. PLHIV=People living with HIV. POC=Point-of-care. SSM= Sputum Smear Microscopy. GRADE=Grading of Recommendations, Assessment, Development and Evaluations.*

4.1.1 Research area No. 1 - What is the molecular structure of LAM *in vivo*, what are most promising epitopes to target with existing and new antibodies to maximize performance?

A combination of two antibodies performs best for *in vivo* uLAM detection in immunoassays: (a) the S4-20 capture antibody that targets the unique 5-methylthio-d-xylofuranose (MTX)-dependent epitope in LAM that is specific to *Mtb* complex and shows no cross-reactivity with fast-growing mycobacteria or other bacteria, and (b) the A194-01 detection antibody that possesses high affinity for linear tetra-arabinoside (Ara4) and branched hexa-arabinoside (Ara6) structures in the arabinan domain of LAM and also binds to mannose (Man)-capped structures (Choudhary et al. 2018; Sigal et al. 2018b). The Lowary lab has been instrumental in developing an array of synthetic mycobacterial glycans for epitope mapping of anti-LAM antibodies (Zheng et al. 2017). Several papers build on the glycan arrays and this initial work and summarize currently available antibodies and their epitope binding properties (Cantera et al. 2022; Corrigan et al. 2022). It is important to note that the MTX motif targeted by S4-20 is only present at a ratio of 1 to 2 per LAM molecule (De et al. 2021; Lowary and Achkar 2022). Work on improved antibodies is ongoing (Corrigan et al. 2022; Yan et al. 2021). Recently, *in vivo* uLAM was purified from a large amount of patient urine which is critical for developing and/or assessing new uLAM antibodies as *in vivo* LAM is structurally different from culture-derived LAM (Cantera et al. 2023).

4.1.2 Research area No. 2 - Is urine or blood the better specimen for LAM detection?

Based on current evidence, urine is the better specimen for LAM detection. I led a study that compared serum LAM to urine LAM concentrations in matched samples from smear-positive TB patients (Broger et al. 2019b). The serum assay detected LAM in 55% of smear-positive patients with concentrations of 6 pg/mL to 70000 pg/mL, but 45% of patient's LAM concentrations were below the cut-off of 6 pg/mL, suggesting a median concentration in serum of roughly 10 pg/mL. In the same set of patients, the uLAM assay with a cut-off of 11 pg/mL showed that nearly all patients (93%) had concentrations in the range of 12 pg/mL to 90000 pg/mL (median 111 pg/mL) (Bulterys et al. 2019). In summary, data suggest that median LAM concentrations in urine are roughly ten times higher than in serum.

4.1.3 Research area No. 3 - Is there a difference between LAM detection in fresh vs. biobanked urines?

Current evidence suggests that there are only minor differences between LAM detection in fresh and biobanked urines. In a study with samples from 182 patients that I co-led, FujiLAM test results from urine samples kept at 2-8°C and tested within 4 hours were compared to results from aliquots from the same samples that were stored for 1 month at -20°C in

polypropylene tubes. Overall categorical agreement was 93.4% (95% CI 88.8–96.2) with positive and negative percent agreements of 85.2% and 97.5%, respectively (Broger et al. 2020a). In conclusion, the stability and availability of LAM, as detected by the FujiLAM assay, are high. The use of biobanked specimens delivers nearly equivalent results to the use of fresh specimens (Broger et al. 2020a). A small reduction in positive percentage agreement suggests that marginal increases in diagnostic sensitivity are possible in fresh samples (Broger et al. 2020a). In summary, urines from biobanks can be used for the development of LAM tests, which simplifies test development and comparative performance studies. Prospective studies with fresh samples should be used once test development has been completed.

4.1.4 Research area No. 4 - Is there a difference between LAM detection in early-morning vs. spot urines?

There are no major differences between LAM detection in early-morning and spot urines. I contributed to a study that compared FujiLAM sensitivity and specificity in spontaneously voided urine samples collected at inclusion (spot urine) versus urine first voided in the early morning (morning urine) and for a one (spot urine) versus two samples (spot and morning urine) strategy (Bjerrum et al. 2022). Overall agreement for spot versus morning urine test results was 94.6% with a Cohen's Kappa of 0.81 suggesting almost perfect agreement (Bjerrum et al. 2022). Therefore, the study indicates that FujiLAM testing performs equivalently for spot and early morning urine samples. A recent, large prospective study showed similar performance for Day 1 spot versus Day 2 early morning urine for both AlereLAM and FujiLAM (Székely et al. 2022).

4.1.5 Research area No. 5 - Can a highly sensitive reference assay for LAM be developed?

I was able to lead the development of a highly sensitive reference assay for uLAM detection. This laboratory-based ultrasensitive electrochemiluminescence LAM research assay (EclLAM), using the best antibody pair (see section 4.1.1), achieved a limit of detection (LOD) of 11 pg/ml (Sigal et al. 2018). Later I and my co-lead added an ultrafiltration step to concentrate urine 7-fold to reach an even lower LOD of 1.6 pg/ml (Broger et al. 2020b). Cantera et al. (2022) and Hoa et al. (2023) both used the same EclLAM assay but without ultrafiltration and reported similar LODs of of 3.3 pg/ml and 10 to 63 pg/ml, respectively.

4.1.6 Research area No. 6 - Can low-cost, simple, and rapid platform alternatives be developed that substantially improve POC detection compared to conventional lateral-flow immunoassays (LFIAs)?

My development of FujiLAM (see section 2.5.3) showed that improved alternatives to conventional gold nanoparticle-based LFIAs (such as AlereLAM) can be developed. FujiLAM's

silver amplification step enables the detection of uLAM concentrations approximately 30 times lower than those detected by AlereLAM (Broger et al. 2019a). In 2020, I led a meta-analysis of the diagnostic accuracy of FujiLAM in 1595 adult PLHIV and showed overall sensitivity of 70.7% for FujiLAM compared to 34.9% for AlereLAM and a specificity of 90.9% for FujiLAM compared to 95.3% for AlereLAM when using a microbiological reference standard (Broger et al. 2020c).

Recent results from two large prospective multicentre studies by Székely et al. 2022 (1624 participants) and Huerga et al. 2023b (1575 participants) came close to confirming the initial performance, but showed significant lot-to-lot variability with FujiLAM. Overall sensitivity was 54.8% and 60% and specificity 85.1% and 87% in the studies by Székely et al. and Huerga et al., respectively. However, performance by lot ranged from 33% to 74% sensitivity and 71% to 96% specificity, and a post-hoc comparison of different lots with 118 samples and the EclLAM assay showed positivity rates varying from 13 to 77% between lots (Székely et al. 2022). This suggests that it is challenging to consistently achieve such a low cut-off across production lots with a point-of-care assay, an aspect that requires close attention not just for FujiLAM but for every next-generation LAM test.

Nevertheless, my LAM work including the FujiLAM development paved the way for the development of several other next-generation uLAM tests including Flow-TB (Salus Discovery, USA), High-sensitivity TB LAM (Abbott, USA), ichroma LAM Ag (Boditech, South Korea), Standard F TB LAM Ag FIA (SD Biosensor, South Korea), TB LAM urine LFA (Biopromic, Sweden), and AIMLAM (Leide Biosciences, China) (Huang et al. 2023; Treatment Action Group (TAG) 2023).

4.1.7 Research area No. 7 - Is LAM detectable in people with TB without HIV co-infection and what are concentration ranges and accuracy?

So far, the focus of uLAM testing and guidelines has been on HIV co-infected individuals with severe immunocompromise and/or disseminated TB. Publication 1 (Broger et al. 2020b) of this cumulative dissertation was, at the time of publication, the first study that reported clinical concentration ranges of LAM in a cohort of 372 HIV negative people comparing the highly sensitive EclLAM reference assay (described in research area 5), the next-generation FujiLAM POC test (from research area 6), and the AlereLAM POC test. The study provides a clear answer to the question from Figure 4, “What is the cut-off that a more sensitive LAM test needs to reach to achieve the ≥65% sensitivity WHO TPP target in HIV-negative patients?”: **a cut-off of around 5 pg/ml or below is required to meet the TPP sensitivity target of ≥65%**. At this cut-off, the EclLAM reached 66.7% sensitivity and 98.1% specificity.

Table 3 summarizes uLAM concentration cut-offs in relation to expected test sensitivity in HIV negative people with TB which provides a valuable resource for developers of next-generation LAM tests to set the analytical sensitivity product requirement.

Table 3: uLAM concentration cut-off in relation to expected test sensitivity in HIV negative people with TB.

uLAM concentration cut-off (based on EclLAM and for the used antibodies)	**Percentage of HIV negative people with TB with uLAM concentration above the cut-off**	**Comment**
500 pg/ml	10%	Approximate AlereLAM cut-off and performance
10-20 pg/ml	50%	Approximate FujiLAM cut-off and performance
5 pg/ml	67%	Approximate cut-off required to meet the TPP (98% specificity for EclLAM assay)
1.6 pg/ml	77%	Lowest cut-off studied with the EclLAM. Specificity 90% which is below the TPP specificity target of 98%.

Based on Broger et al. (2020b). Cut-offs are estimates based on the EclLAM assay using A194-1 and S4-20 antibodies and non-standardized LAM calibration material. Thus cut-offs might not be generalizable to other LAM assays with different antibodies, detection technologies, or LAM calibration material. uLAM=urinary Lipoarabinomannan. EclLAM= Laboratory-based ultrasensitive electrochemiluminescence LAM research assay. HIV=Human Immunodeficiency Virus. AlereLAM= Alere Determine TB LAM. FujiLAM=Fujifilm Silvamp TB LAM. TPP=Target product profile. TB=Tuberculosis.

FujiLAM with a cut-off of 10-20 pg/ml came close to the TPP target and identified 53.2% of positive TB cases, representing a 5-fold increase in diagnostic sensitivity among HIV-negative patients compared with AlereLAM and an approximately 25-fold improvement of analytical sensitivity (Broger et al. 2020b). The AlereLAM sensitivity of 10.8% from my study is close to the 16% sensitivity in immunocompetent PLHIV (Bjerrum et al. 2019), suggesting that sensitivity in HIV uninfected people with TB is similar to that of immunocompetent PLHIV.

Since my pivotal study, others studied clinical concentration ranges of LAM in people with TB without HIV-coinfection. Hoa et al. (2023) measured uLAM concentrations with the same EclLAM assay using the same antibodies as ours in 692 participants with pulmonary TB (PTB) in Vietnam reaching a sensitivity of 39% (at a specificity of 97%) with an average cut-off 23 pg/ml (range 10 to 63 pg/ml depending on the immunoassay plate). In contrast to my study, Hoa et al. reported lower sensitivity not meeting the TPP, but a number of factors could explain this. First their cut-off was on average roughly 5-fold above the cut-off of my study (23 pg/ml vs. 5 pg/ml), primarily because the study did not use a sample concentration step like

the ultrafiltration in my study. If I apply the average cut-off of 23 pg/ml from Hoa et al. to my dataset (Figure 3, Broger et al. 2020c) I get a very similar sensitivity (≈40%) suggesting that their slightly less sensitive assay may miss roughly 25% of people with TB with very low LAM concentrations. Next, TB bacterial burden was identified as a driver for LAM positivity and the much lower percentage of sputum smear microscopy positive (SSM) TB patients in the Hoa study (39% (131/335)), over half which had scanty/1+ microscopy grading suggesting low bacillary burden and subclinical TB in a large proportion of participants. This contrasts with my study with 61% (68/111) SSM positive patients and could further explain the lower performance of the EclLAM in the Hoa study compared to ours. In this context, it is important to highlight that in my study, EclLAM performed better in Peruvian participants with TB who had shorter mycobacterial growth indicator tube liquid culture time to detection (MGIT TTD) and a larger proportion of patients had positive SSM and Xpert results compared with patients from South Africa, suggesting higher mycobacterial load in sputum (Broger et al. 2020b). The SSM proportion, Xpert, and MGIT TTD results suggest more advanced disease in the Peruvian cohort but relatively low burden in South Africa, which likely explains the lower EclLAM sensitivity in Peru (78.5%) compared to South Africa (37.5%). This is in-line with the finding of Hoa and colleagues, who suggested that the relatively low TB bacterial burden and presence of subclinical TB in a large fraction of participants has likely led to an underestimation of EclLAM sensitivity. In summary, the lower sensitivity of EclLAM in the Hoa study can be explained by differences in assay cut-off and study population.

In another recent study, Huang et al. (2024) tested cryopreserved urine samples from HIV-negative individuals with presumed TB (n=144) and healthy donors with a quantitative uLAM antigen strip. At a cut-off of 0.134 arbitrary units which is close to the LOD of 50 pg/ml (personal communication with the author) the group reported a sensitivity of 52% at a specificity of 96%. Although, a direct comparison to my study is complicated by the different methodology, antibodies, and LAM calibration material, performance comes very close to the performance of the FujiLAM in my study. This further substantiates my findings that a cut-off in the low pg/ml range is needed to reach clinically meaningful sensitivities.

Finaly, Huang et al. (2023) conducted a retrospective study measuring uLAM in 166 HIV-negative individuals with TB (n=166), compared to 22 participants with NTM (n=22), 69 Non-TB (n=69) with pulmonary diseases other than TB, and 73 samples from healthy controls (n=73) in a Chinese hospital. Their assay employed magnetic beads to enrich uLAM from 1.5 ml of urine for subsequent detection on a chemiluminescence analyser with an assay from Leide Biosciences (China) reaching a sensitivity of 51% and a specificity of 96%. Although the study did not report uLAM concentrations nor a cut-off, results come close to my findings with the FujiLAM (53.2% sensitivity) and the uLAM antigen strip from Huang et al. (2024) (52% sensitivity).

There are a number of other, typically smaller case-control studies, that also showed that lower detection limits translate into higher diagnostic sensitivity in TB people without HIV (Magni et al. 2020; Paris et al. 2017; Wood et al. 2019). However, an absolute comparison of LAM concentrations and test performance is complicated by the different methodologies

used. In general, it is important to mention that the cut-off requirement of 5 pg/ml to reach the TPP sensitivity target of ≥65% is an estimate based on the EclLAM assay and non-standardized LAM calibration material and might not be generalizable to other LAM assays with different antibodies, detection technologies, or LAM calibration material (Broger et al. 2020b). Establishing biological reference materials and sample panels, as has been done by the WHO for other diseases, is an urgent priority to support the development, validation, and comparison of current and future LAM assays (World Health Organization 2018).

Taken together, studies suggest that a highly sensitive LAM assay with a cut-off around 5 pg/ml is needed to reach the TPP sensitivity target of ≥65%.

4.1.8 Research area No. 8 - What are the diagnostic yields of first and next-generation LAM tests alone and in combination with sputum smear microscopy or sputum Xpert?

Sputum is the most widely used sample to diagnose active tuberculosis, but many people are unable to produce sputum. Urine, in contrast, is readily available. In Publication 2 of this dissertation, I hypothesised that sample availability influences the diagnostic yield of TB tests (Broger et al. 2023). When doing all the uLAM research, I realized that the availability of the diagnostic specimen is of upmost importance but that sensitivity doesn't account for it. In my meta-analysis of 20 datasets with 10202 participants I found that nearly all (98%) participants provided urine, and only 82% provided sputum within 2 days. 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%) and FujiLAM (65%) were unaffected by sample provision as samples were readily obtained from almost all PLHIV (Broger et al. 2023). Urine AlereLAM in combination with sputum smear microscopy detected 54% and in combination with sputum Xpert 75% of TB cases (Broger et al. 2023). Eleven of the datasets were also incorporated in a Cochrane systematic review (in preparation, 12651 participants) on the parallel use of sputum Xpert and urine AlereLAM, finding that parallel use of the two tests could diagnose TB in 77.5% of PLHIV which is in line with my findings (Bjerrum et al. 2024).

A main finding of my work is that it is time to change how we think about tuberculosis (TB) test sensitivity and also assess the diagnostic yield of TB tests and testing algorithms. Policy makers like the World Health Organization (WHO), *in vitro* diagnostic regulators, test manufacturers and the scientific community are currently almost exclusively focused on TB test sensitivity. Sensitivity measures how well an individual assay can detect a biomarker in a specimen of a person with the condition of interest. Critically, this measure neglects whether a person has access to testing, the obtainability of the specimen for testing, and whether the specimen and result are available in time to take action to positively impact the lives of people and their communities. I therefore decided on emphasizing diagnostic yield and effective population coverage for WHO policy development in Publication 3, a health policy paper (see research area 15, below).

4.1.9 Research area No. 9 - LAM in children

The 2019 WHO LAM recommendation to use AlereLAM to assist in the diagnosis of active TB in PLHIV with symptoms of TB, advanced HIV diseases, those who are seriously ill, and in inpatients with CD4 count <200 cells/µL irrespective of TB symptoms and outpatients with CD4 count <100 cells/µL also applies to adolescents and children living with HIV (World Health Organization 2019). However, the policy notes that this is based on generalization of data from adults, while acknowledging very limited data for these populations. AlereLAM sensitivity and specificity ranged from 42% to 56% and 80% to 95%, respectively, in three paediatric studies that were mentioned in the Cochrane review related to the WHO 2019 recommendation (Bjerrum et al. 2019; Kroidl et al. 2015; LaCourse et al. 2018; Nicol et al. 2014).

For FujiLAM testing in children, Olbrich et al. (2022), conducted a systematic review and found three studies with sensitivities and specificities ranging from 42% to 63% and 84 to 93%, respectively, against a microbiological reference standard (Comella-del-Barrio et al. 2021; Nicol et al. 2020; Nkereuwem et al. 2020). Nicol et al. 2020 and Nkereuwem et al. 2020 compared AlereLAM to FujiLAM and both found FujiLAM to have higher performance and concluded that the high specificity of FujiLAM suggests utility as a "rule-in" test for children with a high pretest probability of TB, including hospitalized children living with HIV or with malnutrition. In summary, the evidence of LAM in children remains limited.

4.1.10 Research area No. 10 - LAM in extrapulmonary TB (EPTB)

More sensitive uLAM detection with FujiLAM showed substantially higher sensitivity over AlereLAM for detecting extra-pulmonary TB disease (EPTP) in HIV inpatients. In a study with 872 patients that I co-led, FujiLAM was able to detect TB in 67% of EPTB patients. Sensitivities were high for several EPTB sites of disease including 94% in TB mycobacteraemia, and 88% in TB mycobacteriuria (Kerkhoff et al. 2020). The study further demonstrated moderate sensitivity in patients with microbiologically-confirmed pleural TB (68%) and with TB meningitis (47%). Overall, FujiLAM showed substantially higher sensitivity over the commercially available AlereLAM for detecting EPTB in HIV inpatients, suggesting there is a high potential for next-generation LAM tests for the diagnosis of EPTB. FujiLAM's excellent performance in those with mycobacteraemia suggests a mechanistic association between disease dissemination and urinary LAM. This finding is supported by my recent study that showed a good association between detection of LAM in urine and serum of TB patients, independent of HIV status (Broger et al. 2019b). However, even for patients with forms of disease such as pleural TB and TB meningitis that may be compartmentalised, uLAM had moderate sensitivity, which could add substantial benefit in these cases. Taken together, these findings suggest that LAM antigenuria is likely indicative of glomerular filtration of circulating LAM (or LAM fragments) in addition to renal TB (Lawn and Gupta-Wright 2016). Further research, to characterize LAM structure in urine is needed to better understand the mechanisms by which LAM enters the bloodstream and urine. This may help guide the

targeted development of urine-based diagnostics and catalyse the development of blood-based assays.

4.1.11 Research area No. 11 - LAM as a mortality predictor

Time to diagnosis and treatment is key towards decreasing TB-associated mortality. A more sensitive uLAM test, like FujiLAM could rapidly diagnose TB in up to 89% of those who died whereas the probability of three-months survival was above 86% in uLAM negative patients (Sossen et al. 2020). This suggests that the mortality impact of the AlereLAM may be further increased by a more sensitive uLAM assay by enabling rapid, POC diagnosis of TB in a larger proportion of patients at risk (see next section). Compared to AlereLAM, FujiLAM detected TB in 22 to 37% more PLHIV who died.

4.1.12 Research area No. 12 - What is the impact of a next-generation LAM POC test?

AlereLAM use has been shown to reduce all-cause mortality in high-risk groups of PLHIV. Results of a recent Cochrane meta-analysis by Nathavitharana et al. (2021) showed that hospitalized PLHIV who received AlereLAM testing had 15% lower risk of mortality at eight weeks compared to routine TB diagnostic testing without AlereLAM. The absolute effect in inpatients was 34 fewer deaths per 1000 with a pooled risk ratio of 0.85, 95% CI 0.76 to 0.94, based on data from 5102 participants from two randomized trials (Gupta-Wright et al. 2018; Peter et al. 2016). One trial in outpatients did not detect a difference in mortality but the direction of effect was towards a mortality reduction, and the effect size was similar to that in inpatient settings (risk ratio 0.89, 95% CI 0.71 to 1.11; 2972 participants from one randomize trial) (Grant et al. 2020). For other TB tests, like Xpert, there is no evidence of mortality impact (Haraka et al. 2021) which highlights the potential of LAM based testing, even though this mortality reduction is limited to PLHIV with high mortality risks whereas the Xpert study assessed impact in a broad population.

Modelling from Ricks et al. (2020) suggests that if AlereLAM is only used to test for TB amongst PLHIV, it would avert less than 1% of the total TB cases and deaths in South Africa over 15 years. However, the use of an improved, next-generation uLAM test amongst all people with symptoms, regardless of their HIV status, could avert 30% of TB deaths and 18% of TB cases within 15 years in South Africa. The analysis assumed a sensitivity for a next-generation LAM test of 30% in HIV negatives and 70% in HIV positives, both of which are achievable performances (see research areas 7 and 8).

4.1.13 Research area No. 13 - Is the use of a next-generation LAM test cost-effective?

I contributed to an evaluation of cost-effectiveness of FujiLAM for tuberculosis testing among hospitalized PLHIV irrespective of symptoms (Reddy et al. 2021). The study used a microsimulation model to project clinical and economic outcomes of three testing strategies: (1) sputum Xpert, (2) sputum Xpert + urine AlereLAM, and (3) sputum Xpert + urine FujiLAM as an example for a next-generation LAM test. Costs of Xpert/AlereLAM/FujiLAM were US$15/3/6 (South Africa) and $25/3/6 (Malawi). Xpert+FujiLAM was considered cost-effective if its incremental cost-effectiveness ratio (US$/year-of-life saved) was <$940 (South Africa) and <$750 (Malawi). Compared with Xpert+AlereLAM, Xpert+FujiLAM increased life expectancy by 0.2 years for those tested in South Africa and Malawi and was cost-effective in both countries (Reddy et al. 2021).

Fekadu et al. (2023) confirmed the findings that Xpert+FujiLAM for TB diagnosis in HIV-infected individuals is the preferred cost-effective strategy from the perspective of a South African health service provider in both inpatient and outpatient settings.

It is important to note that even the less-sensitive and WHO recommended AlereLAM+Xpert is cost-effective compared to Xpert alone (Ockhuisen et al. 2024). Including AlereLAM in TB diagnostic algorithms is cost-effective for severely ill or immunosuppressed HIV-positive patients (Yakhelef et al. 2020). In an analysis of the STAMP trial in HIV positive inpatients, AlereLAMs high incremental diagnostic yield and low additional cost compared with sputum Xpert alone make a compelling case for expanding its use (Reddy et al. 2019).

4.1.14 Research area No. 14 - What is the user-perspective of new LAM tests and the intended level of healthcare facility and training required?

Using FujiLAM as an example, I contributed to a qualitative study in two high TB/HIV burden countries to assess values and preferences of end-users, along with potential barriers for the implementation of a next-generation uLAM test (Herrmann et al. 2022). In 42 semi-structured interviews with patients, health care providers (HCPs) and decision makers (DMs) from Malawi and Zambia it was found that the ease and convenience of urine-based testing was empowering in light of patients with difficulty to collect sputum. HCPs expressed concerns that a shift in agency to the patient may affect clinical workflows (i.e. less control over collection). Shorter turnaround times were welcomed by operators and patients alike. The decentralization of diagnostics was considered possible with FujiLAM by HCPs and DMs due to the test's low infrastructure requirements. Finally, the findings support efforts for eliminating the CD4 count as an eligibility criterion for LAM testing to facilitate implementation and benefit a wider range of patients. Interviewees view next-generation uLAM implementation as a viable, acceptable, and likely sustainable option in low- and middle-income countries, though adaptations may be required to current health care processes for deployment (Herrmann et al. 2022).

Regarding training requirements, professionals who have operated the test consider it simple, and the steps are easy to follow using manufacturer's instructions (Herrmann et al. 2022). Therefore, while training is necessary prior to operating the test, no specific laboratory expertise is needed. In addition, HCPs consider FujiLAM easy to interpret, particularly when compared to AlereLAM. However, some users indicate that there are too many timed steps, and the test requires too much hands-on time (Herrmann et al. 2022). Consequently, professionals suggested that FujiLAM should be operated by technicians in the lab, not by nurses or clinicians (e.g., at the bedside) as they hold other roles which could be interrupted while waiting for the timed steps (Herrmann et al. 2022).

Regarding level of healthcare facility, FujiLAM can be run without relying on much additional equipment or maintenance (besides personal protective equipment, urine cup, etc.). This is contrasted with Xpert's infrastructure requirements, whose machinery and modules need installation, electricity, and maintenance. For these reasons, interviewees consider FujiLAM suitable to use in remote areas and decentralized laboratories in Malawi and Zambia. Some of them recommend that implementation should be prioritized in facilities that do not have technologies such as Xpert and microscopy to expand the availability of TB diagnostics in remote areas (Herrmann et al. 2022).

4.1.15 Research area No. 15 - Emphasizing diagnostic yield and effective population coverage for WHO policy development.

My IPD meta-analysis of diagnostic yield (research area 8, Broger et al. (2023)) clearly showed the need to emphasize diagnostic yield and effective population coverage for WHO policy development which I addressed in Publication 3 of this cumulative dissertation (Broger et al. 2024). The objective of this health policy paper is to initiate a rethinking of the methods used to assess TB tests. Conventional diagnostic sensitivity and specificity miss important aspects of tests (such as specimen availability) that are better covered by diagnostic yield and effective population coverage. My approach was to get as many TB key opinion leaders as possible on board and seek consensus on diagnostic yield. This paper proved to be the most challenging article of my career, primarily due to the diverse perspectives among the 21 co-authors which required 14 iterations to reach a consensus. The interaction helped sharpen the concept of diagnostic yield and how to evaluate it, which was critical to convey the right message and to the quality of the paper. I will refrain from delving into the significance of diagnostic yield and effective population coverage in this discussion, as these aspects have been thoroughly addressed in the publication above and revisiting them here would be redundant. The influence of this paper hinges on the extent to which stakeholders assess diagnostic yield in clinical studies, communicate findings, and subsequently integrate diagnostic yield and effective population coverage into health policy development. It is encouraging to observe that some stakeholders are beginning to explore and write about this important subject. In "Transforming Tuberculosis Diagnosis" Pai et al. (2023a), one of the co-authors of my paper and a key opinion leader, proposes a transition from test accuracy to yield and population

coverage as one of the seven transitions needed to close the TB diagnostic gap. Although the WHO is still putting a lot of focus on sensitivity and specificity (and not yield) in its ongoing LAM policy update, combined sensitivity of parallel sputum Xpert and uLAM testing is a focus and sample provision and indeterminates (both aspects of yield) were reported for the first time (Bjerrum et al. 2024). The recent WHO standard on universal access to rapid tuberculosis diagnostic is a major step towards highlighting that effective population coverage for WRDs needs to be improved (Ismail et al. 2023; World Health Organization 2023b).

Although this dissertation focusses on uLAM, diagnostic yield is also highly relevant for other non-sputum based TB tests in the pipeline. The diagnostic industry saw notable levels of investment and accelerated research during the COVID-19 pandemic which sparked a wave of innovation, summarized in Yerlikaya et al. (2023b). Simple non-invasive sampling methods such as swabs have become widely used down to the community setting. Even self-testing has become common and facilitated high diagnostic population coverage in many geographies during the pandemic. These innovations can now be leveraged for TB diagnosis and could help to push diagnostic yield and population coverage to a totally new level. Other than urine-based testing, tongue swab sampling combined with PCR for TB detection on simple, lower cost platforms is an area with extensive development activities (Andama et al. 2022; Church et al. 2024; Steadman et al. 2024). Transitioning from a sole focus on accuracy to a more nuanced consideration of diagnostic yield and coverage will be crucial in evaluating these new, potentially transformative tools for TB diagnosis. Failing to do so might result in overlooking their impact on people's lives, particularly if they are unable to reach the same sensitivity than sputum-based tests.

4.2 Is LAM a suitable biomarker for TB diagnosis in general populations? - TPP fulfilment of a predicted next generation uLAM test

Niels Bohr's quote, "Prediction is very difficult, especially if it's about the future," underscores the inherent uncertainty in forecasting what is to come. However, despite this, it's important to dare to make predictions, particularly when it comes to significant advancements like the development of a next-generation uLAM test meeting the TPP. Attempting to predict the success or failure can help identify the critical aspects that will determine its outcome. While acknowledging the difficulty, it's through this process that we can uncover the essential factors that contribute to success and avoid potential pitfalls.

To make it short, my prediction is **yes; LAM is a suitable marker for TB diagnosis in general populations and a future test can meet the TPP!** Ten years after embarking on my LAM strategy, there has been clear progress for all 15 areas of research and development suggesting a pathway towards a next-generation LAM TB test that meets the TPP. LAM is present in the urine of at least 66.7% of people with TB without HIV co-infection. However, reaching such sensitivity requires highly sensitive detection. Therefore, a critical success factor, or even potential pitfall, is whether a simple-enough, rapid uLAM POC test with an extremely low and stable cut-off of around 5 pg/ml or below, a requirement for detecting

uLAM in general populations and reaching sufficient sensitivity, can be developed. FujiLAM, with its silver amplification step and with a cut-off of 10-20 pg/ml, comes close but recent studies have demonstrated major issues in achieving lot-to-lot reproducibility of such a low cut-off (Huerga et al. 2023; Székely et al. 2022) with a simple instrument-free assay, as very weak test lines need to be interpreted with the naked eye. FujiLAM is pushing the boundaries of what is possible with LFIAs, and I predict that a conventional LFIA, with gold-labelled antibodies, will not get us to the goal target. Alternative labels (e.g. fluorescent, luminescent, or upconverting phosphors) or amplification (e.g. like the FujiLAM silver amplification) are necessary to achieve a reliable cut-off in decentralized settings and in the hands of different users. This requires the use of a reader instrument. An instrument readout also facilitates signal normalization strategies such as using the control band signal as a normalizer for the test band signal. This could lower the variance of the signal, which is typically high so close to the limit of detection of an assay.

Approaches for uLAM analyte concentration via sample processing prior to detection have been proposed. However, such approaches add complexity and training requirements, steps that take time and make the test expensive. Therefore, I do not expect sample concentration to be successful unless it is highly integrated, simple and fast (e.g. a simple magnetic-bead based concentration step). Table 4 summarizes the TPP fulfilment of my predicted, hypothetical future uLAM test using the same criteria as Table 1. But what if developers do not get there with the sensitivity? In this case, modelling and the consensus of the publication 3 strongly suggest to consider the TPP sensitivity target in context because of the following arguments:

(a) sensitivity does not factor in diagnostic yield and undervalues the benefits around the highly available urine specimen,

(b) a simple antigen LFIA with a simple reading device has the potential to reach high population coverage if it is low-cost,

(c) even the AlereLAM with very low sensitivity has shown impact in PLHIV and a next-generation more sensitive uLAM test will have a higher impact, potentially even in other populations, and

(d) the uLAM test can be used to rule-in TB and start treatment, potentially during the same clinical encounter as we wait for results from the slower, more expensive, sputum-based molecular tests.

In summary, my research shows that that there is high promise to achieve a breakthrough in TB diagnostics if researchers, developers, donors, advocates, and policymakers are persistent and intensify investment into uLAM tests despite the recent challenges around FujiLAM. True innovation with significant impact often requires time and multiple iterations, as has been the case with many other diagnostic innovations.

Table 4: TPP fulfilment of a predicted next generation uLAM test

Test (Product names)	**Principle**	**Healthcare level**	**Specimen**	**WHO recom-mended**	**A** **Affordable**	**S** **Sensitive**	**S** **Specific**	**U** **User-friendly**	**R** **Rapid**	**E** **Equipment free**	**D** **Delivered to end-users**
Next Generation uLAM test	Highly sensitive detection of Lipoarabinomannan antigen	L0, L1, L2, L3	Urine	yes	yes	yes	yes	yes	yes	no	unclear
Target Product Profile Minimal Requirement		L1, L2, L3	Non-sputum		<$6	>65%	>98%	yes	<1 hour	yes	yes

For healthcare system levels, L0 – local community health post or care by local healthcare worker, L1 – microscopy centre or primary healthcare centre, L2 – district hospital or community health centre, L3 – reference or tertiary hospital or laboratory. Target product profile minimal requirements are based on Denkinger et al. (2015) and World Health Organization (2014). ASSURED criteria are based on Mabey et al. (2004). TPP criteria are described in the bottom row of the table. Green highlighted are test criteria that meet the TPP requirements and red criteria that do not. uLAM=urinary Lipoarabinomannan. TPP=Target product profile. WHO=World Health Organization.

4.3 Conclusion

LAM is one of the most promising tuberculosis (TB) diagnostic biomarkers, and significant progress has been made in the last decade across 15 areas, including the three areas that form the core of this cumulative dissertation. These advancements suggest that ultra-sensitive urinary LAM (uLAM) detection could form the basis for the first non-sputum TB test meeting the World Health Organization's TPP, potentially having a significant impact on people's lives. As we prioritize the development of this 'holy grail' TB test, it is crucial to ensure the uptake of the first-generation AlereLAM test, which will spur the adoption of next-generation tests.

5 SUMMARY

Background and research questions: This cumulative doctoral dissertation aims to investigate whether urinary Lipoarabinomannan (uLAM) is a suitable biomarker for tuberculosis (TB) diagnosis in general populations. The uLAM-based Alere Determine TB LAM test (Abbott, USA) is currently the only test meeting operational characteristics of the Target Product Profile of the World Health Organization (WHO) and being recommended to assist in the TB diagnosis of patients with HIV co-infection. Its use in general populations is currently not recommended due to the low sensitivity of the test. At the core of this dissertation are three underlying and interconnected research questions, that enable the development and use of uLAM-based tests:

(1) What are the clinical concentration ranges of Lipoarabinomannan (LAM) in urine samples in general (non-HIV) patient populations?

(2) What is the diagnostic yield of urine LAM tests compared to sputum-based tests in people living with human immunodeficiency virus (HIV)?

(3) How can diagnostic yield be incorporated into policy development for novel TB tests?

Methods and findings: The three publications address these research questions:

(1) A clinical multi-centre diagnostic accuracy study in South Africa and Peru that measured LAM concentrations in urine samples from HIV-negative outpatients using a highly sensitive electrochemiluminescence assay and comparing performance to two point-of-care uLAM tests, the existing Alere Determine TB LAM, and the novel Fujifilm Silvamp TB LAM tests (Fujifilm, Japan). The study found that 66.7% of people with TB without HIV co-infection had uLAM concentrations ≥5 pg/ml when using the highly sensitive electrochemiluminescence assay. Fujifilm Silvamp TB LAM, with a cut-off of 10-20 pg/ml, came close to the Target Product Profile (TPP) and identified 53.2% of positive TB cases, representing a 5-fold increase in sensitivity among HIV-negative patients compared with AlereLAM at 10.8% sensitivity.

(2) A systematic review and meta-analysis of individual participant data from existing TB studies employing Bayesian random-effects and mixed-effects meta-analyses to predict the diagnostic yield of urine LAM tests and sputum-based TB tests in people living with HIV. The study, based on 20 datasets with 10202 participants, found that nearly all (98%) participants provided urine, and only 82% provided sputum within 2 days. Due to the lower availability of sputum, the diagnostic yields of sputum Xpert MTB/RIF (Cepheid, USA) (61%) and sputum smear microscopy tests (32%) were lower than their sensitivities. By contrast, the diagnostic yield of urine Alere Determine TB LAM (41%) and Fujifilm Silvamp TB LAM tests (65%) were unaffected by sample provision as urine was readily obtained from almost all people living with HIV.

(3) A health policy review and guidance paper that emphasized the importance of diagnostic yield as a measure of the utility of novel TB tests, especially those that use non-sputum specimens. This paper provides clear definitions and examples of diagnostic yield, highlights its relevance for different TB populations and settings, and proposes ways to incorporate and

evaluate diagnostic yield in the World Health Organization (WHO) guideline development process.

Discussion: The discussion provides an overarching update on 15 LAM research and development areas that form a research strategy and bring the three publications of this dissertation into context of the agenda. After a decade of dedicated research following my LAM strategy, I and others made progress on all 15 research and development areas, paving the way for a next-generation LAM-based TB diagnostic test. The discussion concludes that LAM is a suitable marker for TB diagnosis in general populations and a future test can meet the TPP. However, a critical success factor and possible pitfall is the development of a simple-enough, rapid uLAM point-of-care (POC) test with an extremely low cut-off of around 5 pg/ml, which is required to reach sufficient sensitivity in an unselected population of people with presumed TB. Even if a test does not reach the required sensitivity of 65%, modelling suggests to prioritize uLAM test development, as the benefits of highly available urine-based point-of-care tests can be substantial and tests can reach high population coverage as long as it is low cost.

Conclusion: LAM is one of the most promising tuberculosis diagnostic biomarkers, and significant progress showing its potential has been made. These advancements suggest that ultra-sensitive uLAM detection could form the basis for the first non-sputum TB test meeting the World Health Organization's Target Product Profile (TPP), potentially transforming TB diagnosis.

6 ZUSAMMENFASSUNG

Hintergrund und Forschungsfragen: Diese kumulative Dissertation untersucht, ob Lipoarabinomannan im Urin (uLAM) ein geeigneter diagnostischer Biomarker für die Diagnose von Tuberkulose (TB) ist. Der auf uLAM-Detektion basierende Alere Determine TB LAM Schnelltest (Abbott, USA), ist momentan der einzige Test, der die operativen Anforderungen des Zielproduktprofils (Target Product Profile, TPP) der Weltgesundheitsorganisation erfüllt und wird derzeit zur Unterstützung bei der TB-Diagnose bei Patient*innen mit HIV-Coinfektion empfohlen. Eine breitere Anwendung in allgemeinen Patient*innenpopulationen wird derzeit aufgrund der niedrigen Sensitivität nicht empfohlen. Diese Dissertation konzentriert sich auf drei miteinander verbundene Forschungsfragen, die die Weiterentwicklung von uLAM-basierten Tests und deren breitere Nutzung untersuchen:

(1) Was sind die klinischen Konzentrationsbereiche von Lipoarabinomannan (LAM) in Urinproben von Patient*innen mit Verdacht auf TB in allgemeinen Patient*innenpopulationen, mit Fokus auf Patient*innen ohne HIV-Coinfektion?

(2) Was ist die diagnostische Ausbeute von uLAM-Tests im Vergleich zu Sputum-basierten Tests?

(3) Wie kann die diagnostische Ausbeute in die Entwicklung von Leitlinien für neuartige TB-Tests integriert werden?

Methoden und Ergebnisse: Es werden drei Publikationen vorgestellt, um diese Forschungsfragen zu adressieren:

(1) Eine klinische multizentrische Studie zur diagnostischen Genauigkeit in Südafrika und Peru, die uLAM-Konzentrationen in 372 HIV-negativen Patient*innen mit einem hochsensitiven quantitativen Test nachwies und die Genauigkeit mit zwei qualitativen uLAM-Schnelltests verglichen hat: dem Alere Determine TB LAM- und dem Fujifilm Silvamp TB LAM-Test (Fujifilm, Japan). Die Studie ergab, dass 66.7% der Personen mit TB ohne HIV Co-Infektion uLAM-Konzentrationen von ≥5 pg/ml aufwiesen, wenn der hochsensitive Elektrochemilumineszenz-Assay verwendet wurde. Fujifilm Silvamp TB LAM mit einem Schwellenwert von 10-20 pg/ml kam dem Target Product Profile (TPP) der WHO nahe und identifizierte (bei hoher Spezifizität) 53.2% der TB-Fälle, was im Vergleich zu Alere Determine TB LAM mit einer Sensitivität von 10.8% eine 5-fache Sensitivitätssteigerung bei HIV-negativen Patient*innen darstellte.

(2) Eine systematische Übersichtsarbeit und Metaanalyse von individuellen Teilnehmer*Innendaten aus vorhandenen TB-Studien, die Bayesian Random-Effects- und Mixed-Effects-Metaanalysen verwendete, um die diagnostische Ausbeute von uLAM- und Sputum-basierten TB-Tests bei Menschen mit HIV vorherzusagen. Die auf 20 Datensätzen mit 10202 Teilnehmer*Innen basierende Studie ergab, dass fast alle (98%) Teilnehmer*Innen Urinproben bereitstellten konnten, jedoch bei nur 82% Sputum innerhalb von 2 Tagen verfügbar war. Aufgrund dieser geringeren Verfügbarkeit von Sputum waren die diagnostischen Ausbeuten von Sputum Xpert MTB/RIF (Cepheid, USA) (61%) und der Sputum-Mikroskopie (32%) und somit niedriger als ihre Sensitivitäten. Im Gegensatz dazu waren die

diagnostischen Ausbeuten von Alere Determine TB LAM (41%) und Fujifilm Silvamp TB LAM-Tests (65%) von der Probennahme unbeeinflusst, da Urin von fast allen Studienteilnehmer*Innen verfügbar war.

(3) Ein Leitfaden, der die Bedeutung der diagnostischen Ausbeute als Messgröße für den klinischen Nutzen neuartiger TB-Tests betonte, insbesondere solcher, die alternative Probentypen mit hoher Verfügbarkeit im Vergleich zu Sputum verwenden. Der Artikel enthält klare Definitionen und Beispiele zur diagnostischen Ausbeute, hebt deren Relevanz für verschiedene Populationen hervor und zeigt auf, wie die diagnostische Ausbeute als wichtige Messgröße in den Prozess der Leitlinienentwicklung der Weltgesundheitsorganisation (WHO) integriert werden kann.

Diskussion: Die Diskussion bietet, aufbauend auf meiner LAM-Forschungsstrategie, ein umfassendes Update zu 15 LAM-Forschungs- und Entwicklungsgebieten und integriert die drei Publikationen, die den Kern dieser Dissertation bilden. Nach einem Jahrzehnt intensiver Forschung basierend auf der LAM-Forschungsstrategie, wurden in allen fünfzehn Forschungs- und Entwicklungsgebieten Fortschritte erzielt, was den Weg für einen LAM-basierten, sensitiveren TB-Schnelltest der nächsten Generation aufzeigt. Die Diskussion kommt zu dem Schluss, dass LAM ein geeigneter Marker für die TB-Diagnose in allgemeinen Patient*Innenpopulationen ist und dass ein zukünftiger Test das Target Product Profile (TPP) der Weltgesundheitsorganisation erfüllen kann. Ein kritischer Erfolgsfaktor ist jedoch die Entwicklung eines einfachen, schnellen und hochsensitiven uLAM Point-of-Care Tests mit einem äußerst niedrigen Schwellenwert von etwa 5 pg/ml, um eine ausreichende Sensitivität zu erreichen. Selbst wenn ein Test die erforderliche Sensitivität von 65% nicht erreicht, legen Modellierungsstudien nahe, die Entwicklung von uLAM-Tests trotzdem zu priorisieren, da die Vorteile eines hochverfügbaren, urinbasierten Point-of-Care-Tests erheblich sein können und ein Test eine hohe Bevölkerungsabdeckung erreichen kann, sofern er kostengünstig ist.

Fazit: LAM ist einer der vielversprechendsten diagnostischen Biomarker für Tuberkulose, und es wurden signifikante Fortschritte bei der Aufzeigung seines Potenzials gemacht. Diese Fortschritte deuten darauf hin, dass der ultrasensitive LAM-Nachweis die Grundlage für den ersten nicht-Sputum-TB-Test bilden könnte, der das Target Product Profile der Weltgesundheitsorganisation erfüllt, mit dem Potenzial, die TB-Diagnose zu transformieren.

7 REFERENCES

Afonso-Barroso, A., Clark, S. O., Williams, A., Rosa, G. T., Nóbrega, C., Silva-Gomes, S., Vale-Costa, S., Ummels, R., Stoker, N., Movahedzadeh, F., van der Ley, P., Sloots, A., Cot, M., Appelmelk, B. J., Puzo, G., Nigou, J., Geurtsen, J. and Appelberg, R. (2013). **Lipoarabinomannan mannose caps do not affect mycobacterial virulence or the induction of protective immunity in experimental animal models of infection and have minimal impact on in vitro inflammatory responses**. Cell Microbiol *15*, 660–674, doi:10.1111/cmi.12065.

Aledort, J. E., Ronald, A., Rafael, M. E., Girosi, F., Vickerman, P., Le Blancq, S. M., Landay, A., Holmes, K., Ridzon, R., Hellmann, N., Shea, M. V. and Peeling, R. W. (2006). **Reducing the burden of sexually transmitted infections in resource-limited settings: the role of improved diagnostics.** Nature *444 Suppl*, 59–72, doi:10.1038/nature05447.

Amin, A. G., De, P., Spencer, J. S., Brennan, P. J., Daum, J., Andre, B. G., Joe, M., Bai, Y., Laurentius, L., Porter, M. D., Honnen, W. J., Choudhary, A., Lowary, T. L., Pinter, A. and Chatterjee, D. (2018). **Detection of lipoarabinomannan in urine and serum of HIV-positive and HIV-negative TB suspects using an improved capture-enzyme linked immuno absorbent assay and gas chromatography/mass spectrometry**. Tuberculosis *111*, 178–187, doi:10.1016/j.tube.2018.06.004.

Andama, A., Whitman, G. R., Crowder, R., Reza, T. F., Jaganath, D., Mulondo, J., Nalugwa, T. K., Semitala, F. C., Worodria, W., Cook, C., Wood, R. C., Weigel, K. M., Olson, A. M., Lohmiller Shaw, J., Kato-Maeda, M., Denkinger, C. M., Nahid, P., Cangelosi, G. A. and Cattamanchi, A. (2022). **Accuracy of Tongue Swab Testing Using Xpert MTB-RIF Ultra for Tuberculosis Diagnosis.** J Clin Microbiol *60*, e0042122, doi:10.1128/jcm.00421-22.

Angala, S. K., McNeil, M. R., Shi, L., Joe, M., Pham, H., Zuberogoitia, S., Nigou, J., Boot, C. M., Lowary, T. L., Gilleron, M. and Jackson, M. (2017). **Biosynthesis of the Methylthioxylose Capping Motif of Lipoarabinomannan in Mycobacterium tuberculosis**. ACS Chem Biol *12*, 682–691, doi:10.1021/acschembio.6b01071.

Bahr, N. C., Nuwagira, E., Evans, E. E., Cresswell, F. V., Bystrom, P. V., Byamukama, A., Bridge, S. C., Bangdiwala, A. S., Meya, D. B., Denkinger, C. M., Muzoora, C., Boulware, D. R., Williams, D. A., Taseera, K., Nyehangane, D., Ivan, M., Orikiriza, P., Rhein, J., Hullsiek, K. H., Musubire, A., Pastick, K., Nabeta, P., Mwesigye, J. and Rajasingham, R. (2018). **Diagnostic accuracy of Xpert MTB/RIF Ultra for tuberculous meningitis in HIV-infected adults: a prospective cohort study**. Lancet Infect Dis *18*, 68–75, doi:10.1016/S1473-3099(17)30474-7.

Barr, D. A., Lewis, J. M., Feasey, N., Schutz, C., Kerkhoff, A. D., Jacob, S. T., Andrews, B., Kelly, P., Lakhi, S., Muchemwa, L., Bacha, H. A., Hadad, D. J., Bedell, R., van Lettow, M., Zachariah, R., Crump, J. A., Alland, D., Corbett, E. L., Gopinath, K., Singh, S., Griesel, R., Maartens, G., Mendelson, M., Ward, A. M., Parry, C. M., Talbot, E. A., Munseri, P., Dorman, S. E., Martinson, N., Shah, M., Cain, K., Heilig, C. M., Varma, J. K., von Gottberg,

A., Sacks, L., Wilson, D., Squire, S. B., Lalloo, D. G., Davies, G. and Meintjes, G. (2020). **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 *20*, 742–752, doi:10.1016/S1473-3099(19)30695-4.

Bates, M., O'Grady, J., Maeurer, M., Tembo, J., Chilukutu, L., Chabala, C., Kasonde, R., Mulota, P., Mzyece, J., Chomba, M., Mukonda, L., Mumba, M., Kapata, N., Rachow, A., Clowes, P., Hoelscher, M., Mwaba, P. and Zumla, A. (2013). **Assessment of the Xpert MTB/RIF assay for diagnosis of tuberculosis with gastric lavage aspirates in children in sub-Saharan Africa: a prospective descriptive study**. Lancet Infect Dis *13*, 36–42, doi:10.1016/S1473-3099(12)70245-1.

Besra, G. S. and Brennan, P. J. (1997). **The Mycobacterial Cell Envelope**. Journal of Pharmacy and Pharmacology *49*, 25–30, doi:10.1111/j.2042-7158.1997.tb06146.x.

Bjerrum, S., Åhsberg, J., Szekely, R., Opintan, J., Lartey, M., Shah, M., Mitarai, S., Broger, T., Ruhwald, M. and Johansen, I. S. (2022). **Diagnostic Accuracy of Urine Lipoarabinomannan Testing in Early Morning Urine versus Spot Urine for Diagnosis of Tuberculosis among People with HIV**. Microbiol Spectr *10*, doi:10.1128/spectrum.00208-22.

Bjerrum, S., Schiller, I., Dendukuri, N., Kohli, M., Nathavitharana, R. R., Zwerling, A. A., Denkinger, C. M., Steingart, K. R. and Shah, M. (2019). **Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in people living with HIV**. Cochrane Database of Systematic Reviews *2019*, doi:10.1002/14651858.CD011420.pub3.

Bjerrum, S., Yang, B., Ahsberg, J., Olbrich, L., Weis Damkjaer, M., Nathavitharana, R., Broger, T., Olaru, I. D., Sweetser, B., Poore, H., Razid, A., Kay, A., Denkinger, C. M., Schiller, I., Dendukuri, N., Jaganath, D., Lundh, A. and Shah, M. (2024). **Parallel use of low-complexity automated nucleic acid amplification tests and lateral flow urine lipoarabinomannan assay to detect tuberculosis disease in adults and adolescents with HIV: a systematic review**. Cochrane Database of Systematic Reviews (in preparation).

Boehme, C. C., Nabeta, P., Hillemann, D., Nicol, M. P., Shenai, S., Krapp, F., Allen, J., Tahirli, R., Blakemore, R., Rustomjee, R., Milovic, A., Jones, M., O'Brien, S. M., Persing, D. H., Ruesch-Gerdes, S., Gotuzzo, E., Rodrigues, C., Alland, D. and Perkins, M. D. (2010). **Rapid molecular detection of tuberculosis and rifampin resistance.** N Engl J Med *363*, 1005–1015, doi:10.1056/NEJMoa0907847.

Bordet, J. and Gengou, O. (1910). **Sur l'existence de substances sensibilisatrices dans la plupart des sérums antimicrobiens.** Ann Inst Pasteur *15*, 289–302.

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J. G., Moher, D., Rennie, D., de Vet, H. C. W., Kressel, H. Y., Rifai, N., Golub, R. M., Altman, D. G., Hooft, L., Korevaar, D. A. and Cohen, J. F. (2015). **STARD 2015: an updated list of**

essential items for reporting diagnostic accuracy studies. BMJ *351*, 1446–1452, doi:10.1373/clinchem.2015.246280.

Boyles, T. H., Griesel, R., Stewart, A., Mendelson, M. and Maartens, G. (2018). **Incremental yield and cost of urine Determine TB-LAM and sputum induction in seriously ill adults with HIV**. International Journal of Infectious Diseases *75*, 67–73, doi:10.1016/j.ijid.2018.08.005.

Briken, V., Porcelli, S. a., Besra, G. S. and Kremer, L. (2004). **Mycobacterial lipoarabinomannan and related lipoglycans: From biogenesis to modulation of the immune response**. Mol Microbiol *53*, 391–403, doi:10.1111/j.1365-2958.2004.04183.x.

Brock, M., Hanlon, D., Zhao, M. and Pollock, N. R. (2020). **Detection of mycobacterial lipoarabinomannan in serum for diagnosis of active tuberculosis**. Diagn Microbiol Infect Dis *96*, doi:10.1016/j.diagmicrobio.2019.114937.

Broger, T., Basu Roy, R., Filomena, A., Greef, C. H., Rimmele, S., Havumaki, J., Danks, D., Schneiderhan-Marra, N., Gray, C. M., Singh, M., Rosenkrands, I., Andersen, P., Husar, G. M., Joos, T. O., Gennaro, M. L., Lochhead, M. J., Denkinger, C. M. and Perkins, M. D. (2017). **Diagnostic Performance of Tuberculosis-Specific IgG Antibody Profiles in Patients with Presumptive Tuberculosis from Two Continents**. Clinical Infectious Diseases *64*, 947–955, doi:10.1093/cid/cix023.

Broger, T., Koeppel, L., Huerga, H., Miller, P., Gupta-Wright, A., Blanc, F.-X., Esmail, A., Reeve, B. W. P., Floridia, M., Kerkhoff, A. D., Ciccacci, F., Kasaro, M. P., Thit, S. S., Bastard, M., Ferlazzo, G., Yoon, C., Van Hoving, D. J., Sossen, B., García, J. I., Cummings, M. J., Wake, R. M., Hanson, J., Cattamanchi, A., Meintjes, G., Maartens, G., Wood, R., Theron, G., Dheda, K., Olaru, I. D., Denkinger, C. M. (2023). **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 *11*, e903–e916, doi:10.1016/S2214-109X(23)00135-3.

Broger, T., Marx, F. M., Theron, G., Marais, B. J., Nicol, M. P., Kerkhoff, A. D., Nathavitharana, R., Huerga, H., Gupta-Wright, A., Kohli, M., Nichols, B. E., Muyoyeta, M., Meintjes, G., Ruhwald, M., Peeling, R. W., Pant Pai, N., Pollock, N. R., Pai, M., Cattamanchi, A., Dowdy, D. W., Dewan, P. and Denkinger, C. M. (2024). **Diagnostic yield as an important metric for the evaluation of novel tuberculosis tests – rationale and guidance for future research**. Lancet Glob Health (accepted).

Broger, T., Muyoyeta, M., AD, K., CM, D., Moreau, E., Kerkhoff, A. D., Denkinger, C. M. and Moreau, E. (2020a). **Tuberculosis test results using fresh versus biobanked urine samples with FujiLAM**. Lancet Infect Dis *20*, 22–23, doi:10.1016/S1473-3099(19)30684-X.

Broger, T., Nicol, M. P., Sigal, G. B., Gotuzzo, E., Zimmer, A. J., Surtie, S., Caceres-Nakiche, T., Mantsoki, A., Reipold, E. I., Székely, R., Tsionsky, M., van Heerden, J., Plisova, T., Chikamatsu, K., Lowary, T. L., Pinter, A., Mitarai, S., Moreau, E., Schumacher, S. G. and Denkinger, C. M. (2020b). **Diagnostic accuracy of 3 urine lipoarabinomannan**

tuberculosis assays in HIV-negative outpatients. Journal of Clinical Investigation *130*, 5756–5764, doi:10.1172/JCI140461.

Broger, T., Nicol, M. P., Székely, R., Bjerrum, S., Sossen, B., Schutz, C., Opintan, J. A., Johansen, I. S., Mitarai, S., Chikamatsu, K., Kerkhoff, A. D., Macé, A., Ongarello, S., Meintjes, G., Denkinger, C. M., Schumacher, S. G. (2020c). **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 *17*, e1003113, doi:10.1371/journal.pmed.1003113.

Broger, T., Sossen, B., du Toit, E., Kerkhoff, A. D., Schutz, C., Ivanova Reipold, E., Ward, A., Barr, D. A., Macé, A., Trollip, A., Burton, R., Ongarello, S., Pinter, A., Lowary, T. L., Boehme, C., Nicol, M. P., Meintjes, G., Denkinger, C. M. (2019a). **Novel lipoarabinomannan point-of-care tuberculosis test for people with HIV: a diagnostic accuracy study**. Lancet Infect Dis *19*, 852–861, doi:10.1016/S1473-3099(19)30001-5.

Broger, T., Tsionksy, M., Mathew, A., Lowary, T. L., Pinter, A., Plisova, T., Bartlett, D., Barbero, S., Denkinger, C. M., Moreau, E., Katsuragi, K., Kawasaki, M., Nahid, P. and Sigal, G. B. (2019b). **Sensitive electrochemiluminescence (ECL) immunoassays for detecting lipoarabinomannan (LAM) and ESAT-6 in urine and serum from tuberculosis patients**. PLoS One *14*, e0215443, doi:10.1371/journal.pone.0215443.

Brümmer, L. E., Katzenschlager, S., McGrath, S., Schmitz, S., Gaeddert, M., Erdmann, C., Bota, M., Grilli, M., Larmann, J., Weigand, M. A., Pollock, N. R., Macé, A., Erkosar, B., Carmona, S., Sacks, J. A., Ongarello, S. and Denkinger, C. M. (2022). **Accuracy of rapid point-of-care antigen-based diagnostics for SARS-CoV-2: An updated systematic review and meta-analysis with meta-regression analyzing influencing factors**. PLoS Med *19*, e1004011, doi:10.1371/journal.pmed.1004011.

Bulterys, M. A., Wagner, B., Redard-Jacot, M. M., Suresh, A., Pollock, N. R., Moreau, E., Denkinger, C. M., Drain, P. K. and Broger, T. (2019). **Point-Of-Care Urine LAM Tests for Tuberculosis Diagnosis: A Status Update**. J Clin Med *9*, 111, doi:10.3390/jcm9010111.

Cantera, J. L., Lillis, L. M., Peck, R. B., Moreau, E., Schouten, J. A., Davis, P., Drain, P. K., Andama, A., Pinter, A., Kawasaki, M., Källenius, G., Sundling, C., Dobos, K. M., Flores, D., Chatterjee, D., Murphy, E., Halas, O. R. and Boyle, D. S. (2022). **Performance of novel antibodies for lipoarabinomannan to develop diagnostic tests for Mycobacterium tuberculosis**. PLoS One *17*, 1–17, doi:10.1371/journal.pone.0274415.

Cantera, J. L., Rashid, A. A., Lillis, L. M., Peck, R. B., Drain, P. K., Shapiro, A. E., Wilson, D. P. K., Pinter, A., Kawasaki, M., Moreau, E. and Boyle, D. S. (2023). **Isolation and purification of lipoarabinomannan from urine of adults with active TB.** Int J Tuberc Lung Dis *27*, 75–77, doi:10.5588/ijtld.22.0372.

Choudhary, A., Patel, D., Honnen, W., Lai, Z., Prattipati, R. S., Zheng, R. B., Hsueh, Y.-C., Gennaro, M. L., Lardizabal, A., Restrepo, B. I., Garcia-Viveros, M., Joe, M., Bai, Y., Shen, K., Sahloul, K., Spencer, J. S., Chatterjee, D., Broger, T., Lowary, T. L. and Pinter, A. (2018). **Characterization of the Antigenic Heterogeneity of Lipoarabinomannan, the Major**

Surface Glycolipid of Mycobacterium tuberculosis , and Complexity of Antibody Specificities toward This Antigen. The Journal of Immunology *200*, 3053–3066, doi:10.4049/jimmunol.1701673.

Church, E. C., Steingart, K. R., Cangelosi, G. A., Ruhwald, M., Kohli, M. and Shapiro, A. E. (2024). **Oral swabs with a rapid molecular diagnostic test for pulmonary tuberculosis in adults and children: a systematic review**. Lancet Glob Health *12*, e45–e54, doi:10.1016/S2214-109X(23)00469-2.

Comella-del-Barrio, P., Molina-Moya, B., Gautier, J., Villar-Hernández, R., Doresca, M. J. C., Sallés-Mingels, B., Canales-Aliaga, L., Narcisse, M., Pérez-Porcuna, T. M., Creswell, J., Cuevas, L. E. and Domínguez, J. (2021). **Diagnostic Performance of the Fujifilm SILVAMP TB-LAM in Children with Presumptive Tuberculosis**. J Clin Med *10*, 1914, doi:10.3390/jcm10091914.

Connelly, J. T., Grant, B., Munsamy, V., Pym, A. and Somoskovi, A. (2019). **Lipoarabinomannan point-of-care tests: evaluation with fresh samples needed**. Lancet Infect Dis *19*, 1053, doi:10.1016/S1473-3099(19)30475-X.

Corral, J. E., Mareth, K. F., Riegert-Johnson, D. L., Das, A. and Wallace, M. B. (2019). **Diagnostic Yield From Screening Asymptomatic Individuals at High Risk for Pancreatic Cancer: A Meta-analysis of Cohort Studies**. Clinical Gastroenterology and Hepatology *17*, 41–53, doi:10.1016/j.cgh.2018.04.065.

Correia-Neves, M., Fröberg, G., Korshun, L., Viegas, S., Vaz, P., Ramanlal, N., Bruchfeld, J., Hamasur, B., Brennan, P., Källenius, G., Froberg, G., Korshun, L., Viegas, S., Vaz, P., Ramanlal, N., Bruchfeld, J., Hamasur, B., Brennan, P. and Kallenius, G. (2019). **Biomarkers for tuberculosis: the case for lipoarabinomannan**. ERJ Open Res *5*, 00115–02018, doi:10.1183/23120541.00115-2018.

Corrigan, D. T., Ishida, E., Chatterjee, D., Lowary, T. L. and Achkar, J. M. (2022). **Monoclonal antibodies to lipoarabinomannan/arabinomannan – characteristics and implications for tuberculosis research and diagnostics**. Trends Microbiol *31*, 22–35, doi:10.1016/j.tim.2022.07.001.

Cox, J. a., Lukande, R. L., Kalungi, S., Van Marck, E., Van de Vijver, K., Kambugu, A., Nelson, A. M., Colebunders, R., Manabe, Y. C., de Vijver, K., Kambugu, A., Nelson, A. M., Colebunders, R. and Manabe, Y. C. (2015). **Is Urinary Lipoarabinomannan the Result of Renal Tuberculosis? Assessment of the Renal Histology in an Autopsy Cohort of Ugandan HIV-Infected Adults**. PLoS One *10*, e0123323, doi:10.1371/journal.pone.0123323.

Craig, C. F. (1910). **Observations upon the noguchi modification of the wassermann complement fixation test in the diagnosis of lues in the military service**. Journal of Experimental Medicine *12*, 726–745, doi:10.1084/jem.12.6.726.

Cunningham, J., Jones, S., Gatton, M. L., Barnwell, J. W., Cheng, Q., Chiodini, P. L., Glenn, J., Incardona, S., Kosack, C., Luchavez, J., Menard, D., Nhem, S., Oyibo, W., Rees-Channer, R. R., Gonzalez, I. and Bell, D. (2019). **A review of the WHO malaria rapid diagnostic test**

product testing programme (2008–2018): performance, procurement and policy. Malar J *18*, 387, doi:10.1186/s12936-019-3028-z.

Das, S., Jang, I. K., Barney, B., Peck, R., Rek, J. C., Arinaitwe, E., Adrama, H., Murphy, M., Imwong, M., Ling, C. L., Proux, S., Haohankhunnatham, W., Rist, M., Seilie, A. M., Hanron, A., Daza, G., Chang, M., Nakamura, T., Kalnoky, M., Labarre, P., Murphy, S. C., McCarthy, J. S., Nosten, F., Greenhouse, B., Allauzen, S. and Domingo, G. J. (2017). **Performance of a High-Sensitivity Rapid Diagnostic Test for Plasmodium falciparum Malaria in Asymptomatic Individuals from Uganda and Myanmar and Naive Human Challenge Infections.** Am J Trop Med Hyg *97*, 1540–1550, doi:10.4269/ajtmh.17-0245.

De, P., Amin, A. G., Flores, D., Simpson, A., Dobos, K. and Chatterjee, D. (2021). **Structural implications of lipoarabinomannan glycans from global clinical isolates in diagnosis of Mycobacterium tuberculosis infection**. Journal of Biological Chemistry *297*, doi:10.1016/j.jbc.2021.101265.

Denkinger, C. M., Kik, S. V., Cirillo, D. M., Casenghi, M., Shinnick, T., Weyer, K., Gilpin, C., Boehme, C. C., Schito, M., Kimerling, M. and Pai, M. (2015). **Defining the Needs for Next Generation Assays for Tuberculosis**. J Infect Dis *211*, S29–S38, doi:10.1093/infdis/jiu821.

Denkinger, C. M., Schumacher, S. G., Gilpin, C., Korobitsyn, A., Wells, W. A., Pai, M., Leeflang, M., Steingart, K. R., Bulterys, M., Schünemann, H., Glaziou, P. and Weyer, K. (2019). **Guidance for the Evaluation of Tuberculosis Diagnostics That Meet the World Health Organization (WHO) Target Product Profiles: An Introduction to WHO Process and Study Design Principles**. J Infect Dis *220*, S91–S98, doi:10.1093/infdis/jiz097.

Dhana, A., Hamada, Y., Kengne, A. P., Kerkhoff, A. D., Broger, T., Denkinger, C. M., Rangaka, M. X., Gupta-Wright, A., Fielding, K., Wood, R., Huerga, H., Rücker, S. C. M., Bjerrum, S., Johansen, I. S., Thit, S. S., Kyi, M. M., Hanson, J., Barr, D. A., Meintjes, G. and Maartens, G. (2022). **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**. Journal of Infection *85*, 40–48, doi:10.1016/j.jinf.2022.05.010.

Dorman, S. E., Schumacher, S. G., Alland, D., Nabeta, P., Armstrong, D. T., King, B., Hall, S. L., Chakravorty, S., Cirillo, D. M., Tukvadze, N., Bablishvili, N., Stevens, W., Scott, L., Rodrigues, C., Kazi, M. I., Joloba, M., Nakiyingi, L., Nicol, M. P., Ghebrekristos, Y., Anyango, I., Murithi, W., Dietze, R., Lyrio Peres, R., Skrahina, A., Auchynka, V., Chopra, K. K., Hanif, M., Liu, X., Yuan, X., Boehme, C. C., Ellner, J. J., Denkinger, C. M. (2018). **Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis and rifampicin resistance: a prospective multicentre diagnostic accuracy study.** Lancet Infect Dis *18*, 76–84, doi:10.1016/S1473-3099(17)30691-6.

Drain, P. K., Gardiner, J. L., Hannah, H., Broger, T., Dheda, K., Fielding, K., Walzl, G., Kaforou, M., Kranzer, K., Joosten, S. A., Gilpin, C., Wayer, K., Denkinger, C. M. and Schumacher, S. G. (2019). **Guidance for Studies Evaluating the Accuracy of Biomarker-Based**

Nonsputum Tests to Diagnose Tuberculosis. J Infect Dis *220*, S108–S115, doi:10.1093/infdis/jiz356.

Fekadu, G., Wang, Y. and You, J. H. S. (2023). **Standard diagnostics with and without urine-based lipoarabinomannan testing for tuberculosis disease in HIV-infected patients in a high-burden setting-A cost-effectiveness analysis.** PLoS One *18*, e0288605, doi:10.1371/journal.pone.0288605.

Fraaij, P. L. A., Geelhoed-Mieras, M. M., van Baalen, C. A., Tiddens, H. A. W. M., van Rossum, A. M. C., van der Klis, F. R., Osterhaus, A. D. M. E. and Rimmelzwaan, G. F. (2011). **Structural and enzymatic studies on a novel substituent of Mycobacterium tuberculosis lipoarabinomannan**. J Virol *85*, 11995–12000, doi:10.1128/JVI.05213-11.

Fratti, R. a., Backer, J. M., Gruenberg, J., Corvera, S. and Deretic, V. (2001). **Role of phosphatidylinositol 3-kinase and Rab5 effectors in phagosomal biogenesis and mycobacterial phagosome maturation arrest**. Journal of Cell Biology *154*, 631–644, doi:10.1083/jcb.200106049.

Garg, A., Barnes, P. F., Roy, S., Quiroga, M. F., Wu, S., García, V. E., Krutzik, S. R., Weis, S. E. and Vankayalapati, R. (2008). **Mannose-capped lipoarabinomannan- and prostaglandin E2-dependent expansion of regulatory T cells in humanMycobacterium tuberculosisinfection**. Eur J Immunol *38*, 459–469, doi:10.1002/eji.200737268.

Geijtenbeek, T. B. H., van Vliet, S. J., Koppel, E. a, Sanchez-Hernandez, M., Vandenbroucke-Grauls, C. M. J. E., Appelmelk, B. and van Kooyk, Y. (2002). **Mycobacteria Target DC-SIGN to Suppress Dendritic Cell Function**. Journal of Experimental Medicine *197*, 7–17, doi:10.1084/jem.20021229.

Giguère, K., Eaton, J. W., Marsh, K., Johnson, L. F., Johnson, C. C., Ehui, E., Jahn, A., Wanyeki, I., Mbofana, F., Bakiono, F., Mahy, M. and Maheu-Giroux, M. (2021). **Trends in knowledge of HIV status and efficiency of HIV testing services in sub-Saharan Africa, 2000–20: a modelling study using survey and HIV testing programme data**. Lancet HIV *8*, e284–e293, doi:10.1016/S2352-3018(20)30315-5.

Grant, A. D., Charalambous, S., Tlali, M., Karat, A. S., Dorman, S. E., Hoffmann, C. J., Johnson, S., Vassall, A., Churchyard, G. J. and Fielding, K. L. (2020). **Algorithm-guided empirical tuberculosis treatment for people with advanced HIV (TB Fast Track): an open-label, cluster-randomised trial**. LANCET HIV *7*, E27–E37, doi:10.1016/S2352-3018(19)30266-8.

Guerardel, Y., Maes, E., Elass, E., Leroy, Y., Timmerman, P., Besra, G. S., Locht, C., Strecker, G. and Kremer, L. (2002). **Structural Study of Lipomannan and Lipoarabinomannan from Mycobacterium chelonae.** Journal of Biological Chemistry *277*, 30635–30648, doi:10.1074/jbc.M204398200.

Gupta, R. K., Lucas, S. B., Fielding, K. L. and Lawn, S. D. (2015). **Prevalence of tuberculosis in post-mortem studies of HIV-infected adults and children in resource-limited settings**. AIDS *29*, 1987–2002, doi:10.1097/QAD.0000000000000802.

Gupta-Wright, A., Corbett, E. L., van Oosterhout, J. J., Wilson, D., Grint, D., Alufandika-Moyo, M., Peters, J. A., Chiume, L., Flach, C., Lawn, S. D. and Fielding, K. (2018). **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**. The Lancet *392*, 292–301, doi:10.1016/S0140-6736(18)31267-4.

Gupta-Wright, A., Peters, J. A., Flach, C. and Lawn, S. D. (2016). **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 *14*, 53, doi:10.1186/s12916-016-0603-9.

Hamasur, B., Bruchfeld, J., Haile, M., Pawlowski, A., Bjorvan, B., Kallenius, G., Svenson, S. B., Bjorvatn, B., Källenius, G. and Svenson, S. B. (2001). **Rapid diagnosis of tuberculosis by detection of mycobacterial lipoarabinomannan in urine**. J Microbiol Methods *45*, 41–52, doi:10.1016/S0167-7012(01)00239-1.

Haraka, F., Kakolwa, M., Schumacher, S. G., Nathavitharana, R. R., Denkinger, C. M., Gagneux, S., Reither, K. and Ross, A. (2021). **Impact of the diagnostic test Xpert MTB/RIF on patient outcomes for tuberculosis.** Cochrane Database Syst Rev *5*, CD012972, doi:10.1002/14651858.CD012972.pub2.

Hatzold, K., Gudukeya, S., Mutseta, M. N., Chilongosi, R., Nalubamba, M., Nkhoma, C., Munkombwe, H., Munjoma, M., Mkandawire, P., Mabhunu, V., Smith, G., Madidi, N., Ahmed, H., Kambeu, T., Stankard, P., Johnson, C. C. and Corbett, E. L. (2019). **HIV self-testing: breaking the barriers to uptake of testing among men and adolescents in sub-Saharan Africa, experiences from STAR demonstration projects in Malawi, Zambia and Zimbabwe.** J Int AIDS Soc *22 Suppl 1*, e25244, doi:10.1002/jia2.25244.

Herrmann, Y., Lainati, F., Castro, M. del M., Mwamba, C. P., Kumwenda, M., Muyoyeta, M., Broger, T., Heinrich, N., Olbrich, L., Corbett, E. L., McMahon, S. A., Engel, N. and Denkinger, C. M. (2022). **User perspectives and preferences on a novel TB LAM diagnostic (Fujifilm SILVAMP TB LAM)–a qualitative study in Malawi and Zambia**. PLOS Global Public Health *2*, e0000672, doi:10.1371/journal.pgph.0000672.

Hoa, N. B., Fajans, M., Van, H. N., Ngoc, B. V., Viet, N. N., Thi, H. N., Huong, L. T. T., Minh, D. T., Kim, C. N., Tuyet, T. H. T., Huu, T. N., Ngoc, D. B., Viet, H. N., Khanh, A. T., Lillis, L., Perez, M., Thomas, K. K., Peck, R. B., Cantera, J. L., Murphy, E., Halas, O. R., Storey, H. L., Pinter, A., Ruhwald, M., Drain, P. K. and Boyle, D. S. (2023). **Urine lipoarabinomannan concentrations among HIV-uninfected adults with pulmonary or extrapulmonary tuberculosis disease in Vietnam**. medrxiv *7*, doi:10.1101/2023.07.17.23292752.

Hsu, J., Brozek, J. L., Terracciano, L., Kreis, J., Compalati, E., Stein, A. T., Fiocchi, A. and Schünemann, H. J. (2011). **Application of GRADE: Making evidence-based recommendations about diagnostic tests in clinical practice guidelines**. Implementation Science *6*, doi:10.1186/1748-5908-6-62.

Huang, L., Niu, Y., Zhang, L., Yang, R. and Wu, M. (2023). **Diagnostic value of chemiluminescence for urinary lipoarabinomannan antigen assay in active tuberculosis: insights from a retrospective study**. Front Cell Infect Microbiol *13*, doi:10.3389/fcimb.2023.1291974.

Huang, Z., Huang, H., Hu, J., Xia, L., Liu, X., Qu, R., Huang, X., Yang, Y., Wu, K., Ma, R., Xu, J., Chen, Z., Wu, Y., Yang, J., Fang, Y., Zeng, J., Lai, W., Sui, G., Sha, W., Xiong, Y., Lu, S. and Fan, X.-Y. (2024). **A novel quantitative urine LAM antigen strip for point-of-care tuberculosis diagnosis in non-HIV adults**. Journal of Infection *88*, 194–198, doi:10.1016/j.jinf.2023.11.014.

Huerga, H., Bastard, M., Lubega, A. V., Akinyi, M., Antabak, N. T., Ohler, L., Muyindike, W., Taremwa, I. M., Stewart, R., Bossard, C., Nkosi, N., Ndlovu, Z., Hewison, C., Stavia, T., Okomo, G., Ogoro, J. O., Ngozo, J., Mbatha, M., Aleny, C., Wanjala, S., Musoke, M., Atwine, D., Ascorra, A., Ardizzoni, E., Casenghi, M., Ferlazzo, G., Nakiyingi, L., Gupta-Wright, A. and Bonnet, M. (2023). **Novel FujiLAM assay to detect tuberculosis in HIV-positive ambulatory patients in four African countries: a diagnostic accuracy study**. Lancet Glob Health *11*, e126–e135, doi:10.1016/S2214-109X(22)00463-6.

Huerga, H., Ferlazzo, G., Bevilacqua, P., Kirubi, B., Ardizzoni, E., Wanjala, S., Sitienei, J. and Bonnet, M. (2017). **Incremental Yield of Including Determine-TB LAM Assay in Diagnostic Algorithms for Hospitalized and Ambulatory HIV-Positive Patients in Kenya**. PLoS One *12*, e0170976, doi:10.1371/journal.pone.0170976.

Ishida, E., Corrigan, D. T., Malonis, R. J., Hofmann, D., Chen, T., Amin, A. G., Chatterjee, D., Joe, M., Lowary, T. L., Lai, J. R. and Achkar, J. M. (2021). **Monoclonal antibodies from humans with Mycobacterium tuberculosis exposure or latent infection recognize distinct arabinomannan epitopes**. Commun Biol *4*, 1181, doi:10.1038/s42003-021-02714-w.

Ismail, N., Nathanson, C.-M., Zignol, M. and Kasaeva, T. (2023). **Achieving universal access to rapid tuberculosis diagnostics**. BMJ Glob Health *8*, e012666, doi:10.1136/bmjgh-2023-012666.

Jannati, A., Sadeghi, V., Imani, A. and Saadati, M. (2018). **Effective coverage as a new approach to health system performance assessment: a scoping review**. BMC Health Serv Res *18*, 886, doi:10.1186/s12913-018-3692-7.

Joe, M., Sun, D., Taha, H., Completo, G. C., Croudace, J. E., Lammas, D. A., Besra, G. S. and Lowary, T. L. (2006). **The 5-deoxy-5-methylthio-xylofuranose residue in mycobacterial lipoarabinomannan. Absolute stereochemistry, linkage position, conformation, and immunomodulatory activity**. J Am Chem Soc *128*, 5059–5072, doi:10.1021/ja057373q.

Kahn, R. L. (1942). **Serology in Syphilis Control: Principles of Sensitivity and Specificity. With an Appendix for Health Officers and Industrial Physicians**. JAMA: The Journal of the American Medical Association *120*, 724, doi:10.1001/jama.1942.02830440066030.

Källenius, G., Correia-Neves, M., Buteme, H., Hamasur, B. and Svenson, S. B. (2015). **Lipoarabinomannan, and its related glycolipids, induce divergent and opposing**

immune responses to Mycobacterium tuberculosis depending on structural diversity and experimental variations. Tuberculosis (Edinb), doi:10.1016/j.tube.2015.09.005.

Kawasaki, M., Echiverri, C., Raymond, L., Cadena, E., Reside, E., Gler, M. T., Oda, T., Ito, R., Higashiyama, R., Katsuragi, K. and Liu, Y. (2019). **Lipoarabinomannan in sputum to detect bacterial load and treatment response in patients with pulmonary tuberculosis: Analytic validation and evaluation in two cohorts**. PLoS Med *16*, e1002780, doi:10.1371/journal.pmed.1002780.

Kerkhoff, A. D., Sossen, B., Schutz, C., Reipold, E. I., Trollip, A., Moreau, E., Schumacher, S. G., Burton, R., Ward, A., Nicol, M. P., Meintjes, G., Denkinger, C. M. and Broger, T. (2020). **Diagnostic sensitivity of SILVAMP TB-LAM (FujiLAM) point-of-care urine assay for extra-pulmonary tuberculosis in people living with HIV**. European Respiratory Journal *55*, 1901259, doi:10.1183/13993003.01259-2019.

Kik, S. V., Denkinger, C. M., Chedore, P. and Pai, M. (2014). **Replacing smear microscopy for the diagnosis of tuberculosis: what is the market potential?** European Respiratory Journal *43*, 1793–1796, doi:10.1183/09031936.00217313.

Knutson, K. L., Hmama, Z., Herrera-Velit, P., Rochford, R. and Reiner, N. E. (1998). **Lipoarabinomannan of Mycobacterium tuberculosis Promotes Protein Tyrosine Dephosphorylation and Inhibition of Mitogen-activated Protein Kinase in Human Mononuclear Phagocytes**. Journal of Biological Chemistry *273*, 645–652, doi:10.1074/jbc.273.1.645.

Kraef, C., Yedilbayev, A., Seguy, N., Bentzon, A., Podlekareva, D., Cirillo, D. M., van der Werf, M. J. and Kirk, O. (2022). **Uptake of the lateral flow urine LAM test in Europe and Central Asia**. The International Journal of Tuberculosis and Lung Disease *26*, 835–841, doi:10.5588/ijtld.21.0656.

Kroidl, I., Clowes, P., Reither, K., Mtafya, B., Rojas-Ponce, G., Ntinginya, E. N., Kalomo, M., Minja, L. T., Kowuor, D., Saathoff, E., Kroidl, A., Heinrich, N., Maboko, L., Bates, M., O'Grady, J., Zumla, A., Hoelscher, M. and Rachow, A. (2015). **Performance of urine lipoarabinomannan assays for paediatric tuberculosis in Tanzania**. European Respiratory Journal *46*, 761–770, doi:10.1183/09031936.00003315.

Krumholz, R. A. (1966). **Percutaneous Needle Biopsy of the Lung**. JAMA *195*, 38, doi:10.1001/jama.1966.03100010094030.

LaCourse, S. M., Pavlinac, P. B., Cranmer, L. M., Njuguna, I. N., Mugo, C., Gatimu, J., Stern, J., Walson, J. L., Maleche-Obimbo, E., Oyugi, J., Wamalwa, D. and John-Stewart, G. (2018). **Stool Xpert MTB/RIF and urine lipoarabinomannan for the diagnosis of tuberculosis in hospitalized HIV-infected children**. AIDS *32*, 69–78, doi:10.1097/QAD.0000000000001662.

Land, K. J., Boeras, D. I., Chen, X.-S., Ramsay, A. R. and Peeling, R. W. (2018). **REASSURED diagnostics to inform disease control strategies, strengthen health systems and improve patient outcomes**. Nat Microbiol *4*, 46–54, doi:10.1038/s41564-018-0295-3.

Lawn, S. D. (2012). **Point-of-care detection of lipoarabinomannan (LAM) in urine for diagnosis of HIV-associated tuberculosis: a state of the art review.** BMC Infect Dis *12*, 103, doi:10.1186/1471-2334-12-103.

Lawn, S. D., Edwards, D. J., Kranzer, K., Vogt, M., Bekker, L.-G. G. and Wood, R. (2009). **Urine lipoarabinomannan assay for tuberculosis screening before antiretroviral therapy diagnostic yield and association with immune reconstitution disease**. AIDS *23*, 1875–1880, doi:10.1097/QAD.0b013e32832e05c8.

Lawn, S. D. and Gupta-Wright, A. (2016). **Detection of lipoarabinomannan (LAM) in urine is indicative of disseminated TB with renal involvement in patients living with HIV and advanced immunodeficiency: evidence and implications.** Trans R Soc Trop Med Hyg *110*, 180–185, doi:10.1093/trstmh/trw008.

Lawn, S. D., Kerkhoff, A. D., Burton, R. and Meintjes, G. (2014). **Underestimation of the incremental diagnostic yield of HIV-associated tuberculosis in studies of the Determine TB-LAM Ag urine assay**. AIDS *28*, 1846–1848, doi:10.1097/QAD.0000000000000305.

Lawn, S. D., Kerkhoff, A. D., Vogt, M. and Wood, R. (2012a). **High Diagnostic Yield of Tuberculosis From Screening Urine Samples From HIV-Infected Patients With Advanced Immunodeficiency Using The Xpert MTB/RIF Assay**. JAIDS Journal of Acquired Immune Deficiency Syndromes *60*, 289–294, doi:10.1097/QAI.0b013e318258c6af.

Lawn, S. D., Kerkhoff, A. D., Vogt, M. and Wood, R. (2012b). **High Diagnostic Yield of Tuberculosis From Screening Urine Samples From HIV-Infected Patients With Advanced Immunodeficiency Using The Xpert MTB/RIF Assay**. JAIDS Journal of Acquired Immune Deficiency Syndromes *60*, 289–294, doi:10.1097/QAI.0b013e318258c6af.

Lindner, A. K., Nikolai, O., Kausch, F., Wintel, M., Hommes, F., Gertler, M., Krüger, L. J., Gaeddert, M., Tobian, F., Lainati, F., Köppel, L., Seybold, J., Corman, V. M., Drosten, C., Hofmann, J., Sacks, J. A., Mockenhaupt, F. P. and Denkinger, C. M. (2021a). **Head-to-head comparison of SARS-CoV-2 antigen-detecting rapid test with self-collected nasal swab versus professional-collected nasopharyngeal swab**. European Respiratory Journal *57*, 2003961, doi:10.1183/13993003.03961-2020.

Lindner, A. K., Nikolai, O., Rohardt, C., Kausch, F., Wintel, M., Gertler, M., Burock, S., Hörig, M., Bernhard, J., Tobian, F., Gaeddert, M., Lainati, F., Corman, V. M., Jones, T. C., Sacks, J. A., Seybold, J., Denkinger, C. M. and Mockenhaupt, F. P. (2021b). **Diagnostic accuracy and feasibility of patient self-testing with a SARS-CoV-2 antigen-detecting rapid test**. Journal of Clinical Virology *141*, 104874, doi:10.1016/j.jcv.2021.104874.

Lowary, T. L. and Achkar, J. M. (2022). **Tailor made: New insights into lipoarabinomannan structure may improve TB diagnosis**. Journal of Biological Chemistry *298*, 101678, doi:10.1016/j.jbc.2022.101678.

Mabey, D., Peeling, R. W., Ustianowski, A. and Perkins, M. D. (2004). **Diagnostics for the developing world.** Nat Rev Microbiol *2*, 231–240, doi:10.1038/nrmicro841.

MacLean, E., Broger, T., Yerlikaya, S., Fernandez-Carballo, B. L., Pai, M. and Denkinger, C. M. (2019). **A systematic review of biomarkers to detect active tuberculosis**. Nat Microbiol *4*, 748–758, doi:10.1038/s41564-019-0380-2.

Magni, R., Rruga, F., Alsaab, F. M., Sharif, S., Howard, M., Espina, V., Kim, B., Lepene, B., Lee, G., Alayouni, M. A., Steinberg, H., Araujo, R., Kashanchi, F., Riccardi, F., Morreira, S., Araujo, A., Poli, F., Jaganath, D., Semitala, F. C., Worodria, W., Andama, A., Choudhary, A., Honnen, W. J., Petricoin, E. F., Cattamanchi, A., Colombatti, R., de Waard, J. H., Oberhelman, R., Pinter, A., Gilman, R. H., Liotta, L. A. and Luchini, A. (2020). **Lipoarabinomannan antigenic epitope differences in tuberculosis disease subtypes.** Sci Rep *10*, 13944, doi:10.1038/s41598-020-70669-9.

Maruvada, P., Dmitrieva, N. I., East-Palmer, J. and Yen, P. M. (2004). **Cell cycle-dependent expression of thyroid hormone receptor-beta is a mechanism for variable hormone sensitivity.** Mol Biol Cell *15*, 1895–1903, doi:10.1091/mbc.E03.

Mase, S. R., Ramsay, A., Ng, V., Henry, M., Hopewell, P. C., Cunningham, J., Urbanczik, R., Perkins, M. D., Aziz, M. A. and Pai, M. (2007). **Yield of serial sputum specimen examinations in the diagnosis of pulmonary tuberculosis: A systematic review**. International Journal of Tuberculosis and Lung Disease *11*, 485–495.

Meggi, B., Bollinger, T., Mabunda, N., Vubil, A., Tobaiwa, O., Quevedo, J. I., Loquiha, O., Vojnov, L., Peter, T. F. and Jani, I. V. (2017). **Point-Of-Care p24 Infant Testing for HIV May Increase Patient Identification despite Low Sensitivity**. PLoS One *12*, e0169497, doi:10.1371/journal.pone.0169497.

Mina, M. J., Parker, R. and Larremore, D. B. (2020). **Rethinking Covid-19 Test Sensitivity — A Strategy for Containment**. New England Journal of Medicine *383*, e120, doi:10.1056/NEJMp2025631.

Mitamura, K., Shimizu, H., Yamazaki, M., Ichikawa, M., Nagai, K., Katada, J., Wada, A., Kawakami, C. and Sugaya, N. (2013). **Clinical evaluation of highly sensitive silver amplification immunochromatography systems for rapid diagnosis of influenza**. J Virol Methods *194*, 123–128, doi:10.1016/j.jviromet.2013.08.018.

Murray, C. J. and Evans, D. B. (2003). **Health systems performance assessment - debates, methods and empiricism**, URL: https://www.who.int/publications/i/item/9241562455 [Accessed May 2024].

Mwaura, M. and Engel, N. (2021). **Constructing confidence: User perspectives on AlereLAM testing for tuberculosis.** Int J Infect Dis *112*, 237–242, doi:10.1016/j.ijid.2021.09.018.

Nakiyingi, L., Nonyane, B. A. S., Ssengooba, W., Kirenga, B. J., Nakanjako, D., Lubega, G., Byakika-Kibwika, P., Joloba, M. L., Ellner, J. J., Dorman, S. E., Mayanja-Kizza, H. and Manabe, Y. C. (2015). **Predictors for MTB Culture-Positivity among HIV-Infected Smear-Negative Presumptive Tuberculosis Patients in Uganda: Application of New Tuberculosis Diagnostic Technology**. PLoS One *10*, e0133756, doi:10.1371/journal.pone.0133756.

Nathavitharana, R. R., Lederer, P., Chaplin, M., Bjerrum, S., Steingart, K. R. and Shah, M. (2021). **Impact of diagnostic strategies for tuberculosis using lateral flow urine lipoarabinomannan assay in people living with HIV**. Cochrane Database of Systematic Reviews *2021*, doi:10.1002/14651858.CD014641.

NDTV (2024). **"A Very Big Source Of Covid Vaccines For World": Bill Gates Praises India.** 29 February, URL: https://www.ndtv.com/india-news/bill-gates-in-india-a-very-big-source-of-covid-vaccines-for-world-bill-gates-praises-india-5146720 [Accessed May 2024].

Newcombe, R. G. (1998). **Two-sided confidence intervals for the single proportion: comparison of seven methods.** Stat Med *17*, 857–872, doi:10.1002/(SICI)1097-0258(19980430)17.

Ng, M., Fullman, N., Dieleman, J. L., Flaxman, A. D., Murray, C. J. L. and Lim, S. S. (2014). **Effective Coverage: A Metric for Monitoring Universal Health Coverage**. PLoS Med *11*, doi:10.1371/journal.pmed.1001730.

Nicol, M. P., Allen, V., Workman, L., Isaacs, W., Munro, J., Pienaar, S., Black, F., Adonis, L., Zemanay, W., Ghebrekristos, Y. and Zar, H. J. (2014). **Urine lipoarabinomannan testing for diagnosis of pulmonary tuberculosis in children: a prospective study**. Lancet Glob Health *2*, e278–e284, doi:10.1016/S2214-109X(14)70195-0.

Nicol, M. P., Schumacher, S. G., Workman, L., Broger, T., Baard, C., Prins, M., Bateman, L., du Toit, E., van Heerden, J., Szekely, R., Zar, H. J. and Denkinger, C. M. (2020). **Accuracy of a novel urine test, Fujifilm SILVAMP TB LAM, for the diagnosis of pulmonary tuberculosis in children**. Clinical Infectious Diseases *26*, 1–15, doi:10.1093/cid/ciaa1052.

Nicol, M. P. and Zar, H. J. (2011). **New specimens and laboratory diagnostics for childhood pulmonary TB: progress and prospects**. Paediatr Respir Rev *12*, 16–21, doi:10.1016/j.prrv.2010.09.008.

Nigou, J., Gilleron, M. and Puzo, G. (2003). **Lipoarabinomannans: from structure to biosynthesis**. Biochimie *85*, 153–166, doi:10.1016/S0300-9084(03)00048-8.

Nkereuwem, E., Togun, T., MP, G., Székely, R., Macé, A., Jobe, D., SG, S., Kampmann, B. and CM, D. (2020). **Comparing accuracy of lipoarabinomannan urine tests for diagnosis of pulmonary tuberculosis in children from four African countries: a cross-sectional study.** Lancet Infect Dis, doi:10.1016/S1473-3099(20)30598-3.

Ockhuisen, T., de Nooy, A., Jenkins, H. E., Han, A., Russell, C. A., Khan, S., Girdwood, S., Ruhwald, M., Kohli, M. and Nichols, B. E. (2024). **Cost-effectiveness of diagnostic tools and strategies for the screening and diagnosis of tuberculosis disease and infection: a scoping review**. BMJ Public Health *2*, e000276, doi:10.1136/bmjph-2023-000276.

Olbrich, L., Khambati, N., Bijker, E. M., Ruhwald, M., Heinrich, N. and Song, R. (2022). **FujiLAM for the diagnosis of childhood tuberculosis: a systematic review**. BMJ Paediatr Open *6*, e001447, doi:10.1136/bmjpo-2022-001447.

Pai, M., Dewan, P. K. and Swaminathan, S. (2023). **Transforming tuberculosis diagnosis**. Nat Microbiol *8*, 756–759, doi:10.1038/s41564-023-01365-3.

Pai, M., Kasaeva, T. and Swaminathan, S. (2022). **Covid-19's Devastating Effect on Tuberculosis Care — A Path to Recovery**. New England Journal of Medicine 1–3, doi:10.1056/NEJMp2118145.

Pai, N. P., Balram, B., Shivkumar, S., Martinez-Cajas, J. L., Claessens, C., Lambert, G., Peeling, R. W. and Joseph, L. (2012). **Head-to-head comparison of accuracy of a rapid point-of-care HIV test with oral versus whole-blood specimens: a systematic review and meta-analysis**. Lancet Infect Dis *12*, 373–380, doi:10.1016/S1473-3099(11)70368-1.

Pan, Q., Wang, Q., Sun, X., Xia, X., Wu, S., Luo, F. and Zhang, X.-L. (2014). **Aptamer against mannose-capped lipoarabinomannan inhibits virulent Mycobacterium tuberculosis infection in mice and rhesus monkeys.** Mol Ther *22*, 940–951, doi:10.1038/mt.2014.31.

Paris, L., Magni, R., Zaidi, F., Araujo, R., Saini, N., Harpole, M., Coronel, J., Kirwan, D. E., Steinberg, H., Gilman, R. H., Petricoin, E. F., Nisini, R., Luchini, A. and Liotta, L. (2017). **Urine lipoarabinomannan glycan in HIV-negative patients with pulmonary tuberculosis correlates with disease severity**. Sci Transl Med *9*, eaal2807, doi:10.1126/scitranslmed.aal2807.

Pavelka, M., Van-Zandvoort, K., Abbott, S., Sherratt, K., Majdan, M., Jarčuška, P., Krajčí, M., Flasche, S. and Funk, S. (2021). **The impact of population-wide rapid antigen testing on SARS-CoV-2 prevalence in Slovakia**. Science (1979) *372*, 635–641, doi:10.1126/science.abf9648.

Peter, J. G., Zijenah, L. S., Chanda, D., Clowes, P., Lesosky, M., Gina, P., Mehta, N., Calligaro, G., Lombard, C. J., Kadzirange, G., Bandason, T., Chansa, A., Liusha, N., Mangu, C., Mtafya, B., Msila, H., Rachow, A., Hoelscher, M., Mwaba, P., Theron, G. and Dheda, K. (2016). **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**. The Lancet *387*, 1187–1197, doi:10.1016/S0140-6736(15)01092-2.

Polechová, J., Johnson, K. D., Payne, P., Crozier, A., Beiglböck, M., Plevka, P. and Schernhammer, E. (2022). **SARS-CoV-2 rapid antigen tests provide benefits for epidemic control – observations from Austrian schools**. J Clin Epidemiol *145*, 14–19, doi:10.1016/j.jclinepi.2022.01.002.

Portevin, D., Moukambi, F., Clowes, P., Bauer, A., Chachage, M., Ntinginya, N. E., Mfinanga, E., Said, K., Haraka, F., Rachow, A., Saathoff, E., Mpina, M., Jugheli, L., Lwilla, F., Marais, B. J., Hoelscher, M., Daubenberger, C., Reither, K. and Geldmacher, C. (2014). **Assessment of the novel T-cell activation marker–tuberculosis assay for diagnosis of active tuberculosis in children: a prospective proof-of-concept study**. Lancet Infect Dis *14*, 931–938, doi:10.1016/S1473-3099(14)70884-9.

R2D2 (2024), URL: https://www.r2d2tbnetwork.org/productdevelopers [Accessed April 2024].

Rachow, A., Clowes, P., Saathoff, E., Mtafya, B., Michael, E., Ntinginya, E. N., Kowour, D., Rojas-Ponce, G., Kroidl, A., Maboko, L., Heinrich, N., Reither, K. and Hoelscher, M. (2012).

Increased and Expedited Case Detection by Xpert MTB/RIF Assay in Childhood Tuberculosis: A Prospective Cohort Study. Clinical Infectious Diseases *54*, 1388–1396, doi:10.1093/cid/cis190.

Reddy, K. P., Denkinger, C. M., Broger, T., McCann, N. C., Gupta-Wright, A., Kerkhoff, A. D., Pei, P. P., Shebl, F. M., Fielding, K. L., Nicol, M. P., Horsburgh, C. R., Meintjes, G., Freedberg, K. A., Wood, R. and Walensky, R. P. (2021). **Cost-effectiveness of a Novel Lipoarabinomannan Test for Tuberculosis in Patients With Human Immunodeficiency Virus**. Clinical Infectious Diseases *73*, e2077–e2085, doi:10.1093/cid/ciaa1698.

Reddy, K. P., Gupta-Wright, A., Fielding, K. L., Costantini, S., Zheng, A., Corbett, E. L., Yu, L., van Oosterhout, J. J., Resch, S. C., Wilson, D. P., Horsburgh, C. R., Wood, R., Alufandika-Moyo, M., Peters, J. A., Freedberg, K. A., Lawn, S. D. and Walensky, R. P. (2019). **Cost-effectiveness of urine-based tuberculosis screening in hospitalised patients with HIV in Africa: a microsimulation modelling study**. Lancet Glob Health *7*, e200–e208, doi:10.1016/S2214-109X(18)30436-4.

Ricks, S., Denkinger, C. M., Schumacher, S. G., Hallett, T. B. and Arinaminpathy, N. (2020). **The potential impact of urine-LAM diagnostics on tuberculosis incidence and mortality: A modelling analysis**. PLoS Med *17*, e1003466, doi:10.1371/journal.pmed.1003466.

Rojas, M., Barrera, L. F., Puzo, G. and Garcia, L. F. (1997). **Differential induction of apoptosis by virulent Mycobacterium tuberculosis in resistant and susceptible murine macrophages: role of nitric oxide and mycobacterial products.** J Immunol *159*, 1352–1361, doi:9233632.

Rojas, M., García, L. F., Nigou, J., Puzo, G. and Olivier, M. (2000). **Mannosylated lipoarabinomannan antagonizes Mycobacterium tuberculosis-induced macrophage apoptosis by altering Ca+2-dependent cell signaling.** J Infect Dis *182*, 240–251, doi:10.1086/315676.

Ryckman, T. S., Dowdy, D. W. and Kendall, E. A. (2022). **Infectious and clinical tuberculosis trajectories: Bayesian modeling with case finding implications**. Proceedings of the National Academy of Sciences *119*, 2017, doi:10.1073/pnas.2211045119.

Sarkar, P., Biswas, D., Sindhwani, G., Rawat, J., Kotwal, A. and Kakati, B. (2014). **Application of lipoarabinomannan antigen in tuberculosis diagnostics: current evidence**. Postgrad Med J *90*, 155–163, doi:10.1136/postgradmedj-2013-132053.

Schumacher, S. G., van Smeden, M., Dendukuri, N., Joseph, L., Nicol, M. P., Pai, M. and Zar, H. J. (2016). **Diagnostic Test Accuracy in Childhood Pulmonary Tuberculosis: A Bayesian Latent Class Analysis**. Am J Epidemiol *184*, 690–700, doi:10.1093/aje/kww094.

Schünemann, H. J., Mustafa, R. A., Brozek, J., Santesso, N., Bossuyt, P. M., Steingart, K. R., Leeflang, M., Lange, S., Trenti, T., Langendam, M., Scholten, R., Hooft, L., Murad, M. H., Jaeschke, R., Rutjes, A., Singh, J., Helfand, M., Glasziou, P., Arevalo-Rodriguez, I., Akl, E. A., Deeks, J. J. and Guyatt, G. H. (2019). **GRADE guidelines: 22. The GRADE approach for tests and strategies—from test accuracy to patient-important outcomes and recommendations**. J Clin Epidemiol *111*, 69–82, doi:10.1016/j.jclinepi.2019.02.003.

Schünemann, H. J., Mustafa, R., Brozek, J., Santesso, N., Alonso-Coello, P., Guyatt, G., Scholten, R., Langendam, M., Leeflang, M. M., Akl, E. A., Singh, J. A., Meerpohl, J., Hultcrantz, M., Bossuyt, P., Oxman, A. D., Schünemann, H. J., Mustafa, R., Brozek, J., Santesso, N., Alonso-Coello, P., Scholten, R., Langendam, M., Bossuyt, P., Leeflang, M. M., Akl, E. A., Singh, J., Meerpohl, J., Hultcrantz, M., Guyatt, G., Oxman, A. D., Lange, S., Parmelli, E., Moberg, J., Rosenbaum, S., Brignardello-Petersen, R., Wiercioch, W., Davoli, M., Nowak, A. and Dietl, B. (2016). **GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health**. J Clin Epidemiol *76*, 89–98, doi:10.1016/j.jclinepi.2016.01.032.

Schünemann, H. J., Oxman, A. D., Brozek, J., Glasziou, P., Jaeschke, R., Vist, G. E., Williams, J. W., Kunz, R., Craig, J., Montori, V. M., Bossuyt, P. and Guyatt, G. H. (2008). **Grading quality of evidence and strength of recommendations for diagnostic tests and strategies**. BMJ *336*, 1106–1110, doi:10.1136/bmj.39500.677199.AE.

Shengelia, B., Tandon, A., Adams, O. B. and Murray, C. J. L. (2005). **Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy**. Soc Sci Med *61*, 97–109, doi:10.1016/j.socscimed.2004.11.055.

Sigal, G. B., Pinter, A., Lowary, T. L., Kawasaki, M., Li, A., Mathew, A., Tsionsky, M., Zheng, R. B., Plisova, T., Shen, K., Katsuragi, K., Choudhary, A., Honnen, W. J., Nahid, P., Denkinger, C. M. and Broger, T. (2018). **A Novel Sensitive Immunoassay Targeting the 5-Methylthio-D -Xylofuranose–Lipoarabinomannan Epitope Meets the WHO's Performance Target for Tuberculosis Diagnosis**. J Clin Microbiol *56*, 1–17, doi:10.1128/JCM.01338-18.

Silva, S., Arinaminpathy, N., Atun, R., Goosby, E. and Reid, M. (2021). **Economic impact of tuberculosis mortality in 120 countries and the cost of not achieving the Sustainable Development Goals tuberculosis targets: a full-income analysis**. Lancet Glob Health *9*, e1372–e1379, doi:10.1016/S2214-109X(21)00299-0.

Singhroy, D. N., MacLean, E., Kohli, M., Lessem, E., Branigan, D., England, K., Suleiman, K., Drain, P. K., Ruhwald, M., Schumacher, S., Denkinger, C. M., Waning, B., Van Gemert, W. and Pai, M. (2020). **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 *4*, 24, doi:10.12688/gatesopenres.13112.1.

Sossen, B., Broger, T., Kerkhoff, A. D., Schutz, C., Trollip, A., Moreau, E., Schumacher, S. G., Burton, R., Ward, A., Wilkinson, R. J., Barr, D. A., Nicol, M. P., Denkinger, C. M. and Meintjes, G. (2020). **"SILVAMP TB LAM" Rapid Urine Tuberculosis Test Predicts Mortality in Patients Hospitalized With Human Immunodeficiency Virus in South Africa**. Clinical Infectious Diseases *71*, 1973–1976, doi:10.1093/cid/ciaa024.

Stalford, S. a, Fascione, M. a, Sasindran, S. J., Chatterjee, D., Dhandayuthapani, S. and Turnbull, W. B. (2009). **A natural carbohydrate substrate for Mycobacterium tuberculosis methionine sulfoxide reductase A.** Chem Commun (Camb) 110–112, doi:10.1039/b817483k.

Steadman, A., Andama, A., Ball, A., Mukwatamundu, J., Khimani, K., Mochizuki, T., Asege, L., Bukirwa, A., Kato, J. B., Katumba, D., Kisakye, E., Mangeni, W., Mwebe, S., Nakaye, M., Nassuna, I., Nyawere, J., Nakaweesa, A., Cook, C., Phillips, P., Nalugwa, T., Bachman, C. M., Semitala, F. C., Weigl, B. H., Connelly, J., Worodria, W. and Cattamanchi, A. (2024). **New Manual Quantitative Polymerase Chain Reaction Assay Validated on Tongue Swabs Collected and Processed in Uganda Shows Sensitivity That Rivals Sputum-based Molecular Tuberculosis Diagnostics**. Clinical Infectious Diseases *78*, 1313–1320, doi:10.1093/cid/ciae041.

Subbaraman, R., Nathavitharana, R. R., Mayer, K. H., Satyanarayana, S., Chadha, V. K., Arinaminpathy, N. and Pai, M. (2019). **Constructing care cascades for active tuberculosis: A strategy for program monitoring and identifying gaps in quality of care**. PLoS Med *16*, e1002754, doi:10.1371/journal.pmed.1002754.

Székely, R., Sossen, B., Mukoka, M., Muyoyeta, M., Reither, K., Nakiyingi, L., Kerkhoff, A. D., MacPherson, P., Meintjes, G., Denkinger, C. M. and Ruhwald, M. (2022). **Multicentre accuracy trial of FUJIFILM SILVAMP TB LAM test in people with HIV reveals lot variability**. medrxiv, doi:10.1101/2022.09.07.22278961.

Thepsamarn, P., Prayoollawongsa, N., Puksupa, P., Puttoom, P., Thaidumrong, P., Wongchai, S., Doddara, J., Tantayarak, J., Buchachart, K., Wilairatana, P. and Looareesuwan, S. (1997). **The ICT Malaria Pf: a simple, rapid dipstick test for the diagnosis of Plasmodium falciparum malaria at the Thai-Myanmar border.** Southeast Asian J Trop Med Public Health *28*, 723–726.

Torrelles, J. B., Azad, A. K. and Schlesinger, L. S. (2006). **Fine discrimination in the recognition of individual species of phosphatidyl-myo-inositol mannosides from Mycobacterium tuberculosis by C-type lectin pattern recognition receptors.** J Immunol *177*, 1805–1816, doi:177/3/1805 [pii].

Treatment Action Group (TAG) (2023). **Tuberculosis Diagnostics Pipeline Report 2023**, URL: https://www.treatmentactiongroup.org/wp-content/uploads/2023/11/2023_pipeline_TB_diagnostics_final.pdf [Accessed May 2024].

Venisse, A., Berjeaud, J. M., Chaurand, P., Gilleron, M. and Puzo, G. (1993). **Structural features of lipoarabinomannan from Mycobacterium bovis BCG. Determination of molecular mass by laser desorption mass spectrometry.** J Biol Chem *268*, 12401–12411.

Vergne, I., Gilleron, M. and Nigou, J. (2014). **Manipulation of the endocytic pathway and phagocyte functions by Mycobacterium tuberculosis lipoarabinomannan.** Front Cell Infect Microbiol *4*, 187, doi:10.3389/fcimb.2014.00187.

Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., Abdollahi, M., Abdollahpour, I., Abolhassani, H., Aboyans, V., Abrams, E. M., Abreu, L. G., Abrigo, M. R. M., Abu-Raddad, L. J., Abushouk, A. I., Acebedo, A., Ackerman, I. N., Adabi, M., Adamu, A. A., Adebayo, O. M., Adekanmbi, V., Adelson, J. D., Adetokunboh, O. O., Adham, D., Afshari, M., Afshin, A.,

Agardh, E. E., Agarwal, G., Agesa, K. M., Aghaali, M., Aghamir, S. M. K., Agrawal, A., Ahmad, T., Ahmadi, A., Ahmadi, M., Ahmadieh, H., Ahmadpour, E., Akalu, T. Y., Akinyemi, R. O., Akinyemiju, T., Akombi, B., Al-Aly, Z., Alam, K., Alam, N., Alam, S., Alam, T., Alanzi, T. M., Albertson, S. B., Alcalde-Rabanal, J. E., Alema, N. M., Ali, M., Ali, S., Alicandro, G., Alijanzadeh, M., Alinia, C., Alipour, V., Aljunid, S. M., Alla, F., Allebeck, P., Almasi-Hashiani, A., Alonso, J., Al-Raddadi, R. M., Altirkawi, K. A., Alvis-Guzman, N., Alvis-Zakzuk, N. J., Amini, S., Amini-Rarani, M., Aminorroaya, A., Amiri, F., Amit, A. M. L., Amugsi, D. A., Amul, G. G. H., Anderlini, D., Andrei, C. L., Andrei, T., Anjomshoa, M., Ansari, F., Ansari, I., Ansari-Moghaddam, A., Antonio, C. A. T., Antony, C. M., Antriyandarti, E., Anvari, D., Anwer, R., Arabloo, J., Arab-Zozani, M., Aravkin, A. Y., Ariani, F., Ärnlöv, J., Aryal, K. K., Arzani, A., Asadi-Aliabadi, M., Asadi-Pooya, A. A., Asghari, B., Ashbaugh, C., Atnafu, D. D., Atre, S. R., Ausloos, F., Ausloos, M., Ayala Quintanilla, B. P., Ayano, G., Ayanore, M. A., Aynalem, Y. A., Azari, S., Azarian, G., Azene, Z. N., Babaee, E., Badawi, A., Bagherzadeh, M., Bakhshaei, M. H., Bakhtiari, A., Balakrishnan, S., Balalla, S., Balassyano, S., Banach, M., Banik, P. C., Bannick, M. S., Bante, A. B., Baraki, A. G., Barboza, M. A., Barker-Collo, S. L., Barthelemy, C. M., Barua, L., Barzegar, A., Basu, S., Baune, B. T., Bayati, M., Bazmandegan, G., Bedi, N., Beghi, E., Béjot, Y., Bello, A. K., Bender, R. G., Bennett, D. A., Bennitt, F. B., Bensenor, I. M., Benziger, C. P., Berhe, K., Bernabe, E., Bertolacci, G. J., Bhageerathy, R., Bhala, N., Bhandari, D., Bhardwaj, P., Bhattacharyya, K., Bhutta, Z. A., Bibi, S., Biehl, M. H., Bikbov, B., Bin Sayeed, M. S., Biondi, A., Birihane, B. M., Bisanzio, D., Bisignano, C., Biswas, R. K., Bohlouli, S., Bohluli, M., Bolla, S. R. R., Boloor, A., Boon-Dooley, A. S., Borges, G., Borzì, A. M., Bourne, R., Brady, O. J., Brauer, M., Brayne, C., Breitborde, N. J. K., Brenner, H., Briant, P. S., Briggs, A. M., Briko, N. I., Britton, G. B., Bryazka, D., Buchbinder, R., Bumgarner, B. R., Busse, R., Butt, Z. A., Caetano dos Santos, F. L., Cámera, L. L. A., Campos-Nonato, I. R., Car, J., Cárdenas, R., Carreras, G., Carrero, J. J., Carvalho, F., Castaldelli-Maia, J. M., Castañeda-Orjuela, C. A., Castelpietra, G., Castle, C. D., Castro, F., Catalá-López, F., Causey, K., Cederroth, C. R., Cercy, K. M., Cerin, E., Chandan, J. S., Chang, A. R., Charlson, F. J., Chattu, V. K., Chaturvedi, S., Chimed-Ochir, O., Chin, K. L., Cho, D. Y., Christensen, H., Chu, D.-T., Chung, M. T., Cicuttini, F. M., Ciobanu, L. G., Cirillo, M., Collins, E. L., Compton, K., Conti, S., Cortesi, P. A., Costa, V. M., Cousin, E., Cowden, R. G., Cowie, B. C., Cromwell, E. A., Cross, D. H., Crowe, C. S., Cruz, J. A., Cunningham, M., Dahlawi, S. M. A., Damiani, G., Dandona, L., Dandona, R., Darwesh, A. M., Daryani, A., Das, J. K., Das Gupta, Rajat, das Neves, J., Dávila-Cervantes, C. A., Davletov, K., De Leo, D., Dean, F. E., DeCleene, N. K., Deen, A., Degenhardt, L., Dellavalle, R. P., Demeke, F. M., Demsie, D. G., Denova-Gutiérrez, E., Dereje, N. D., Dervenis, N., Desai, R., Desalew, A., Dessie, G. A., Dharmaratne, S. D., Dhungana, G. P., Dianatinasab, M., Diaz, D., Dibaji Forooshani, Z. S., Dingels, Z. V., Dirac, M. A., Djalalinia, S., Do, H. T., Dokova, K., Dorostkar, F., Doshi, C. P., Doshmangir, L., Douiri, A., Doxey, M. C., Driscoll, T. R., Dunachie, S. J., Duncan, B. B., Duraes, A. R., Eagan, A. W., Ebrahimi Kalan, M., Edvardsson, D., Ehrlich, J. R., El Nahas, N., El Sayed, I., El Tantawi, M., Elbarazi, I., Elgendy, I. Y., Elhabashy, H. R., El-Jaafary, S. I., Elyazar, I. R., Emamian, M. H., Emmons-Bell, S., Erskine, H. E., Eshrati, B., Eskandarieh, S., Esmaeilnejad, S., Esmaeilzadeh, F., Esteghamati, A., Estep, K., Etemadi, A., Etisso, A. E., Farahmand, M., Faraj, A., Fareed, M., Faridnia, R., Farinha, C. S. e S., Farioli, A., Faro, A.,

Faruque, M., Farzadfar, F., Fattahi, N., Fazlzadeh, M., Feigin, V. L., Feldman, R., Fereshtehnejad, S.-M., Fernandes, E., Ferrari, A. J., Ferreira, M. L., Filip, I., Fischer, F., Fisher, J. L., Fitzgerald, R., Flohr, C., Flor, L. S., Foigt, N. A., Folayan, M. O., Force, L. M., Fornari, C., Foroutan, M., Fox, J. T., Freitas, M., Fu, W., Fukumoto, T., Furtado, J. M., Gad, M. M., Gakidou, E., Galles, N. C., Gallus, S., Gamkrelidze, A., Garcia-Basteiro, A. L., Gardner, W. M., Geberemariyam, B. S., Gebrehiwot, A. M., Gebremedhin, K. B., Gebreslassie, A. A. A. A., Gershberg Hayoon, A., Gething, P. W., Ghadimi, M., Ghadiri, K., Ghafourifard, M., Ghajar, A., Ghamari, F., Ghashghaee, A., Ghiasvand, H., Ghith, N., Gholamian, A., Gilani, S. A., Gill, P. S., Gitimoghaddam, M., Giussani, G., Goli, S., Gomez, R. S., Gopalani, S. V., Gorini, G., Gorman, T. M., Gottlich, H. C., Goudarzi, H., Goulart, A. C., Goulart, B. N. G., Grada, A., Grivna, M., Grosso, G., Gubari, M. I. M., Gugnani, H. C., Guimaraes, A. L. S., Guimarães, R. A., Guled, R. A., Guo, G., Guo, Y., Gupta, Rajeev, Haagsma, J. A., Haddock, B., Hafezi-Nejad, N., Hafiz, A., Hagins, H., Haile, L. M., Hall, B. J., Halvaei, I., Hamadeh, R. R., Hamagharib Abdullah, K., Hamilton, E. B., Han, C., Han, H., Hankey, G. J., Haro, J. M., Harvey, J. D., Hasaballah, A. I., Hasanzadeh, A., Hashemian, M., Hassanipour, S., Hassankhani, H., Havmoeller, R. J., Hay, R. J., Hay, S. I., Hayat, K., Heidari, B., Heidari, G., Heidari-Soureshjani, R., Hendrie, D., Henrikson, H. J., Henry, N. J., Herteliu, C., Heydarpour, F., Hird, T. R., Hoek, H. W., Hole, M. K., Holla, R., Hoogar, P., Hosgood, H. D., Hosseinzadeh, M., Hostiuc, M., Hostiuc, S., Househ, M., Hoy, D. G., Hsairi, M., Hsieh, V. C., Hu, G., Huda, T. M., Hugo, F. N., Huynh, C. K., Hwang, B.-F., Iannucci, V. C., Ibitoye, S. E., Ikuta, K. S., Ilesanmi, O. S., Ilic, I. M., Ilic, M. D., Inbaraj, L. R., Ippolito, H., Irvani, S. S. N., Islam, M. M., Islam, M. M., Islam, S. M. S., Islami, F., Iso, H., Ivers, R. Q., Iwu, C. C. D., Iyamu, I. O., Jaafari, J., Jacobsen, K. H., Jadidi-Niaragh, F., Jafari, H., Jafarinia, M., Jahagirdar, D., Jahani, M. A., Jahanmehr, N., Jakovljevic, M., Jalali, A., Jalilian, F., James, S. L., Janjani, H., Janodia, M. D., Jayatilleke, A. U., Jeemon, P., Jenabi, E., Jha, R. P., Jha, V., Ji, J. S., Jia, P., John, O., John-Akinola, Y. O., Johnson, C. O., Johnson, S. C., Jonas, J. B., Joo, T., Joshi, A., Jozwiak, J. J., Jürisson, M., Kabir, A., Kabir, Z., Kalani, H., Kalani, R., Kalankesh, L. R., Kalhor, R., Kamiab, Z., Kanchan, T., Karami Matin, B., Karch, A., Karim, M. A., Karimi, S. E., Kassa, G. M., Kassebaum, N. J., Katikireddi, S. V., Kawakami, N., Kayode, G. A., Keddie, S. H., Keller, C., Kereselidze, M., Khafaie, M. A., Khalid, N., Khan, M., Khatab, K., Khater, M. M., Khatib, M. N., Khayamzadeh, M., Khodayari, M. T., Khundkar, R., Kianipour, N., Kieling, C., Kim, D., Kim, Y. J. Y.-E., Kim, Y. J. Y.-E., Kimokoti, R. W., Kisa, A., Kisa, S., Kissimova-Skarbek, K., Kivimäki, M., Kneib, C. J., Knudsen, A. K. S., Kocarnik, J. M., Kolola, T., Kopec, J. A., Kosen, S., Koul, P. A., Koyanagi, A., Kravchenko, M. A., Krishan, K., Krohn, K. J., Kuate Defo, B., Kucuk Bicer, B., Kumar, G. A., Kumar, M., Kumar, P., Kumar, V., Kumaresh, G., Kurmi, O. P., Kusuma, D., Kyu, H. H., La Vecchia, C., Lacey, B., Lal, D. K., Lalloo, R., Lam, J. O., Lami, F. H., Landires, I., Lang, J. J., Lansingh, V. C., Larson, S. L., Larsson, A. O., Lasrado, S., Lassi, Z. S., Lau, K. M.-M., Lavados, P. M., Lazarus, J. V., Ledesma, J. R., Lee, P. H., Lee, S. W. H., LeGrand, K. E., Leigh, J., Leonardi, M., Lescinsky, H., Leung, J., Levi, M., Lewington, S., Li, S., Lim, L.-L., Lin, C., Lin, R.-T., Linehan, C., Linn, S., Liu, H.-C., Liu, S., Liu, Z., Looker, K. J., Lopez, A. D., Lopukhov, P. D., Lorkowski, S., Lotufo, P. A., Lucas, T. C. D., Lugo, A., Lunevicius, R., Lyons, R. A., Ma, J., MacLachlan, J. H., Maddison, E. R., Maddison, R., Madotto, F., Mahasha, P. W., Mai, H. T., Majeed, A.,

Maled, V., Maleki, S., Malekzadeh, R., Malta, D. C., Mamun, A. A., Manafi, A., Manafi, N., Manguerra, H., Mansouri, B., Mansournia, M. A., Mantilla Herrera, A. M., Maravilla, J. C., Marks, A., Martins-Melo, F. R., Martopullo, I., Masoumi, S. Z., Massano, J., Massenburg, B. B., Mathur, M. R., Maulik, P. K., McAlinden, C., McGrath, J. J., McKee, M., Mehndiratta, M. M., Mehri, F., Mehta, K. M., Meitei, W. B., Memiah, P. T. N., Mendoza, W., Menezes, R. G., Mengesha, E. W., Mengesha, M. B., Mereke, A., Meretoja, A., Meretoja, T. J., Mestrovic, T., Miazgowski, B., Miazgowski, T., Michalek, I. M., Mihretie, K. M., Miller, T. R., Mills, E. J., Mirica, A., Mirrakhimov, E. M., Mirzaei, H., Mirzaei, M., Mirzaei-Alavijeh, M., Misganaw, A. T., Mithra, P., Moazen, B., Moghadaszadeh, M., Mohamadi, E., Mohammad, D. K., Mohammad, Y., Mohammad Gholi Mezerji, N., Mohammadian-Hafshejani, A., Mohammadifard, N., Mohammadpourhodki, R., Mohammed, S., Mokdad, A. H., Molokhia, M., Momen, N. C., Monasta, L., Mondello, S., Mooney, M. D., Moosazadeh, M., Moradi, G., Moradi, M., Moradi-Lakeh, M., Moradzadeh, R., Moraga, P., Morales, L., Morawska, L., Moreno Velásquez, I., Morgado-da-Costa, J., Morrison, S. D., Mosser, J. F., Mouodi, S., Mousavi, S. M., Mousavi Khaneghah, A., Mueller, U. O., Munro, S. B., Muriithi, M. K., Musa, K. I., Muthupandian, S., Naderi, M., Nagarajan, A. J., Nagel, G., Naghshtabrizi, B., Nair, S., Nandi, A. K., Nangia, V., Nansseu, J. R., Nayak, V. C., Nazari, J., Negoi, I., Negoi, R. I., Netsere, H. B. N., Ngunjiri, J. W., Nguyen, C. T., Nguyen, J., Nguyen, M. M., Nguyen, M. M., Nichols, E., Nigatu, D., Nigatu, Y. T., Nikbakhsh, R., Nixon, M. R., Nnaji, C. A., Nomura, S., Norrving, B., Noubiap, J. J., Nowak, C., Nunez-Samudio, V., Oţoiu, A., Oancea, B., Odell, C. M., Ogbo, F. A., Oh, I.-H., Okunga, E. W., Oladnabi, M., Olagunju, A. T., Olusanya, B. O., Olusanya, J. O., Oluwasanu, M. M., Omar Bali, A., Omer, M. O., Ong, K. L., Onwujekwe, O. E., Orji, A. U., Orpana, H. M., Ortiz, A., Ostroff, S. M., Otstavnov, N., Otstavnov, S. S., Øverland, S., Owolabi, M. O., P A, M., Padubidri, J. R., Pakhare, A. P., Palladino, R., Pana, A., Panda-Jonas, S., Pandey, A., Park, E.-K., Parmar, P. G. K., Pasupula, D. K., Patel, S. K., Paternina-Caicedo, A. J., Pathak, A., Pathak, M., Patten, S. B., Patton, G. C., Paudel, D., Pazoki Toroudi, H., Peden, A. E., Pennini, A., Pepito, V. C. F., Peprah, E. K., Pereira, A., Pereira, D. M., Perico, N., Pham, H. Q., Phillips, M. R., Pigott, D. M., Pilgrim, T., Pilz, T. M., Pirsaheb, M., Plana-Ripoll, O., Plass, D., Pokhrel, K. N., Polibin, R. V., Polinder, S., Polkinghorne, K. R., Postma, M. J., Pourjafar, H., Pourmalek, F., Pourmirza Kalhori, R., Pourshams, A., Poznańska, A., Prada, S. I., Prakash, V., Pribadi, D. R. A., Pupillo, E., Quazi Syed, Z., Rabiee, M., Rabiee, N., Radfar, A., Rafiee, A., Rafiei, A., Raggi, A., Rahimi-Movaghar, A., Rahman, M. A., Rajabpour-Sanati, A., Rajati, F., Ramezanzadeh, K., Ranabhat, C. L., Rao, P. C., Rao, S. J., Rasella, D., Rastogi, P., Rathi, P., Rawaf, D. L., Rawaf, S., Rawal, L., Razo, C., Redford, S. B., Reiner, R. C., Reinig, N., Reitsma, M. B., Remuzzi, G., Renjith, V., Renzaho, A. M. N., Resnikoff, S., Rezaei, N., Rezai, M. sadegh, Rezapour, A., Rhinehart, P.-A., Riahi, S. M., Ribeiro, A. L. P., Ribeiro, D. C., Ribeiro, D. C., Rickard, J., Roberts, N. L. S., Roberts, S., Robinson, S. R., Roever, L., Rolfe, S., Ronfani, L., Roshandel, G., Roth, G. A., Rubagotti, E., Rumisha, S. F., Sabour, S., Sachdev, P. S., Saddik, B., Sadeghi, E., Sadeghi, M., Saeidi, S., Safi, S., Safiri, S., Sagar, R., Sahebkar, A., Sahraian, M. A., Sajadi, S. M., Salahshoor, M. R., Salamati, P., Salehi Zahabi, S., Salem, H., Salem, M. R. R., Salimzadeh, H., Salomon, J. A., Salz, I., Samad, Z., Samy, A. M., Sanabria, J., Santomauro, D. F., Santos, I. S., Santos, J. V., Santric-Milicevic, M. M.,

Saraswathy, S. Y. I., Sarmiento-Suárez, R., Sarrafzadegan, N., Sartorius, B., Sarveazad, A., Sathian, B., Sathish, T., Sattin, D., Sbarra, A. N., Schaeffer, L. E., Schiavolin, S., Schmidt, M. I., Schutte, A. E., Schwebel, D. C., Schwendicke, F., Senbeta, A. M., Senthilkumaran, S., Sepanlou, S. G., Shackelford, K. A., Shadid, J., Shahabi, S., Shaheen, A. A., Shaikh, M. A., Shalash, A. S., Shams-Beyranvand, M., Shamsizadeh, M., Shannawaz, M., Sharafi, K., Sharara, F., Sheena, B. S., Sheikhtaheri, A., Shetty, R. S., Shibuya, K., Shiferaw, W. S., Shigematsu, M., Shin, J. Il, Shiri, R., Shirkoohi, R., Shrime, M. G., Shuval, K., Siabani, S., Sigfusdottir, I. D., Sigurvinsdottir, R., Silva, J. P., Simpson, K. E., Singh, A., Singh, J. A., Skiadaresi, E., Skou, S. T. S., Skryabin, V. Y., Sobngwi, E., Sokhan, A., Soltani, S., Sorensen, R. J. D., Soriano, J. B., Sorrie, M. B., Soyiri, I. N., Sreeramareddy, C. T., Stanaway, J. D., Stark, B. A., Ştefan, S. C., Stein, C., Steiner, C., Steiner, T. J., Stokes, M. A., Stovner, L. J., Stubbs, J. L., Sudaryanto, A., Sufiyan, M. B., Sulo, G., Sultan, I., Sykes, B. L., Sylte, D. O., Szócska, M., Tabarés-Seisdedos, R., Tabb, K. M., Tadakamadla, S. K., Taherkhani, A., Tajdini, M., Takahashi, K., Taveira, N., Teagle, W. L., Teame, H., Tehrani-Banihashemi, A., Teklehaimanot, B. F., Terrason, S., Tessema, Z. T., Thankappan, K. R., Thomson, A. M., Tohidinik, H. R., Tonelli, M., Topor-Madry, R., Torre, A. E., Touvier, M., Tovani-Palone, M. R. R., Tran, B. X., Travillian, R., Troeger, C. E., Truelsen, T. C., Tsai, A. C., Tsatsakis, A., Tudor Car, L., Tyrovolas, S., Uddin, R., Ullah, S., Undurraga, E. A., Unnikrishnan, B., Vacante, M., Vakilian, A., Valdez, P. R., Varughese, S., Vasankari, T. J., Vasseghian, Y., Venketasubramanian, N., Violante, F. S., Vlassov, V., Vollset, S. E., Vongpradith, A., Vukovic, A., Vukovic, R., Waheed, Y., Walters, M. K., Wang, J., Wang, Y. Y.-P., Wang, Y. Y.-P., Ward, J. L., Watson, A., Wei, J., Weintraub, R. G., Weiss, D. J., Weiss, J., Westerman, R., Whisnant, J. L., Whiteford, H. A., Wiangkham, T., Wiens, K. E., Wijeratne, T., Wilner, L. B., Wilson, S., Wojtyniak, B., Wolfe, C. D. A., Wool, E. E., Wu, A.-M., Wulf Hanson, S., Wunrow, H. Y., Xu, G., Xu, R., Yadgir, S., Yahyazadeh Jabbari, S. H., Yamagishi, K., Yaminfirooz, M., Yano, Y., Yaya, S., Yazdi-Feyzabadi, V., Yearwood, J. A., Yeheyis, T. Y., Yeshitila, Y. G., Yip, P., Yonemoto, N., Yoon, S.-J., Yoosefi Lebni, J., Younis, M. Z., Younker, T. P., Yousefi, Z., Yousefifard, M., Yousefinezhadi, T., Yousuf, A. Y., Yu, C., Yusefzadeh, H., Zahirian Moghadam, T., Zaki, L., Zaman, S. Bin, Zamani, M., Zamanian, M., Zandian, H., Zangeneh, A., Zastrozhin, M. S., Zewdie, K. A., Zhang, Y., Zhang, Z.-J., Zhao, J. T., Zhao, Y., Zheng, P., Zhou, M., Ziapour, A., Zimsen, S. R. M., Naghavi, M. and Murray, C. J. L. (2020). **Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019**. The Lancet *396*, 1204–1222, doi:10.1016/S0140-6736(20)30925-9.

Whiting, P. F. (2011). **QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies**. Ann Intern Med *155*, 529, doi:10.7326/0003-4819-155-8-201110180-00009.

Witzel, T. C., Eshun-Wilson, I., Jamil, M. S., Tilouche, N., Figueroa, C., Johnson, C. C., Reid, D., Baggaley, R., Siegfried, N., Burns, F. M., Rodger, A. J. and Weatherburn, P. (2020). **Comparing the effects of HIV self-testing to standard HIV testing for key populations: a systematic review and meta-analysis**. BMC Med *18*, 1–13, doi:10.1186/s12916-020-01835-z.

Wood, A., Barizuddin, S., Darr, C. M., Mathai, C. J., Ball, A., Minch, K., Somoskovi, A., Hamasur, B., Connelly, J. T., Weigl, B., Andama, A., Cattamanchi, A., Gangopadhyay, K., Bok, S., Gangopadhyay, S., Weigle, B., Andaman, A., Cattamanchi, A., Gangopadhyay, K., Bok, S. and Gangopadhyay, S. (2019). **Ultrasensitive detection of lipoarabinomannan with plasmonic grating biosensors in clinical samples of HIV negative patients with tuberculosis**. PLoS One *14*, e0214161, doi:10.1371/journal.pone.0214161.

World Health Organization (2000). **New perspectives: malaria diagnosis. Report of a Joint WHO/USAID Informal Consultation, 25-27 October 1999**, URL: https://fctc.who.int/publications/i/item/new-perspectives-malaria-diagnosis [Accessed May 2024].

World Health Organization (2013). **Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children policy update**, URL: https://www.who.int/publications/i/item/9789241501545 [Accessed May 2024].

World Health Organization (2014). **High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting**, URL: https://www.who.int/publications/i/item/WHO-HTM-TB-2014.18 [Accessed May 2024].

World Health Organization (2018). **Catalogue of the WHO international reference preparations**, URL: https://www.who.int/teams/health-product-policy-and-standards/standards-and-specifications/catalogue [Accessed March 2024].

World Health Organization (2019). **Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV - Policy update 2019**, URL: https://www.who.int/publications/i/item/9789241550604 [Accessed May 2024].

World Health Organization (2020). **Consolidated Guidelines on Tuberculosis - Module 3: Diagnosis**, URL: https://www.who.int/publications/i/item/9789240089488 [Accessed May 2024].

World Health Organization (2021a). **The selection and use of essential in vitro diagnostics**, URL: https://www.who.int/publications/i/item/9789240019102 [Accessed May 2024].

World Health Organization (2021b). **World Malaria Report 2021**, URL: https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021 [Accessed May 2024].

World Health Organization (2022). **Use of SARS-CoV-2 antigen-detection rapid diagnostic tests for COVID-19 self-testing**, URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Ag-RDTs-Self_testing-2022.1 [Accessed May 2024].

World Health Organization (2023a). **Global Tuberculosis Report 2023**, URL: https://iris.who.int/bitstream/handle/10665/373828/9789240083851-eng.pdf?sequence=1 [Accessed May 2024].

World Health Organization (2023b). **WHO standard Universal access to rapid tuberculosis diagnostics**, URL: https://www.who.int/publications/i/item/9789240071315 [Accessed May 2024].

World Health Organization (2024). **Target Product Profile TB diagnostic tests for peripheral settings**, URL: https://www.who.int/news-room/articles-detail/public-consultation-for-the-target-product-profile-tb-diagnostic-tests-for-peripheral-settings [Accessed May 2024].

World Health Organization (2024). **TSS-23 Rapid diagnostic tests to detect mycobacterial lipoarabinomannan (LAM) antigen in urine, draft for comment Technical specifications series for submission to WHO prequalification-diagnostic assessment**, URL: https://extranet.who.int/prequal/vitro-diagnostics/technical-specifications-series [Accessed May 2024].

Yakhelef, N., Audibert, M., Ferlazzo, G., Sitienei, J., Wanjala, S., Varaine, F., Bonnet, M. and Huerga, H. (2020). **Cost-effectiveness of diagnostic algorithms including lateral-flow urine lipoarabinomannan for HIV-positive patients with symptoms of tuberculosis**. PLoS One *15*, doi:10.1371/journal.pone.0227138.

Yan, Z.-H., Zhao, B., Pang, Y., Wang, X.-J., Yi, L., Wang, H.-L., Yang, B., Wei, P.-J., Jia, H.-Y., Li, S.-P., Zhao, Y.-L. and Zhang, H.-T. (2021). **Generation of mycobacterial lipoarabinomannan-specific monoclonal antibodies and their ability to identify mycobacterium isolates**. Journal of Microbiology, Immunology and Infection *54*, 437–446, doi:10.1016/j.jmii.2020.02.005.

Yerlikaya, S., Broger, T., Isaacs, C., Bell, D., Holtgrewe, L., Gupta-Wright, A., Nahid, P., Cattamanchi, A. and Denkinger, C. M. (2023). **Blazing the trail for innovative tuberculosis diagnostics**. Infection, doi:10.1007/s15010-023-02135-3.

Yerushalmy, J. (1947). **Statistical Problems in Assessing Methods of Medical Diagnosis, with Special Reference to X-Ray Techniques**. Public Health Reports (1896-1970) *62*, 1432, doi:10.2307/4586294.

Yimam, Y., Mohebali, M. and Abbaszadeh Afshar, M. J. (2022). **Comparison of diagnostic performance between conventional and ultrasensitive rapid diagnostic tests for diagnosis of malaria: A systematic review and meta-analysis**. PLoS One *17*, e0263770, doi:10.1371/journal.pone.0263770.

Young, B. C., Eyre, D. W., Kendrick, S., White, C., Smith, S., Beveridge, G., Nonnenmacher, T., Ichofu, F., Hillier, J., Oakley, S., Diamond, I., Rourke, E., Dawe, F., Day, I., Davies, L., Staite, P., Lacey, A., McCrae, J., Jones, F., Kelly, J., Bankiewicz, U., Tunkel, S., Ovens, R., Chapman, D., Bhalla, V., Marks, P., Hicks, N., Fowler, T., Hopkins, S., Yardley, L. and Peto, T. E. A. (2021). **Daily testing for contacts of individuals with SARS-CoV-2 infection and attendance and SARS-CoV-2 transmission in English secondary schools and colleges: an open-label, cluster-randomised trial**. The Lancet *398*, 1217–1229, doi:10.1016/S0140-6736(21)01908-5.

Zachary, D., Mwenge, L., Muyoyeta, M., Shanaube, K., Schaap, A., Bond, V., Kosloff, B., de Haas, P. and Ayles, H. (2012). **Field comparison of OraQuick® ADVANCE Rapid HIV-1/2 antibody test and two blood-based rapid HIV antibody tests in Zambia**. BMC Infect Dis *12*, 183, doi:10.1186/1471-2334-12-183.

Zar, H. J., Workman, L., Isaacs, W., Munro, J., Black, F., Eley, B., Allen, V., Boehme, C. C., Zemanay, W. and Nicol, M. P. (2012). **Rapid Molecular Diagnosis of Pulmonary Tuberculosis in Children Using Nasopharyngeal Specimens**. Clinical Infectious Diseases *55*, 1088–1095, doi:10.1093/cid/cis598.

Zhang, X., Barr, B., Green, M., Hughes, D., Ashton, M., Charalampopoulos, D., García-Fiñana, M. and Buchan, I. (2022). **Impact of community asymptomatic rapid antigen testing on covid-19 related hospital admissions: synthetic control study**. BMJ e071374, doi:10.1136/bmj-2022-071374.

Zheng, R. B., Jégouzo, S. A. F., Joe, M., Bai, Y., Tran, H.-A., Shen, K., Saupe, J., Xia, L., Ahmed, Md. F., Liu, Y.-H., Patil, P. S., Tripathi, A., Hung, S.-C., Taylor, M. E., Lowary, T. L. and Drickamer, K. (2017). **Insights into Interactions of Mycobacteria with the Host Innate Immune System from a Novel Array of Synthetic Mycobacterial Glycans**. ACS Chem Biol *12*, 2990–3002, doi:10.1021/acschembio.7b00797.

8 PERSONAL PUBLICATIONS

8.1 First and last author publications

Broger, T., Koeppel, L., Huerga, H., Miller, P., Gupta-Wright, A., Blanc, F.-X., Esmail, A., Reeve, B. W. P., Floridia, M., Kerkhoff, A. D., Ciccacci, F., Kasaro, M. P., Thit, S. S., Bastard, M., Ferlazzo, G., Yoon, C., Van Hoving, D. J., Sossen, B., García, J. I., Cummings, M. J., Wake, R. M., Hanson, J., Cattamanchi, A., Meintjes, G., Maartens, G., Wood, R., Theron, G., Dheda, K., Olaru, I. D., Denkinger, C. M.. (2023). **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 *11*, e903–e916, doi:10.1016/S2214-109X(23)00135-3.

Broger, T., Marx, F. M., Theron, G., Marais, B. J., Nicol, M. P., Kerkhoff, A. D., Nathavitharana, R., Huerga, H., Gupta-Wright, A., Kohli, M., Nichols, B. E., Muyoyeta, M., Meintjes, G., Ruhwald, M., Peeling, R. W., Pant Pai, N., Pollock, N. R., Pai, M., Cattamanchi, A., Dowdy, D. W., Dewan, P. and Denkinger, C. M. (2024). **Diagnostic yield as an important metric for the evaluation of novel tuberculosis tests – rationale and guidance for future research**. Lancet Glob Health (accepted).

Broger, T., Nicol, M. P., Sigal, G. B., Gotuzzo, E., Zimmer, A. J., Surtie, S., Caceres-Nakiche, T., Mantsoki, A., Reipold, E. I., Székely, R., Tsionsky, M., van Heerden, J., Plisova, T., Chikamatsu, K., Lowary, T. L., Pinter, A., Mitarai, S., Moreau, E., Schumacher, S. G., Denkinger, C. M. (2020b). **Diagnostic accuracy of 3 urine lipoarabinomannan tuberculosis assays in HIV-negative outpatients**. Journal of Clinical Investigation *130*, 5756–5764, doi:10.1172/JCI140461.

Broger, T., Sossen, B., du Toit, E., Kerkhoff, A. D., Schutz, C., Ivanova Reipold, E., Ward, A., Barr, D. A., Macé, A., Trollip, A., Burton, R., Ongarello, S., Pinter, A., Lowary, T. L., Boehme, C., Nicol, M. P., Meintjes, G., Denkinger, C. M. (2019a). **Novel lipoarabinomannan point-of-care tuberculosis test for people with HIV: a diagnostic accuracy study**. Lancet Infect Dis *19*, 852–861, doi:10.1016/S1473-3099(19)30001-5.

Broger, T., Muyoyeta, M., AD, K., CM, D., Moreau, E., Kerkhoff, A. D., Denkinger, C. M. and Moreau, E. (2020a). **Tuberculosis test results using fresh versus biobanked urine samples with FujiLAM**. Lancet Infect Dis *20*, 22–23, doi:10.1016/S1473-3099(19)30684-X.

Broger, T., Nicol, M. P., Székely, R., Bjerrum, S., Sossen, B., Schutz, C., Opintan, J. A., Johansen, I. S., Mitarai, S., Chikamatsu, K., Kerkhoff, A. D., Macé, A., Ongarello, S., Meintjes, G., Denkinger, C. M., Schumacher, S. G. (2020c). **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 *17*, e1003113, doi:10.1371/journal.pmed.1003113.

Broger, T., Tsionksy, M., Mathew, A., Lowary, T. L., Pinter, A., Plisova, T., Bartlett, D., Barbero, S., Denkinger, C. M., Moreau, E., Katsuragi, K., Kawasaki, M., Nahid, P. and Sigal, G. B. (2019b). **Sensitive electrochemiluminescence (ECL) immunoassays for detecting lipoarabinomannan (LAM) and ESAT-6 in urine and serum from tuberculosis patients.** PLoS One *14*, e0215443, doi:10.1371/journal.pone.0215443.

Broger, T. and Nicol, M. P. (2021). **A Novel Bedside Rule-in Test for Tuberculous Meningitis in Human Immunodeficiency Virus–Infected Adults**. Clinical Infectious Diseases *73*, e3435–e3437, doi:10.1093/cid/ciaa1915.

Broger, T., Basu Roy, R., Filomena, A., Greef, C. H., Rimmele, S., Havumaki, J., Danks, D., Schneiderhan-Marra, N., Gray, C. M., Singh, M., Rosenkrands, I., Andersen, P., Husar, G. M., Joos, T. O., Gennaro, M. L., Lochhead, M. J., Denkinger, C. M. and Perkins, M. D. (2017). **Diagnostic Performance of Tuberculosis-Specific IgG Antibody Profiles in Patients with Presumptive Tuberculosis from Two Continents**. Clinical Infectious Diseases *64*, 947–955, doi:10.1093/cid/cix023.

Broger, T., Odermatt, R. P., Huber, P. and Sonnleitner, B. (2011). **Real-time on-line flow cytometry for bioprocess monitoring**. J Biotechnol *154*, 240–247, doi:10.1016/j.jbiotec.2011.05.003.

Broger, T., Odermatt, R. P., Ledergerber, P. and Sonnleitner, B. (2009). **Exploiting Fluorescent Reporter Molecules for Process Analytical Technology (PAT)**. CHIMIA International Journal for Chemistry *63*, 171–173, doi:10.2533/chimia.2009.171.

Bulterys, M. A., Wagner, B., Redard-Jacot, M. M., Suresh, A., Pollock, N. R., Moreau, E., Denkinger, C. M., Drain, P. K. and Broger, T. (2019). **Point-Of-Care Urine LAM Tests for Tuberculosis Diagnosis: A Status Update**. J Clin Med *9*, 111, doi:10.3390/jcm9010111.

Kerkhoff, A. D., Sossen, B., Schutz, C., Reipold, E. I., Trollip, A., Moreau, E., Schumacher, S. G., Burton, R., Ward, A., Nicol, M. P., Meintjes, G., Denkinger, C. M. and Broger, T. (2020). **Diagnostic sensitivity of SILVAMP TB-LAM (FujiLAM) point-of-care urine assay for extra-pulmonary tuberculosis in people living with HIV**. European Respiratory Journal *55*, 1901259, doi:10.1183/13993003.01259-2019.

MacLean, E. and Broger, T. (2017). **A 10-Gene Signature for the Diagnosis and Treatment Monitoring of Active Tuberculosis Using a Molecular Interaction Network Approach**. EBioMedicine *16*, 22–23, doi:10.1016/j.ebiom.2017.01.017.

MacLean, E., Broger, T., Yerlikaya, S., Fernandez-Carballo, B. L., Pai, M. and Denkinger, C. M. (2019). **A systematic review of biomarkers to detect active tuberculosis**. Nat Microbiol *4*, 748–758, doi:10.1038/s41564-019-0380-2.

Sigal, G. B., Pinter, A., Lowary, T. L., Kawasaki, M., Li, A., Mathew, A., Tsionsky, M., Zheng, R. B., Plisova, T., Shen, K., Katsuragi, K., Choudhary, A., Honnen, W. J., Nahid, P., Denkinger, C. M.

and Broger, T. (2018). **A Novel Sensitive Immunoassay Targeting the 5-Methylthio-D-Xylofuranose–Lipoarabinomannan Epitope Meets the WHO's Performance Target for Tuberculosis Diagnosis.** J Clin Microbiol *56*, 1–17, doi:10.1128/JCM.01338-18.

Sossen, B., Broger, T., Kerkhoff, A. D., Schutz, C., Trollip, A., Moreau, E., Schumacher, S. G., Burton, R., Ward, A., Wilkinson, R. J., Barr, D. A., Nicol, M. P., Denkinger, C. M. and Meintjes, G. (2020). **"SILVAMP TB LAM" Rapid Urine Tuberculosis Test Predicts Mortality in Patients Hospitalized With Human Immunodeficiency Virus in South Africa**. Clinical Infectious Diseases *71*, 1973–1976, doi:10.1093/cid/ciaa024.

Yerlikaya, S., Broger, T., MacLean, E., Pai, M. and Denkinger, C. M. (2017). **A tuberculosis biomarker database: the key to novel TB diagnostics**. International Journal of Infectious Diseases *56*, 253–257, doi:10.1016/j.ijid.2017.01.025.

8.2 Middle author publications

Ahmad, R., Xie, L., Pyle, M., Suarez, M. F., Broger, T., Steinberg, D., Ame, S. M., Lucero, M. G., Szucs, M. J., MacMullan, M., Berven, F. S., Dutta, A., Sanvictores, D. M., Tallo, V. L., Bencher, R., Eisinger, D. P., Dhingra, U., Deb, S., Ali, S. M., Mehta, S., Fawzi, W. W., Riley, I. D., Sazawal, S., Premji, Z., Black, R., Murray, C. J. L., Rodriguez, B., Carr, S. A., Walt, D. R. and Gillette, M. A. (2019). **A rapid triage test for active pulmonary tuberculosis in adult patients with persistent cough**. Sci Transl Med *11*, eaaw8287, doi:10.1126/scitranslmed.aaw8287.

Bjerrum, S., Åhsberg, J., Szekely, R., Opintan, J., Lartey, M., Shah, M., Mitarai, S., Broger, T., Ruhwald, M. and Johansen, I. S. (2022). **Diagnostic Accuracy of Urine Lipoarabinomannan Testing in Early Morning Urine versus Spot Urine for Diagnosis of Tuberculosis among People with HIV**. Microbiol Spectr *10*, doi:10.1128/spectrum.00208-22.

Bjerrum, S., Broger, T., Szekely, R., Mitarai, S., Opintan, J. A., Kenu, E., Lartey, M., Addo, K. K., Chikamatsu, K., Mace, A., Schumacher, S. G., Moreau, E., Shah, M., Johansen, I. S., Denkinger, C. M. (2020). **Diagnostic Accuracy of a Novel and Rapid Lipoarabinomannan Test for Diagnosing Tuberculosis Among People With Human Immunodeficiency Virus**. Open Forum Infect Dis *7*, ofz530, doi:10.1093/ofid/ofz530.

Bjerrum, S., Yang, B., Ahsberg, J., Olbrich, L., Weis Damkjaer, M., Nathavitharana, R., Broger, T., Olaru, I. D., Sweetser, B., Poore, H., Razid, A., Kay, A., Denkinger, C. M., Schiller, I., Dendukuri, N., Jaganath, D., Lundh, A. and Shah, M. (2024). **Parallel use of low-complexity automated nucleic acid amplification tests and lateral flow urine lipoarabinomannan assay to detect tuberculosis disease in adults and adolescents with HIV: a systematic review**. Cochrane (in preparation).

Choudhary, A., Patel, D., Honnen, W., Lai, Z., Prattipati, R. S., Zheng, R. B., Hsueh, Y.-C., Gennaro, M. L., Lardizabal, A., Restrepo, B. I., Garcia-Viveros, M., Joe, M., Bai, Y., Shen, K., Sahloul, K., Spencer, J. S., Chatterjee, D., Broger, T., Lowary, T. L. and Pinter, A. (2018). **Characterization**

of the Antigenic Heterogeneity of Lipoarabinomannan, the Major Surface Glycolipid of Mycobacterium tuberculosis , and Complexity of Antibody Specificities toward This Antigen. The Journal of Immunology *200*, 3053–3066, doi:10.4049/jimmunol.1701673.

Dhana, A., Hamada, Y., Kengne, A. P., Kerkhoff, A. D., Broger, T., Denkinger, C. M., Rangaka, M. X., Gupta-Wright, A., Fielding, K., Wood, R., Huerga, H., Rücker, S. C. M., Bjerrum, S., Johansen, I. S., Thit, S. S., Kyi, M. M., Hanson, J., Barr, D. A., Meintjes, G. and Maartens, G. (2022). **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**. Journal of Infection *85*, 40–48, doi:10.1016/j.jinf.2022.05.010.

Drain, P. K., Gardiner, J. L., Hannah, H., Broger, T., Dheda, K., Fielding, K., Walzl, G., Kaforou, M., Kranzer, K., Joosten, S. A., Gilpin, C., Wayer, K., Denkinger, C. M. and Schumacher, S. G. (2019). **Guidance for Studies Evaluating the Accuracy of Biomarker-Based Nonsputum Tests to Diagnose Tuberculosis**. J Infect Dis *220*, S108–S115, doi:10.1093/infdis/jiz356.

Falzon, D., Timimi, H., Kurosinski, P., Migliori, G. B., Van Gemert, W., Broger, T., Denkinger, C., Isaacs, C., Story, A., Garfein, R. S., do Valle Bastos, L. G., Yassin, M. A., Rusovich, V., Skrahina, A., Van Hoi, L., Abubakar, I., Hayward, A., Thomas, B. V., Temesgen, Z., Quraishi, S., von Delft, D., Jaramillo, E., Weyer, K. and Raviglione, M. C. (2016). **Digital health for the End TB Strategy: developing priority products and making them work**. European Respiratory Journal *48*, 29–45, doi:10.1183/13993003.00424-2016.

Fernández-Carballo, B. L., Broger, T., Wyss, R., Banaei, N. and Denkinger, C. M. (2018). **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 *57*, doi:10.1128/JCM.01234-18.

Hammes, F., Broger, T., Weilenmann, H.-U., Vital, M., Helbing, J., Bosshart, U., Huber, P., Peter Odermatt, R. and Sonnleitner, B. (2012). **Development and laboratory-scale testing of a fully automated online flow cytometer for drinking water analysis**. Cytometry Part A *81A*, 508–516, doi:10.1002/cyto.a.22048.

Herrmann, Y., Lainati, F., Castro, M. del M., Mwamba, C. P., Kumwenda, M., Muyoyeta, M., Broger, T., Heinrich, N., Olbrich, L., Corbett, E. L., McMahon, S. A., Engel, N. and Denkinger, C. M. (2022). **User perspectives and preferences on a novel TB LAM diagnostic (Fujifilm SILVAMP TB LAM)–a qualitative study in Malawi and Zambia.** PLOS Global Public Health *2*, e0000672, doi:10.1371/journal.pgph.0000672.

Ibarra-Meneses, A. V., Cruz, I., Chicharro, C., Sánchez, C., Biéler, S., Broger, T., Moreno, J. and Carrillo, E. (2018). **Evaluation of fluorimetry and direct visualization to interpret results of a loop-mediated isothermal amplification kit to detect Leishmania DNA**. Parasit Vectors *11*, 250, doi:10.1186/s13071-018-2836-2.

Koeppel, L., Denkinger, C. M., Wyss, R., Broger, T., Chegou, N. N., Dunty, J. M., Scott, K., Cáceres, T., Dutoit, E., Ugarte-Gil, C., Nicol, M., Gotuzzo, E., Corstjens, P. L. A. M., Geluk, A., Sutherland, J., Sigal, G. B., Moreau, E., Albertini, A., Mantsoki, A., Ongarello, S., Walzl, G. and Fernandez

Suarez, M. (2023). **Diagnostic performance of host protein signatures as a triage test for active pulmonary TB.** J Clin Microbiol *61*, e0026423, doi:10.1128/jcm.00264-23.

Milewski, M. C., Broger, T., Kirkpatrick, J., Filomena, A., Komadina, D., Schneiderhan-Marra, N., Wilmanns, M. and Parret, A. H. A. (2016). **A standardized production pipeline for high profile targets from Mycobacterium tuberculosis**. Proteomics Clin Appl *10*, 1049–1057, doi:10.1002/prca.201600033.

Nathavitharana, R. R., Yoon, C., Macpherson, P., Dowdy, D. W., Cattamanchi, A., Somoskovi, A., Broger, T., Ottenhoff, T. H. M., Arinaminpathy, N., Lonnroth, K., Reither, K., Cobelens, F., Gilpin, C., Denkinger, C. M. and Schumacher, S. G. (2019). **Guidance for Studies Evaluating the Accuracy of Tuberculosis Triage Tests.** J Infect Dis *220*, S116–S125, doi:10.1093/infdis/jiz243.

Nicol, M. P., Schumacher, S. G., Workman, L., Broger, T., Baard, C., Prins, M., Bateman, L., du Toit, E., van Heerden, J., Szekely, R., Zar, H. J. and Denkinger, C. M. (2020). **Accuracy of a novel urine test, Fujifilm SILVAMP TB LAM, for the diagnosis of pulmonary tuberculosis in children**. Clinical Infectious Diseases *26*, 1–15, doi:10.1093/cid/ciaa1052.

Nonyane, B. A. S., Nicol, M. P., Andreas, N. J., Rimmele, S., Schneiderhan-Marra, N., Workman, L. J., Perkins, M. D., Joos, T., Broger, T., Ellner, J. J., Alland, D., Kampmann, B., Dorman, S. E. and Zar, H. J. (2018). **Serologic Responses in Childhood Pulmonary Tuberculosis**. Pediatr Infect Dis J *37*, 1–9, doi:10.1097/INF.0000000000001683.

Reddy, K. P., Denkinger, C. M., Broger, T., McCann, N. C., Gupta-Wright, A., Kerkhoff, A. D., Pei, P. P., Shebl, F. M., Fielding, K. L., Nicol, M. P., Horsburgh, C. R., Meintjes, G., Freedberg, K. A., Wood, R. and Walensky, R. P. (2021). **Cost-effectiveness of a Novel Lipoarabinomannan Test for Tuberculosis in Patients With Human Immunodeficiency Virus**. Clinical Infectious Diseases *73*, e2077–e2085, doi:10.1093/cid/ciaa1698.

Yerlikaya, S., Broger, T., Isaacs, C., Bell, D., Holtgrewe, L., Gupta-Wright, A., Nahid, P., Cattamanchi, A. and Denkinger, C. M. (2023a). **Blazing the trail for innovative tuberculosis diagnostics**. Infection, doi:10.1007/s15010-023-02135-3.

Yerlikaya, S., Holtgrewe, L. M.-L., Broger, T., Isaacs, C., Nahid, P., Cattamanchi, A. and Denkinger, C. M. (2023b). **Innovative COVID-19 point-of-care diagnostics suitable for tuberculosis diagnosis: a scoping review protocol**. BMJ Open *13*, e065194, doi:10.1136/bmjopen-2022-065194.

9 CURRICULUM VITAE

PERSONAL DATA

Name: Tobias Broger
Date of Birth: 11. October 1983
Place of Birth: Appenzell, AI, Switzerland
Nationality: Swiss
Address: Dorfstrasse 46, CH-8102 Oberengstringen, Switzerland
Phone: +41 78 843 21 12, e-mail : tobias.broger@uni-heidelberg.de

SUMMARY

Tobias Broger has a BSc in biological chemistry, a MSc in Engineering, and received business leadership training from IMD Lausanne, Switzerland. Tobias has more than 10 years of experience in *in vitro* diagnostic research, product development, clinical studies and commercialisation. Tobias has written >30 articles in refereed journals and is a holder of patent applications. In 2020, he started a doctorate at the Department of Infectious Diseases and Tropical Medicine in the Heidelberg University Hospital. Tobias Broger's goal is to deliver people-centric diagnostic solutions with global impact.

EDUCATION

04/2020 – present	**Dr. sc hum candidate, Infectious Diseases and Tropical Medicine University, Heidelberg University Hospital,** Heidelberg (DE) ▪ Thesis: Can Lipoarabinomannan Antigen Enable Point-of-Care TB Diagnosis? Exploring the Nexus of Clinical Concentration Ranges, Accuracy, and Diagnostic Yield
10/2017 – 11/2017	**Foundations for Business Leadership (FBL), IMD Business School**, Lausanne (CH) ▪ Strategy, entrepreneurship, marketing, finance, operations, business models
09/2008 – 04/2010	**Master of Science in Engineering**, ZHAW, Wädenswil (CH) ▪ Thesis: „On-line flow cytometry for bioprocess monitoring"
10/2003 – 12/2006	**Bachelor of Science in Biological Chemistry**, ZHAW, Wädenswil (CH) ▪ Learned all modern methods of biology, chemistry and engineering in theory and practice, ▪ Thesis: „High-density cultivation of *Pichia pastoris*"
08/1999 – 07/2002	**Apprenticeship in Analytical Chemistry,** Swiss Federal Laboratories for Materials Science and Technology (EMPA), St. Gallen (CH)

PROFESSIONAL EXPERIENCE

12/2020 – present	**Chief Technology Officer (CTO), Avelo AG**, diagnostics start-up company in Zurich (CH) ▪ Detection of respiratory tract infections with breath aerosols

02/2019 – 11/2020 **Global Product Manager IVD, Proteomedix**, diagnostics company in Zurich (CH)

- Launch of Proclarix IVD assays and software for prostate cancer diagnosis
- IVD product lifecycle management interfacing medical affairs, manufacturing, quality and regulatory, marketing, sales, and product development

04/2013 – 01/2019 **Technical and Senior Technical Officer IVD, FINDdx,** Foundation for Innovative New Diagnostics, Geneva (CH)

- Led development of Fujifilm SILVAMP TB LAM including fundraising $6m, set-up and scientific lead of global consortium, project management, design and conduct of clinical studies in South Africa, Peru and Japan
- Oversaw FIND's global health infectious diseases technology scouting program including identification and assessment of >100 technologies resulting in >15 new projects
- Business development by establishing and following collaborations with global healthcare companies and NGO's
- Contributed to FIND's digital health strategy

01/2007 – 12/2012 **Research & teaching positions in life sciences**

- Teaching biochemical engineering and Matlab at Zurich University of Applied Sciences, Wädenwil (CH)
- Engineer in Environmental Microbiology and Microtechnology, EAWAG, Dübendorf (CH) and Frauenhofer Inst. für Mikrotechn. Mainz (DE): developed an automated liquid handling system for real-time drinking water monitoring
- Scientific assistant Zurich University of Applied Sciences, Dr. Sonnleitner, Wädenswil (CH): acquired and led research projects and prototype development of real-time sensors and cytometry for bioprocess monitoring

Research Areas, Skills and Specialities

- *In vitro* diagnostics (IVD) test development and evaluation
- Tuberculosis
- Lipoarabinomannan (LAM)
- Clinical studies
- Global health
- Exhaled breath aerosols
- Matlab, R, and LabView

Awards

- Swiss TB Award 2020. Swiss Foundation for Tuberculosis Research

10 ACKNOWLEDGEMENTS

I am deeply grateful to my supervisor, Prof. Dr. med. Claudia M Denkinger, for her guidance, encouragement, and support beyond the call of duty before and throughout my doctoral journey. Her expertise, feedback, mentorship, and unwavering dedication have been invaluable in shaping my research and helping me overcome the challenges I faced along the way.

I would also like to thank Drs. Mark D Perkins and Jennifer Lee Gardiner for their guidance during the development of the LAM strategy in 2015 and Dr. Ranald Sutherland for his crucial support during the early and difficult phases of the FujiLAM development project.

I would also like to thank all collaborators, scientific colleagues, and co-authors that worked with me on the important LAM topic. Your expertise and contributions have been critical in ensuring the quality and rigor of my work.

I am grateful to all the study participants whose willingness to contribute their time and effort made this research possible. Your involvement is invaluable, and your commitment to advancing scientific knowledge is greatly appreciated.

I owe a debt of gratitude to my love, Janina Krepart, for her unwavering support and understanding, and to our sunshine, Mija Ylvi Broger. Although Mija Yilvi is too young to understand why I am not driving her around in her little plastic train and instead staring at this bright screen, she brings joy to my life every day.

I also want to acknowledge the tremendous support and understanding of my extended family: Helena Broger, Armin Broger, Ulrike Krepart, Heinrich Krepart, and my sisters Melanie Broger and Anja Broger - I miss you dearly.

Special thanks to Melanie Aregger, my co-founder and partner in crime for her understanding and patience.

Special thanks to Miriam Compton for her meticulous proofreading and invaluable feedback.

Finally I'd like to thank all the donors that invested in LAM work over the years including the Japanese Global Health Innovative Technology Fund, UK Department for International Development, Dutch Ministry of Foreign Affairs, Bill & Melinda Gates Foundation, Australian Department of Foreign Affairs and Trade, the German Federal Ministry of Education and Research through Kreditanstalt für Wiederaufbau, the National Institute of Allergy and Infectious Diseases, and the Foundation for Innovative New Diagnostics.

Thank you all for your contributions and support. This achievement would not have been possible without you.

11 DECLARATION OF GENERATIVE AI IN THE WRITING PROCESS

During the preparation of this work, I used ChatGPT-3.5 for proofreading and shortening sections of text. After utilizing this tool, I carefully reviewed and edited the content as necessary. I take full responsibility for the content of this dissertation. All concepts, strategies, and opinions were developed without the assistance of AI.

12 EIDESSTATTLICHE VERSICHERUNG

1. Bei der eingereichten Dissertation zu dem Thema „Can Lipoarabinomannan Antigen Enable Point-of-Care TB Diagnosis? Exploring the Nexus of Clinical Concentration Ranges, Accuracy, and Diagnostic Yield “ handelt es sich um meine eigenständig erbrachte Leistung.
2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.
3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.
4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.
5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Ort und Datum

Zürich, 30.05.2024

Unterschrift